

**REPUBLIQUE DU CAMEROUN**

**Paix-Travail-Patrie**

-----  
**UNIVERSITE DE YAOUNDE I**

-----  
**FACULTE DES SCIENCES**

-----  
**DEPARTEMENT DE MATHEMATIQUES-**



**REPUBLIC OF CAMEROON**

**Peace-Work-Fatherland**

-----  
**UNIVERSITY OF YAOUNDE I**

-----  
**FACULTY OF SCIENCE**

-----  
**DEPARTMENT OF MATHEMATICS**

## **ETUDE ASYMPTOTIQUE DE L'ANALYSE CANONIQUE GENERALISEE A NOYAU SUR LES DONNEES FONCTIONNELLES**

**Mémoire présenté en vue de l'obtention du diplôme de**

**MASTER en Mathématiques**

**Option : Probabilités, Statistiques et Finance**

**Par : DONFACK Leonel Julio**

**Licencié en Mathématiques**

**Matricule : 15D2495**

**Sous la direction de :**

**Dr. AGHOUKENG Gérard**

**Chargé de Cours**

**Enseignant, Université de Yaoundé I**

**ANNEE ACADEMIQUE : 2021-2022**

# **ETUDE ASYMPTOTIQUE DE L'ANALYSE CANONIQUE GENERALISEE A NOYAU SUR LES DONNEES FONCTIONNELLES**

**Mémoire présenté en vue de l'obtention du diplôme de**

**MASTER en Mathématiques**

**Option : Probabilités, Statistiques et Finance**

**Par : DONFACK Leonel Julio**

**Licencié en Mathématiques**

**Matricule : 15D2495**

**Sous la direction de :**

**Dr. AGHOUKENG Gérard**

**Chargé de Cours**

**Enseignant, Université de Yaoundé**

**19 Septembre 2022**

---

---

## ✠ Dédicace ✠

---

Je dédie ce travail,

À

mon grand frère.

Quoi que je fasse ou quoi que je dise, je ne saurai te remercier comme il se doit. Tu m'as été d'un soutien de tous les instants de mon parcours universitaire. Que ces quelques

mots soient le témoignage de ma gratitude, mon admiration et mon

plus profond respect pour toi. Puisse Dieu te donner

santé, bonheur, courage et surtout succès.

**VOFACK ARNOLD RAPHAEL**

---

---

## ✠ Remerciements ✠

---

Je rends grâce au Dieu tout puissant de m'avoir permis de parachever ce travail dans la paix et la santé.

J'adresse aussi mes vifs et sincères remerciements à tous ceux qui de près comme de loin m'ont aidé et soutenu dans la réalisation de ce travail.

Je pense particulièrement à :

♥ Mon Directeur de mémoire le Dr. **AGHOUKENG Gérard**, pour sa confiance placée en moi en acceptant de diriger ce travail, la pertinence du thème proposé et pour sa disponibilité et son soutien sans faille qu'il m'a accordé pendant la rédaction de ce mémoire. J'ai appris énormément à vos côtés.

♥ Mes **Enseignants de l'Université de Yaoundé 1**, en particulier ceux du département de Mathématiques et d'informatiques pour ma formation académique et sociale. Je pense particulièrement au Pr. **MBELE Bidima Martin**, Pr. **TAKAM Soh Patrice** et au Pr. **TSOPZE Norbert**.

♥ Ma famille pour les soutiens moraux, financiers et matériels qu'elle a eu à faire jusqu'ici pour ma réussite académique. Je pense particulièrement à : mon père **FOUAZEU Thomas**, ma mère **SONNA Philomène**.

♥ Mes amis et connaissances. Je pense particulièrement à : ma meilleure ami **MFEGUE Bana fidèle Linda**

♥ Mes camarades de promotion pour leur formidable esprit d'équipe. Je pense particulièrement à **NDONTSOP Cabrel**, **NGUELE Evrard**, **DONDJIO Rutherford**, **TAKONGMO Muriel**, **DOUGUE Yvan**, **ATEBA Carles**, **LEMOFOUET Ulich**, **MOYO Guilaine**, **YAKAM YAKAM Brice**, **TATSA Josselin**, **TSUATA Wilson**.

Que tous ceux dont les noms ne sont pas mentionnés et qui de quelque manière que ce soit, ont contribué à la réalisation de ce travail, trouvent ici l'expression de ma profonde gratitude.

---

---

## ✠ Déclaration sur l'honneur ✠

---

Le présent mémoire est une œuvre originale du candidat et n'a été soumis nulle part ailleurs, en partie ou en totalité, pour une autre évaluation académique. Les contributions externes ont été dûment mentionnées et recensées en bibliographie.

Signature du candidat

**DONFACK Leonel Julio**

---

---

# ✠ Table des matières ✠

---

<b>Dédicace</b>	<b>i</b>
<b>Remerciements</b>	<b>ii</b>
<b>Déclaration sur l'honneur</b>	<b>iii</b>
<b>Résumé</b>	<b>vi</b>
<b>Abstract</b>	<b>viii</b>
<b>GLOSSAIRE DES NOTATIONS ET ABREVIATIONS</b>	<b>ix</b>
<b>Introduction générale</b>	<b>1</b>
<b>1 Notions Préliminaires</b>	<b>3</b>
1.1 Outils mathématiques . . . . .	3
1.1.1 Rappels . . . . .	3
1.1.2 Noyaux . . . . .	4
1.1.3 Données fonctionnelles et approximation . . . . .	8
1.1.4 Forme quadratique de vecteur gaussien . . . . .	10
1.1.5 Etude asymptotique de l'ACGN et analyse factorielle . . . . .	11
1.2 Analyse canonique linéaire simple . . . . .	11
1.3 Analyse canonique généralisée des V.a euclidiennes . . . . .	13
<b>2 Analyse canonique généralisée à noyau</b>	<b>19</b>
2.1 Analyse canonique généralisée à noyau des V.a hilbertiennes . . . . .	19
2.2 Étude asymptotique de l'analyse canonique généralisée à noyau . . . . .	27

2.2.1	Estimations et convergence presque sûre des operateurs, des vecteurs de directions et des coefficients canoniques. . . . .	28
2.2.2	Distribution asymptotique des opérateurs de covariances empiriques . . . . .	29
<b>3</b>	<b>SIMULATIONS ET APPLICATIONS</b>	<b>38</b>
3.1	Introduction . . . . .	38
3.2	Test de non-corrélation mutuelle multidimensionnelle . . . . .	38
3.2.1	Une statistique de test . . . . .	39
3.2.2	Distribution asymptotique sous l'hypothèse nulle . . . . .	39
3.3	Test de non-corrélation de variables aleatoire hilbertiennes pour $K=2$ . . . . .	42
3.3.1	Hypothèse de test . . . . .	42
3.3.2	Statistique de test . . . . .	43
3.3.3	Loi limite de la statistique de test sous hypothèse nulle . . . . .	44
3.3.4	Propriétés asymptotiques de $\widehat{R}^{(n)}(\alpha)$ . . . . .	44
3.3.5	Loi limite de $\widehat{T}^{(n)}(\alpha)$ sous $\mathcal{H}_0$ . . . . .	48
3.3.6	Construction du test . . . . .	49
3.3.7	Convergence du test . . . . .	51
3.4	APPLICATION AUX DONNEES FONCTIONNELLES . . . . .	52
3.4.1	Application de la méthode . . . . .	53
3.4.2	Cas de données dans l'espace de Sobolev . . . . .	53
3.5	Simulations . . . . .	53
3.5.1	Le modèle utilisé . . . . .	53
3.5.2	Données météorologiques utilisées . . . . .	54
3.6	Approche fonctionnelle . . . . .	60
3.7	Implémentation sur des données réelles . . . . .	63
3.7.1	Test de non-corrélation . . . . .	63
3.7.2	Utilisation de la méthode proposée . . . . .	64
	<b>Conclusion et perspectives</b>	<b>66</b>
	<b>Bibliographie</b>	<b>66</b>
	<b>Bibliographie</b>	<b>67</b>

---

---

## ✠ Résumé ✠

---

Les données fonctionnelles ont beaucoup été ces deux dernières décennies utilisées dans plusieurs domaines, notamment en météorologie. En pratique, ce type de données ne sont pas manipulables facilement et nécessitent le lissage. Ce mémoire s'intéresse à l'analyse canonique généralisée en faisant recours aux méthodes à noyaux. Dans un premier temps, nous présentons les préliminaires qui nous seront utiles dans la suite. Ensuite, nous présentons dans le chapitre 2 l'étude asymptotique de l'analyse canonique généralisée (ACG) à noyau des variables aléatoires euclidiennes. Pour ce faire, nous définissons d'abord, dans la section 2.1, la notion de l'ACG à noyau des variables aléatoires euclidiennes. Cette analyse est définie à partir d'un problème de maximisation sous des contraintes spécifiques, et nous montrons qu'elle est obtenue à partir de l'analyse spectrale d'un opérateur approprié. Dans la section 2.2, nous abordons le problème de l'estimation de l'ACG à noyau. Plus précisément, nous introduisons des estimateurs basés sur des opérateurs de covariance empiriques. Ensuite, la cohérence des estimateurs obtenus est prouvée. En outre, nous dérivons la distribution asymptotique de l'estimateur utilisé de l'opérateur susmentionné, ainsi que celle de l'estimateur du vecteur des coefficients canoniques dans le cas général. La section 2.3 est consacrée aux simulations et applications pour évaluer la performance de nos travaux. Pour cela, nous construisons un test de non-corrélation mutuelle entre les variables aléatoires impliquées dans l'ACG à noyau. Les résultats obtenus pour la théorie asymptotique de l'ACG à noyau sont ensuite utilisés afin de dériver la distribution asymptotique de la statistique de test utilisée sous l'hypothèse nulle. Nous clôturons notre étude dans cette même section destinée aux simulations et applications par des applications qui montrent comment ces variables fonctionnelles s'appliquent sur des fonctions B-splines en considérant l'espace de Sobolev et présenter un cas pratique de test de non-corrélation sur des données de température et précipitation. Nous avons réalisé ces simulations avec le logiciel R.



**Mots clés** : Données fonctionnelles, Etude asymptotique, Analyse canonique généralisée à noyau, test de non-corrélation, **B-Spline**, **KSIR** .

---

---

## ✠ Abstract ✠

---

Functional data has been widely used in the last two decades in several fields, notably in meteorology. In practice, this type of data is not easily manipulated and requires smoothing. This thesis focuses on the canonical analysis generalized using kernel methods. Firstly, we present all the preliminaries which we will use later. Next, we present in chapter 2 the asymptotic study of generalized canonical analysis (GCA) with kernel variables euclidean random. To do this, we first define, in section 2.1, the notion kernel ACG of Euclidean random variables. This analysis is defined from a maximization problem under specific constraints, and we show that it is obtained from the spectral analysis of an appropriate operator. In section 2.2, we discuss the kernel ACG estimation problem. More precisely, we introduce estimators based on empirical covariance operators. Then, the consistency of the estimators obtained is proven. Furthermore, we derive the asymptotic distribution of the estimator used from the aforementioned operator, as well as that of the estimator of the vector of the canonical coefficients in the general case. Section 2.3 is devoted to simulations and applications to evaluate the performance of our work. For this, we build a test of mutual non-correlation between the random variables involved in kernel ACG. The results obtained for the theory distribution of the kernel ACG are then used to derive the asymptotic distribution of the test statistic used under the null hypothesis. We end our study in this same section intended for simulations and applications by applications that show how these functional variables apply to B-spline functions considering Sobolev case of non-correlation test on data from temperature and precipitation. We performed these simulations with the R software

**Keywords :** Functional data, Asymptotic study, Generalized canonical analysis with kernel, non-correlation test, B-Spline, KSIR .

---

---

# ✠ GLOSSAIRE DES NOTATIONS ET ABBREVIATIONS ✠

---

---

Dans cette section, nous présentons les symboles ou les abréviations qui nous aideront à comprendre leur utilisation dans notre mémoire.

$\mathbb{R}$	Ensemble des nombres réels
$\chi$	Espace des observations
$T^\dagger$	inverse généralisée de Moore-Penrose de l'opérateur $T$
$x^T$	transposée de $x$
$rg(A)$	rang de l'opérateur $A$
$\langle \cdot, \cdot \rangle_{\mathcal{H}}$	produit scalaire dans l'espace $\mathcal{H}$
$\  \cdot \ _{\mathcal{H}}$	norme associée au produit scalaire $\langle \cdot, \cdot \rangle_{\mathcal{H}}$
$\otimes^K$	produit de Kronecker
$\otimes$	produit tensoriel
AD	Analyse Discriminante
ACP	Analyse en Composantes Principales
ACL	Analyse Canonique Linéaire
ACG	Analyse Canonique Généralisée
ACGN	Analyse Canonique Généralisée à Noyau

---

---

## ✠ Introduction ✠

---

A l'ère du numérique, la multiplication des systèmes d'informations est à l'origine d'un accroissement constant de la masse de données à traiter. Cette inflation de plus en plus importante révèle les limites des techniques d'exploitation et d'analyse actuelle, et impose la mise en oeuvre d'outils nouveaux, adaptés à cette profusion de données. L'analyse de données fonctionnelles est un domaine de la statistique qui apporte entre autres plusieurs éléments de réponse au problème soulevé ci-dessus. Le livre[10] de Ramsey et Silverman est une introduction de choix dans le domaine. Cet ouvrage montre à suffisance que ce champ de recherche a trouvé un réel écho auprès de la communauté scientifique et statisticienne, suscitant ainsi l'engouement de nombreux experts qui se traduit par la production de nombreux travaux tant théoriques que pratiques.

L'hypothèse sur laquelle repose l'Analyse de Données Fonctionnelles est que les données à traiter possèdent une structure sous-jacente plus ou moins apparente, et que l'identification et la prise en compte explicite de cette structure peut être utilisée afin d'étendre efficacement les techniques de l'analyse de données traditionnelles.

Cette modélisation est particulièrement fructueuse dans le cas où les données présentent une variabilité temporelle : Nous avons par exemple le domaine de la météorologie où on est amené à traiter des teras octets de données par jour ; Ramsay et Silverman[10] font une autre illustration de ce cas en s'intéressant à la croissance d'un groupe d'enfants au cours du temps. Ils montrent que la description de chaque individu par une fonction régulière (la taille de chaque enfant en fonction de l'âge) est particulièrement bien adaptée à la nature du problème. De plus, dans ce cas précis, cette approche fonctionnelle permet de mettre en évidence des caractéristiques non décelables par des techniques traditionnelles : par exemple l'identification des phases de croissance grâce à l'étude de la dérivée de la fonction. C'est un domaine de la Statistique en plein essor comme en témoignent les hors-séries qui lui ont été consacrés dans des revues reconnues comme Statistica Sinica(2004), Computational Statistics Data Analysis (2006), Computational

Statistics (2007), Journal of Multivariate Analysis (2008).

Les méthodes à noyaux occupent ces derniers temps une place importante dans le développement des algorithmes non-linéaire. Ces techniques reposent sur la théorie des noyaux reproduisants permettent de développer une classe d'algorithme d'apprentissage dont la mise en place ne dépend pas de la nature des données traitées, ni de l'espace de représentation adopté pour résoudre les problèmes. L'approche principale consiste à utiliser le coup du noyau, qui permet de transformer les données linéaires en données non linéaires dans un espace à dimension élevée. Ces méthodes à noyaux ont été appliquées à des nombreuses applications et présentant des résultats satisfaisants sans que le coût calculatoire soit énormément ressenti. Parmi lesquels, on trouve dans la littérature l'ACP à noyau et l'AD à noyau, qui sont rien d'autre que les extensions non-linéaires d'algorithmes classiques d'analyse en composantes principales (ACP) et d'analyses discriminantes (AD). La non linéarité s'obtient implicitement par une transformation de l'espace de données vers un espace de représentation de dimension plus élevée. pour une description assez large sur l'ACP et l'ACP à noyaux, on peut consulter [11], [12], [13], [14]. Quand à l'AD et l'AD à noyau , nous pouvons voir [15], [16]

L'objectif de ce mémoire est de proposer un nouveau cadre d'étude de la statistique fonctionnelle, et tente d'introduire une méthode permettant d'effectuer l'ACG à noyau pour variables fonctionnelles, notamment en parlant de la notion de non-correlation. L'essence de cette question est motivée par le fait que plusieurs travaux se soient consacrés à l'Analyse Canonique Généralisée[17] de variables aléatoires hilbertiennes. Il existe très peu de travaux concernant l'étude asymptotique de l'ACG de telles variables avec l'approche du noyau malgré ses multiples avantages, à part les cas évoqué ci haut. Pour remédier à ce manque, une étude asymptotique de l'ACG à noyau est introduit au chapitre 2 . L'une de difficultés montrées lorsque l'on travaille avec de données fonctionnelles est liée à la dimension grande (voire infinie ) de l'espace de travail ; ce problème de dimension élevée de données est corrigé par les méthodes à noyaux . Ce mémoires s'organise de la manière suivante :

- ♣ Dans le premier chapitre intitulé Notions préliminaires , nous présentons dans un premier temps les outils mathématiques utilisés pour la définition de l'ACG à noyau et ensuite nous définissons l'analyse canonique linéaire (ACL) et l'analyse canonique généralisée (ACG) .
- ♣ Dans le deuxième chapitre nous nous intéressons à l'étude asymptotique de l'analyse canonique généralisée (ACG) à noyau des variables aléatoires euclidiennes.
- ♣ Le troisième chapitre est consacré aux simulations et à l'application aux données réelles.

# Notions Préliminaires

---

Toute modélisation mathématiques nécessite l'utilisation des outils mathématiques et des espaces bien précis dans lesquels les variables utilisées prennent leurs valeurs. Dans ce chapitre divisé en trois grandes parties, nous présentons dans la première partie (1.1), les outils mathématiques utilisés pour la réalisation de notre travail. Dans la seconde partie noté (1.2), Nous présentons quelques méthodes de statistique multidimensionnelle et fonctionnelle, plus précisément l'analyse canonique simple (ACL). Enfin, dans la troisième partie notée (1.3), nous présentons l'analyse canonique généralisée ou multiple (ACG).

## 1.1. Outils mathématiques

### 1.1.1. Rappels

**Définition 1.1.1.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Un processus stochastique  $X = (X_t)_{t \geq 0}$  est une famille de variables aléatoires indexées sur le temps telle que :

- pour  $t$  fixé,  

$$X_t : \Omega \longrightarrow \mathbb{R}$$

est une variable aléatoire

$$\omega \longmapsto X_t(\omega)$$
- pour  $\omega$  fixé  

$$X : [0, +\infty[ \longrightarrow \mathbb{R}$$

est une trajectoire

$$t \longmapsto X_t(\omega)$$

**Définition 1.1.2.** Un espace vectoriel  $\mathcal{H}$  muni du produit scalaire  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$  est un espace de Hilbert s'il est complet pour la norme associée au produit scalaire  $\| \cdot \|_{\mathcal{H}} = \sqrt{\langle \cdot, \cdot \rangle_{\mathcal{H}}}$  (autrement dit, toute suite de Cauchy converge dans  $\mathcal{H}$ ).

---

**Définition 1.1.3.** Un espace de Sobolev est un espace vectoriel des fonctions muni de la norme obtenue par la combinaison de la norme  $L^p$  de la fonction elle-même et de ses dérivées jusqu'à un certain ordre.

**Remarque 1.1.1.** Les dérivées sont comprises dans un sens faible, au sens des distributions afin de rendre l'espace complet. Donc les espaces de Sobolev sont des espaces de Banach.

**Définition 1.1.4.**

Soit  $X$  et  $Y$  deux espaces réels normés, une application linéaire de  $X$  vers  $Y$  est appelée opérateur linéaire.

**Définition 1.1.5.**

On appelle variable aléatoire euclidienne définie sur  $(\Omega, A, P)$  et à valeur dans  $(E, B_E)$ , espace euclidien muni de la tribu de ses boréliens, toute application linéaire de  $X$  de  $\Omega$  vers  $E$  telle que , pour tout  $u$  de  $E$  l'application

$$\langle u, X \rangle : \omega \in \Omega \mapsto \langle u, X(\omega) \rangle$$

soit une V.a.r.

## 1.1.2. Noyaux

### 1.1.2.1. Intérêt du noyaux

L'engouement des chercheurs pour cette méthode est qu'elle consiste à transformer tout algorithme linéaire, s'exprimant sous forme de produits scalaires, en un algorithme non linéaire et de plus nécessite peu d'hypothèse sur le type de données, assure les garanties théoriques et permettent la flexibilité.

**Définition 1.1.6. (Noyau)**

Un noyau désigne une fonction  $X \times X$  dans  $\mathbb{R}$  à symétrie hermitienne, c'est à dire telle que

$$K(x_i, x_j) = K(x_j, x_i)$$

La construction de matrice de Gram s'obtient à partir de noyau comme le montre dans la définition 1.1.6.

---

Parmi des noyaux régulièrement utilisés, nous avons :

- ♣ Le noyau linéaire :  $K(x, y) = \langle x; y \rangle_{\mathcal{H}}$ 
  - ♣ Le noyau gaussien :  $K(x, y) = \exp(-\alpha \|x - y\|^2)$
- ♣ Le noyau polynomial :  $K(x, y) = (\langle x; y \rangle_{\mathcal{H}} + \varepsilon)^{\text{degre}}$

**Définition 1.1.7. (Matrice de Gram)**

Soit  $k(., .)$  un noyau et  $n$  observations  $x_1, \dots, x_n \in \mathcal{X}$ , la matrice définie par

$$K_{i,j} = K(x_i, x_j),$$

pour tout  $i, j = 1, 2, \dots, n$  est appelée la matrice de Gram de  $k$  associée à l'ensemble de  $x_1, \dots, x_n \in \mathcal{X}$ . C'est une matrice de dimension  $n \times n$ .

On rappelle les deux définitions fondamentales suivantes et les propriétés qui en découlent.

**Définition 1.1.8. (Noyau défini positif)**

Un noyau est dit positif sur  $\mathcal{X}$  s'il vérifie

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j K(x_i, x_j) \geq 0$$

pour tout  $x_1, \dots, x_n \in \mathcal{X}$  et  $a_1, \dots, a_n \in \mathbb{R}$

Les noyaux définis positifs sont considérés comme une généralisation du produit scalaire.

## 1.1.2.2. Espace de Hilbert à noyau reproduisant (RKHS)

Dépuis les travaux précurseurs de Aronszajn dans [Aro50] sur les noyaux reproduisants, on a de plus en plus recours à ce type d'espaces, notamment depuis qu'ils ont été retenus pour la résolution de problèmes d'interpolation par Parzen [Par70], Kailath [Kai71] et Wahba [Wah90]. Plus récemment, on peut se référer aux travaux de Saito [SAB09]. On présente ici un aperçu des noyaux reproduisants avant d'exposer à la section suivante les éléments clés qui ont contribué à leurs succès dans le cadre de problèmes d'apprentissage. Dans ce qui suit, on considère un espace mesurable  $\mathcal{X}$  auquel on associe le produit scalaire  $\langle ., . \rangle$  et la norme correspondante  $\| . \|$ .

**Définition 1.1.9. (Espace de Hilbert à noyau reproduisant (RKHS))**

Soit  $(\mathcal{H}, \langle ., . \rangle_{\mathcal{H}})$  un espace de Hilbert constitué de fonctions de  $\mathcal{X}$  dans  $\mathbb{C}$ . La fonction  $k(x_i, x_j)$  de  $\mathcal{X} \times \mathcal{X}$  dans  $\mathbb{C}$  est un noyau reproduisant de l'espace  $\mathcal{H}$  sous réserve que celui-ci en admette



un si et seulement si.

- la fonction  $k_{x_i} : x_j \mapsto k_{x_i}(x_j) = k(x_i, x_j)$  appartient à  $\mathcal{H}$  pour tout  $x_i$  fixé.
- On a  $\langle \psi, k_{x_j} \rangle_{\mathcal{H}} = \psi(x_j)$  pour tout  $x_j \in \mathcal{X}$  et  $\psi \in \mathcal{H}$ .

On dit que  $\mathcal{H}$  est un espace de Hilbert à noyau reproduisant, ou encore un RKHS, acronyme de Reproducing Kernel Hilbert Space.

#### ♣ Espace de Sobolev

Soient  $m \in \mathbb{N}$  et  $p \in \mathbb{R}$ ,  $1 \leq p \leq +\infty$ . On définit l'espace de Sobolev  $W^{m,p}(\Omega)$  par

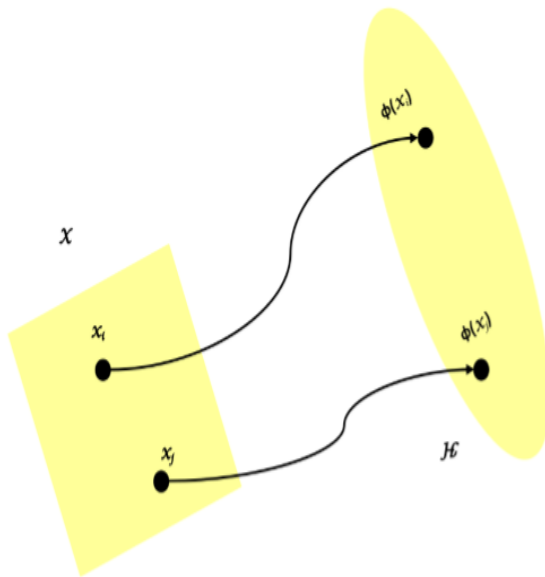
$$W^{m,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega), \forall \alpha \in \mathbb{N}^N, |\alpha| \leq m\}$$

♣ Pour  $p = 2$  et  $\forall m \in \mathbb{N}$ , on a

$$W^{m,2}(\Omega) = H^m(\Omega) := \{u \in L^2(\Omega) : D^\alpha u \in L^2(\Omega), \forall \alpha \in \mathbb{N}^N, |\alpha| \leq m\}$$

muni du produit scalaire défini par  $(u, v)_{H^m} := \sum_{|\alpha| \leq m} \int_{\Omega} D^\alpha u \cdot D^\alpha v dx$ , est un espace de Hilbert.

On désigne par  $k_{x_i} = k(x_i, \cdot)$



où  $\phi$  est une transformation linéaire et continue  $\phi$  de  $\mathcal{X}$  vers un autre espace  $\mathcal{H}$

FIGURE 1.1-Espace des observations  $\mathcal{X}$  et espace de Hilbert à noyau reproduisant  $\mathcal{H}$  induit par le noyau  $k$

#### Corollaire[ cout du noyau]

Tout noyau reproduisant  $k$  issu d'un espace de Hilbert  $\mathcal{H}$ , se traduit comme un produit scalaire dans cet espace comme suit :

$$k(x_i, x_j) = \langle k_{x_i}, k_{x_j} \rangle$$

pour tout  $x_i, x_j \in \mathcal{X}$ .

Ce présent corollaire constitue une propriété fondamentale des noyaux reproduisant et des méthodes à noyau. Sa démonstration découle de la propriété de reproduisant  $f(x_j) = \langle f, k_{x_j} \rangle$  des noyaux, en remplaçant  $f$  dans cette expression par la fonction noyau  $k_{x_i}$ .

### **Théorème 1.1.1. (Moore-Aronszajn(1950))**

A tout noyau défini positif, il correspond un RKHS unique, et réciproquement.

La preuve de ce théorème se trouve dans [18](Paul Honeine 2007).

Le théorème de Moore-Aronszajn établit le lien entre noyau défini positif et RKHS. Par abus de langage, on remplace dans la suite la démonstration de noyau défini positif par noyau reproduisant.

**Corollaire :** Pour tout  $x, x' \in \mathcal{X}$ ,  $|k(x, x')|^2 \leq k(x, x)k(x', x')$

**Démonstration.** Ceci est une conséquence de l'inégalité de Cauchy-Schwartz et de la dernière affirmation du théorème.

## **1.1.2.3. Théorème de représentation**

Le coup du noyau offre une interprétation du noyau reproduisant en tant que produit scalaire, et permet d'élaborer des méthodes non linéaire à partir d'algorithmes linéaires. Pour tenir compte de ce principe, il est très souvent nécessaire de l'associer au Théorème de Représentation. Ce dernier, stipule que les solutions de certaines fonctions de régularisation dans des espaces de dimension élevée ou infinie se trouve dans de sous espaces de dimension finie couvrant par les représentants des données. Il permet de réduire efficacement les problèmes lourds ou infaisables sur le plan informatique dans les espaces de dimension élevée ou infinie à des problèmes d'optimisation sur les coefficients scalaires. Ce résultat fascinant est apparu pour la première fois dans les travaux de [19](Wahba 1990), [20](Kimeldorf and G.Wahba 1971). Les applications multidimensionnelles du Théorème de Représentation ont commencé beaucoup plus tard à la fin des années 1980, avec l'explosion des grandes données complexes et de la puissance de calcul.

### **Théorème 1.1.2.**

Soit un espace non vide  $\mathcal{X}$ , un noyau défini positif  $\kappa$  sur  $\mathcal{X} \times \mathcal{X}$ , un ensemble de données fonctionnelles d'étude  $(x_1, y_1), \dots, (x_n, y_n) \in \mathcal{X} \times \mathbb{R}$ , une fonction réelle  $g$  strictement monotone croissante

sur  $[0, +\infty[$  et une fonction coût arbitraire  $d$ . Soit  $\mathcal{H}$  l'espace de Hilbert associé au noyau  $\kappa$ , avec  $\Phi(x_i) = \kappa(x_i, \cdot)$ . Toute fonction  $f_{\mathcal{H}} \in \mathcal{H}$  minimisant la fonctionnelle de risque régularisée

$$d((x_1, y_1, f_{\mathcal{H}}(x_1), \dots, (x_n, y_n, f_{\mathcal{H}}(x_n))) + g(\|f\|),$$

admet une représentation de la forme

$$f = \sum_{i=1}^n \alpha_i \Phi(x_i).$$

où  $\Phi$  est une transformation linéaire et continue de  $\mathcal{X}$  vers un autre espace  $\mathcal{H}$

La preuve de ce présent théorème se trouve dans [20] (Maya Kallas 2012)

### 1.1.3. Données fonctionnelles et approximation

#### 1.1.3.1. Données fonctionnelles

##### **Définition 1.1.10. (Données fonctionnelles)**

Les données fonctionnelles constituent des observations d'une variable fonctionnelle  $X(t)_{t \in I}$  où  $I = [a, b]$  est un intervalle de  $\mathbb{R}$ , qui n'est autre qu'une variable aléatoire à valeurs dans un espace de Hilbert fonctionnel.

En pratique, on ne dispose, même si la grille de discrétisation est très fine, que d'un nombre fini de valeurs de la réalisation d'une variable fonctionnelle mais pas de cette réalisation fonctionnelle elle-même. Ces données sont constituées d'observations de  $X$  à différents instants. Soit  $(X^{(i)})_{1 \leq i \leq n}$  un échantillon iid de  $X$ , on considère  $t_1, t_2, \dots, t_L$  des points de  $I$  satisfaisant  $a = t_1 < t_2 < \dots < t_L < t_{L+1} = b$  et on note  $X^{(i)}(t_j)$  la variable aléatoire correspondant à la valeur de la  $i$ -ème trajectoire à l'instant  $t_j$ . Nous pouvons donc ranger ces données dans une matrice  $X = X^{(i)}(t_j)_{1 \leq i \leq n, 1 \leq j \leq L}$  qui contiendra toutes les observations de l'échantillon. De plus, il se peut qu'il y ait des valeurs manquantes ou bien que la discrétisation ne soit pas la même d'une données fonctionnelle à l'autre. D'autre part, il est assez fréquent d'être confronté à des courbes déphasées alors que l'on veut appliquer une méthode adaptée à des courbes alignées.

Par exemple, pour passer d'un ensemble de données discrétisées à une données fonctionnelle, l'approche la plus simple et la plus souvent utilisée consiste à lisser les données, c'est à dire à estimer de façon non-paramétrique chaque données fonctionnelle à partir des valeurs dis-

crétiées. Parmi les nombreuses méthodes existantes, les plus couramment utilisées se servent de splines de lissage, d'ondolettes, méthodes des polynômes locaux.

### 1.1.3.2. Approximation par les fonctions splines

#### **Définition 1.1.11.**

Les fonctions splines polynomiales sont des fonctions constituées de morceaux de polynômes se recordant plus ou moins.

On appelle espace de splines polynomiales d'ordre  $m \geq 2$  sur un intervalle  $I = [a, b]$ , en des noeuds simples  $a = x_0 < x_1 < \dots < x_N < x_{N+1} = b$ , l'espace  $S^m(\Delta)$  des fonctions  $s$  sur  $I$  telles qu'il existe  $s_0, s_1, \dots, s_N \in P^{(m)}(I)$  vérifiant  $s|_{[x_i, x_{i+1}]} = s_i$  et  $s_{i-1}^{(j)}(x_i) = s_i^{(j)}(x_i)$  pour tout  $i = 1, \dots, N, j = 0, 1, \dots, m-2$  et  $s \in P^{(m)}(I)$ .

Où  $s|_{[x_i, x_{i+1}]}$  représente la restriction de  $s$  sur l'intervalle  $[x_i, x_{i+1}]$ ,  $f^{(j)}$  désigne la dérivée d'ordre  $j$  de  $f$  et  $P^{(m)}(I)$  est l'espace des polynômes d'ordre  $m$  sur  $I$ .

$S^m(\Delta)$  est un sous espace de dimension  $r = N + L$  de  $H^{m-1}(I)$  où  $H^p(I)$  est l'espace de sobolev des fonctions  $f : I \rightarrow \mathbb{R}$  telle que pour  $0 \leq k \leq p-1$  la dérivée  $k$ -ème  $f^{(k)}$  est continue sur  $I$  et la dérivée  $p$ -ème  $f^{(p)} \in L^2(I)$ . Lorsqu'il est muni du produit scalaire défini par :  $\langle f; g \rangle_{H^p(I)} = \sum_{j=0}^p \int_a^b f^{(j)}(t)g^{(j)}(t)dt$ , l'espace  $H^p(I)$  est un espace de Hilbert. Il existe plusieurs bases de  $S^m(\Delta)$  parmi lesquelles la base B-spline qui est d'ailleurs l'une des bases les plus utilisées en raison des propriétés intéressantes tels que le calcul de ses dérivées successives qui est programmées dans une librairie du logiciel R. Soit  $\tilde{\Delta} = (\delta_i)_{i=1, \dots, m+r}$  avec  $\delta_1 = \delta_2 = \dots = \delta_m = a$ ,  $\delta_{m+i} = x_i$  pour  $i = 1, \dots, N$  et  $\delta_{r+1} = \delta_{r+2} = \dots = \delta_{r+m} = b$ , la partition associé à  $S^m(\Delta)$ . On construit la base B-spline par recurrence en posant :

$$B_j^1(t) = \begin{cases} 1 & \text{si } \delta_j \leq t \leq \delta_{j+1} \\ 0 & \text{sinon} \end{cases}$$

et pour  $s = 2, \dots, m$ ,

$$B_j^s = \frac{t - \delta_j}{\delta_{j+s-1} - \delta_j} B_j^{s-1}(t) + \frac{\delta_{j+s} - t}{\delta_{j+s} - \delta_{j+1}} B_{j+1}^{s-1}(t)$$

Comme indiqué dans schumaker [8], toute fonction  $f$  de  $H^p(I)$  peut être approchée par sa projection sur l'espace engendré par les B-splines d'ordre  $P+1$ . La détermination de cette projection nécessite le calcul des produits scalaires  $\langle B_j^{p+1}; f \rangle_{H^p(I)} = \sum_{k=0}^p \int_a^b \frac{d^k}{dt^k} B_j^{p+1}(t) f^{(k)}(t) dt$ .

---

Lorsque l'expression analytique de la fonction  $f$  n'est pas connue et qu'on ne dispose des valeurs de  $f$  qu'en certains points, ces produits scalaires peuvent être approchés en utilisant les formules classiques d'approximation des dérivées successives  $f^{(k)}$  sur la base des points d'observations, puis une formule de quadrature (par exemple la méthode des trapèzes, rectangle,...) pour le calcul de l'intégrale. Le calcul des fonctions B-splines et de leurs dérivées successives est programmé dans les librairies du logiciel R.

### 1.1.4. Forme quadratique de vecteur gaussien

La détermination de la loi d'une variable aléatoire est un problème réel en statistique surtout lorsqu'on est devant un test statistique. Dans cette partie, nous donnons une approximation de la loi d'une variable aléatoire qui s'exprime sous forme quadratique d'un vecteur gaussien.

Soit  $Q$  une variable aléatoire qui s'exprime sous forme quadratique d'un vecteur gaussien  $M \in \mathbb{R}^p$ , alors  $Q$  s'écrit :

$$Q = M^t A M$$

où  $A$  est une matrice symétrique appelée matrice de la forme quadratique. Nous supposons dans toute la suite que  $M$  est une variable aléatoire centrée. Soit  $\Sigma$  l'opérateur de covariance de  $M$  et posons

$$U = P^t \Sigma^{-1/2} M$$

où  $P$  est la matrice carrée orthogonale de dimension  $p$  qui diagonalise l'opérateur  $\Sigma^{1/2} A \Sigma^{1/2}$ , alors  $Q = \sum_{j=1}^p \lambda_j U_j^2$  où les  $\lambda_j$  sont les valeurs propres de l'opérateur  $\Sigma^{1/2} A \Sigma^{1/2}$  et  $U = (U_1, U_2, \dots, U_p)$ .

Les méthodes existantes pour le calcul de la loi exacte d'une forme quadratique de vecteur gaussien nécessite en général beaucoup de calcul. Plusieurs méthodes d'approximation ont été proposées pour palier à cette difficulté. La méthode d'approximation la plus simple est celle proposée par **Patnaik en 1949** (voir Mathai et Provost (1992)).

Elle consiste à remplacer la loi de  $Q \sim \sum_{j=1}^p \lambda_j U_j^2$  où les  $U_j$  sont des variables aléatoires indépendantes et de loi normale par celle de  $c\chi_\nu^2$  où  $c$  et  $\nu$  sont choisis de telle sorte que les deux premiers moments de  $Q$  et  $c\chi_\nu^2$  soient égaux. Ainsi, on obtient :

$$c = \frac{\left( \sum_{j=1}^p \lambda_j^2 \right)}{\left( \sum_{j=1}^p \lambda_j \right)} \quad \text{et} \quad v = \frac{\left( \sum_{j=1}^p \lambda_j \right)^2}{\left( \sum_{j=1}^p \lambda_j^2 \right)}$$

On peut donc approcher la loi de  $Q$  par

$$\mathbb{P}(Q \leq x) \simeq \mathbb{P}\left(\chi_v^2 < \frac{x}{c}\right)$$

### 1.1.5. Etude asymptotique de l'ACGN et analyse factorielle

**Définition 1.1.12.** L'étude asymptotique de l'ACGN consiste à définir l'analyse théorique puis l'analyse empirique sur un échantillon de taille  $n$  ainsi que leurs opérateurs, enfin d'étudier la convergence des éléments de l'analyse empiriques (fonction des éléments propres de l'opérateur d'échantillonnage) vers les éléments correspondants de l'analyse théoriques lorsque la taille de l'échantillon augmente indéfiniment.

Nous entendons par analyse factorielles les méthodes statistiques donc les résultats sont obtenus à partir de la décomposition spectrale d'un opérateur (en général positif et symétrique) : A.C.P, A.C, A.F.D, A.F.C. Les modèles associés sont à effets fixes ou à effets aléatoires pour lesquels l'estimateur des paramètres est obtenu à partir de la décomposition spectrale d'un opérateur.

Le cadre opératoire présente l'avantage d'être facilement étendu à la dimension infinie et permet d'obtenir les résultats d'une grande généralité.

## 1.2. Analyse canonique linéaire simple

L'analyse canonique est une méthode de statistique multidimensionnelle connue depuis les travaux de Hotelling (1936). Plutard, Dauxois et Pousse (1975) (voir [21]) définissent l'analyse canonique de sous espaces. Cette définition permet de voir l'analyse canonique linéaire comme un cas particulier de l'analyse canonique de sous espaces. L'engouement des chercheurs pour cette méthode est dû au fait qu'elle permet l'étude des liaisons entre groupes de variables. La recherche de la liaison entre variables aléatoires a eu beaucoup d'impact dans les séries temporelles.

Lorsque les variables aléatoires sont réelles, cette liaison entre variables aléatoires s'évalue par le coefficient de corrélation linéaire

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \text{var}(Y)}}$$

Lorsqu'on est en présence des groupes de variables, on procède à une analyse canonique linéaire décrite comme suit :

Considérons  $X = (X_1, X_2, \dots, X_p)$  et  $Y = (Y_1, Y_2, \dots, Y_q)$  deux vecteurs aléatoires de  $(\Omega, \mathcal{A}, \mathbb{P})$  dans  $(\mathbb{R}^p, \mathcal{B}_{\mathbb{R}^p})$  et  $(\mathbb{R}^q, \mathcal{B}_{\mathbb{R}^q})$  respectivement et supposons que :

- (i)  $p < q$
- (ii) pour tout  $i \in \{1, \dots, p\}$ ,  $X_i \in L^2(P)$ ,  $\mathbb{E}(X_i) = 0$
- (iii) pour tout  $j \in \{1, \dots, q\}$ ,  $Y_j \in L^2(P)$ ,  $\mathbb{E}(Y_j) = 0$ .

Soient  $V_1$  (resp.  $V_2$ ) l'opérateur de covariance de  $X$  (resp.  $Y$ ) et  $V_{12}$  l'opérateur de covariance croisé de  $X$  et  $Y$ . Soit  $(\alpha, \beta) \in \mathbb{R}^p \times \mathbb{R}^q$ , on va considérer le coefficient de corrélation des variables aléatoires projetées sur des axes dirigés par  $\alpha$  et  $\beta$  respectivement. On pose :

$$\begin{aligned} \rho(\langle \alpha, X \rangle_{\mathbb{R}^p}, \langle \beta, Y \rangle_{\mathbb{R}^q}) &= \frac{\text{cov}(\langle \alpha, X \rangle_{\mathbb{R}^p}, \langle \beta, Y \rangle_{\mathbb{R}^q})}{\sqrt{\text{var}(\langle \alpha, X \rangle_{\mathbb{R}^p}) \text{var}(\langle \beta, Y \rangle_{\mathbb{R}^q})}} \\ &= \frac{\langle \alpha, V_{12}\beta \rangle_{\mathbb{R}^p}}{\sqrt{\langle \alpha, V_1\alpha \rangle_{\mathbb{R}^p} \langle \beta, V_2\beta \rangle_{\mathbb{R}^q}}} \end{aligned}$$

L'analyse canonique linéaire est donc la recherche de la maximisation de ce coefficient. Alors,

**Définition 1.2.1.** On appelle **analyse canonique linéaire (ACL)** de  $X$  et  $Y$  la recherche du triplet  $(\rho_i, \alpha_i, \beta_i)_{1 \leq i \leq p}$  telle que

$$\rho_i = \langle \alpha_i, V_{12}\beta_i \rangle_{\mathbb{R}^p} = \max\{\langle \alpha, V_{12}\beta \rangle_{\mathbb{R}^p}; (\alpha, \beta) \in \mathbb{R}^p \times \mathbb{R}^q\}$$

avec pour contrainte

$$\langle \alpha, V_1\alpha \rangle_{\mathbb{R}^p} = \langle \beta, V_2\beta \rangle_{\mathbb{R}^q} = 1$$

et pour  $i > 1$ , on a :

$$\langle \alpha_i, V_1\alpha_j \rangle_{\mathbb{R}^p} = \langle \beta_i, V_2\beta_j \rangle_{\mathbb{R}^q} = 0, \quad j < i$$

On montre (voir [6] ) que cette ACL se ramène à l'analyse spectrale de l'un des opérateurs

$$R_1 = \left(V_1^\dagger\right)^{1/2} V_{12} V_2^\dagger V_{21} \left(V_1^\dagger\right)^{1/2} \quad \text{et} \quad R_2 = V_2^\dagger V_{21} V_1^\dagger V_{12} \left(V_2^\dagger\right)^{1/2}$$

Avec  $V_1 = \mathbb{E}(X \otimes X)$ ,  $V_2 = \mathbb{E}(Y \otimes Y)$  et  $V_{12} = \mathbb{E}(X \otimes Y)$

où  $T^\dagger$  est l'**inverse généralisée de Moore-Penrose** de l'opérateur  $T$ .

On appelle **inverse généralisée de Moore-Penrose**, ou **inverse symétrique** d'une matrice  $A$  de type  $m \times n$ , une matrice  $X$  de type  $n \times m$  vérifiant  $AXA = A$  et  $XAX = X$  simultanément.

Si  $A$  est inversible, inversible à gauche, inversible à droite, alors ces inverses sont des inverses généralisés.

On note l'ensemble des inverses généralisés par  $A(1, 2)$

Ainsi, lorsque  $\rho_i = 0$ , pour tout  $i \in \{1, \dots, p\}$ , on conclut que  $X$  et  $Y$  sont **non corrélées**. En d'autres termes, il n'existe aucun lien entre  $X$  et  $Y$ . Ceci a permis de construire les tests de non-corrélation entre  $X$  et  $Y$  (voir par exemple Dauxois et Nkiet (1997a) dans [6]).

Le test de l'hypothèse nulle définie par :

$$\mathcal{H}_0 : \rho_i = 0, \quad \text{pour tout } i \in \{1, \dots, p\}$$

se ramène à tester l'hypothèse nulle

$$\mathcal{H}_0 : \sum_{i=1}^p \rho_i^2 = 0 \iff \text{tr}(R_1) = 0$$

### 1.3. Analyse canonique généralisée des V.a euclidiennes

L'analyse canonique généralisée (ACG) est une méthode statistique multidimensionnelle connue depuis les travaux de Kitterring (1971), Gifi (1991), Gardner et al. (2006), Takane et al. (2008), Tenenhaus et Tenenhaus (2011), ainsi que les autres références qui y figurent tel que ceux de Sumcor (1961) et Carroll (1968). Plutard, Guy Martial Nkiet (23 septembre 2017) définit l'ACG des variables aléatoires euclidiennes. Cette définition permet de voir l'analyse canonique généralisée comme un cas général de l'analyse canonique simple, mais aussi de l'ACP, l'AFD et l'ACM, la régression simple, et multiple, l'analyse de la covariance et l'analyse discriminante. Il s'agit d'une méthode statistique qui généralise l'analyse canonique linéaire (ACL) au cas où plus de deux ensembles de variables sont considérées. L'engouement des chercheurs pour cette méthode est dû au fait qu'elle permet de décrire et de résumer les liaisons linéaires existantes



entre plusieurs groupes de variables mesurées sur les même individus.

Pour un entier  $K \geq 2$ , considérons des variables aléatoires  $X_1, \dots, X_K$  définies sur un espace probabilisé  $(\Omega, \mathcal{A}, P)$  et évaluées dans des espaces vectoriels euclidiens  $\mathcal{X}_1, \dots, \mathcal{X}_K$  respectivement. En désignant par  $\mathbb{E}$  l'espérance mathématique liée à  $P$ , nous supposons que, pour tout  $k \in \{1, \dots, K\}$ , nous avons  $\mathbb{E}(\|X_k\|_k^2) < +\infty$  où  $\|\cdot\|_k$  indique la norme induite par le produit interne  $\langle \cdot, \cdot \rangle_k$  de  $\mathcal{X}_k$  et sans perte de généralité, que  $\mathbb{E}(X_k) = 0$ . Chaque vecteur  $\alpha$  dans l'espace vectoriel  $\mathcal{X} : \mathcal{X}_1 \times \dots \times \mathcal{X}_K$  est définie par

$$\alpha = \begin{pmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_K \end{pmatrix},$$

et nous rappelons que  $\mathcal{X}$  est un espace vectoriel euclidien équipé du produit interne  $\langle \cdot, \cdot \rangle_{\mathcal{X}}$  défini par :

$$\forall (\alpha, \beta) \in \mathcal{X} \times \mathcal{X}, \langle \alpha, \beta \rangle_{\mathcal{X}} = \sum_{k=1}^K \langle \alpha_k, \beta_k \rangle_k.$$

Nous désignons par  $\|\cdot\|_{\mathcal{X}}$  la norme induite par ce produit interne. En considérant la V.a à valeur dans  $\mathcal{X}$

$$X = \begin{pmatrix} X_1 \\ \cdot \\ \cdot \\ \cdot \\ X_K \end{pmatrix},$$

nous pouvons donner la définition suivante, qui adapte la définition classique de l'analyse canonique à ensembles multiples (par exemple, Gifi(1991), Gardner et al. (2006), Takane et al.(2008)) au contexte des V.a euclidiennes.

**Définition 1.3.1. (Analyse canonique linéaire à ensemble multiples))**

L'analyse canonique linéaire à ensembles multiples (ACLG) de  $X$  est la recherche d'une séquence  $(\alpha^{(j)})_{1 \leq j \leq q}$  de vecteurs de  $E$ , où  $q = \dim(\mathcal{X})$ , satisfaisant :

$$\alpha^{(j)} = \underset{\alpha \in C_j}{\operatorname{argmax}} \mathbb{E}(\langle \alpha, X \rangle_{\mathcal{X}}^2), \quad (1)$$

où

$$C_1 = \{\alpha \in \mathcal{X} / \sum_{k=1}^K \text{var}(\langle \alpha_k, X_k \rangle_k) = 1\}, \quad (2)$$

et, pour  $j \geq 2$

$$C_j = \{\alpha \in C_1 / \sum_{k=1}^K \text{cov}(\langle \alpha_k^{(r)}, X_k \rangle_k, \langle \alpha_k, X_k \rangle_k) = 0, \forall r \in \{1, \dots, j-1\}\}, \quad (3)$$

Avec

$$\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_K \end{pmatrix} \quad \text{et} \quad X = \begin{pmatrix} X_1 \\ \vdots \\ X_K \end{pmatrix}$$

**Remarque 1.3.1.** L'ensembles de contraintes donnés dans (2) et (3) peuvent être exprimés en utilisant des opérateurs de covariance définis pour

1)  $\forall (k, l) \in \{1, \dots, K\}^2$  par :

$$V_{kl} = E(X_l \otimes X_k) = V_{lk}^* \text{ et } V_k = V_{kk},$$

où  $\otimes$  désigne le produit tensoriel tel que, pour tout  $(x, y)$ ,  $x \otimes y$  est l'application linéaire :  $h \mapsto \langle x, h \rangle y$ , et  $T^*$  désigne l'adjoint de  $T$ . En effet, il est facile de voir que, pour  $(\alpha, \beta) \in \mathcal{X}^2$ ,

$$\begin{aligned} \text{var}(\langle \alpha_k, X_k \rangle_k) &= E(\langle \alpha_k, X_k \rangle_k^2) \text{ car } E(\langle \alpha_k, X_k \rangle_k) = 0 \quad \forall k \\ &= E(\langle \alpha_k, (X_k \otimes X_k)(\alpha_k) \rangle_k) \\ &= \langle \alpha_k, V_k \alpha_k \rangle_k \end{aligned}$$

et

$$\begin{aligned} \text{cov}(\langle \alpha_k, X_k \rangle_k, \langle \beta_l, X_l \rangle_l) &= E(\langle \alpha_k, X_k \rangle_k \langle \beta_l, X_l \rangle_l) \\ &= E(\langle \alpha_k, (X_l \otimes X_k)(\beta_l) \rangle_k) \\ &= \langle \alpha_k, V_{kl} \beta_l \rangle_k. \end{aligned}$$

Donc

$$C_1 = \{\alpha \in \mathcal{X} / \sum_{k=1}^K \langle \alpha_k, V_k \alpha_k \rangle_k = 1\}, \quad (4)$$

et

$$C_j = \{\alpha \in C_1 / \sum_{k=1}^K < \alpha_k^{(r)}, V_k \alpha_k >_k = 0, \forall r \in \{1, \dots, j-1\}\}. \quad (5)$$

2) Pour tout  $\alpha \in C_1$ , on a :

$$\begin{aligned} E(< \alpha, X >_{\mathcal{X}}^2) &= E\left(\sum_{k=1}^K < \alpha_k, X_k >_k\right)^2 \\ &= \sum_{k=1}^K \sum_{l=1}^K E(< \alpha_k, X_k >_k < \alpha_l, X_l >_l) \\ &= \sum_{k=1}^K E(< \alpha_k, X_k >_k^2) + \sum_{k=1}^K \sum_{l=1, l \neq k}^K E(< \alpha_k, X_k >_k < \alpha_l, X_l >_l) \\ &= \sum_{k=1}^K \text{var}(< \alpha_k, X_k >_k) + \sum_{k=1}^K \sum_{l=1, l \neq k}^K < \alpha_k, V_{kl} \alpha_l >_k \\ &= 1 + \sum_{k=1}^K \sum_{l=1, l \neq k}^K < \alpha_k, V_{kl} \alpha_l >_k \\ &= 1 + \varphi(\alpha) \end{aligned}$$

où

$$\varphi(\alpha) = \sum_{k=1}^K \sum_{l=1, l \neq k}^K < \alpha_k, V_{kl} \alpha_l >_k \quad (6)$$

Ensuite, l'ACLG de  $X$  est obtenu en minimisant  $\varphi(\alpha)$  sous les contraintes exprimées dans (4) et (5).

Pour  $k \in \{1, \dots, K\}$ , l'opérateur de covariance  $V_k$  est un opérateur auto-adjoint non-négatif. A partir de maintenant, nous supposons qu'il est inversible. Soit  $\tau_k$  la projection canonique définie comme suit :

$$\tau_k : \alpha \in \mathcal{X} \mapsto \alpha_k \in \mathcal{X}_k :$$

son adjoint  $\tau_k^*$  de  $\tau_k$  est la carte donnée par :

$$\tau_k^* : t \in \mathcal{X}_k \mapsto (\underbrace{0, \dots, 0}_{k-1 \text{ fois}}, t, 0, \dots, 0)^T \in \mathcal{X},$$

où nous désignons par  $a^T$  la transposée de  $a$ . Considérons maintenant les opérateurs de  $\mathcal{L}(\mathcal{X})$  donnés par :

$$\Phi = \sum_{k=1}^K \tau_k^* V_k \tau_k \quad \text{et} \quad \Psi = \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* V_{kl} \tau_l.$$

Du fait que  $\tau_k \tau_l^* = \delta_{kl} I_k$ , où  $\delta_{kl}$  est le symbole de Kronecker habituel et  $I_k$  est l'opérateur d'identité de  $\mathcal{X}_k$ , on voit facilement que  $\Phi$  est aussi un opérateur auto-adjoint inversible et non négatif, avec  $\Phi^{-1} = \sum_{k=1}^K \tau_k^* V_k^{-1} \tau_k$  et  $\Phi^{-\frac{1}{2}} = \sum_{k=1}^K \tau_k^* V_k^{-\frac{1}{2}} \tau_k$ . Le théorème suivant montre comment obtenir

une ACLG de X.

### **Théorème 1.3.1.**

Soit  $\{\beta^{(1)}, \dots, \beta^{(q)}\}$  une base orthonormale de  $\mathcal{X}$  telle que  $\beta^j$  soit un vecteur propre de l'opérateur  $T = \Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}}$  associé à la j-ième plus grande valeur propre  $\rho_j$  de T. Alors, la séquence  $(\alpha^j)_{1 \leq j \leq q}$  donnée par :

$$\alpha^{(j)} = \Phi^{-\frac{1}{2}} \beta^{(j)} = (V_1^{-\frac{1}{2}} \beta_1^{(j)}, \dots, V_K^{-\frac{1}{2}} \beta_K^{(j)}),$$

est constitué des solutions de (1) sous les contraintes (2) et (3), et nous avons :

$$\rho_j = \langle \beta^{(j)}, T \beta^{(j)} \rangle_E = \rho(\alpha^{(j)}).$$

**Preuve.** En mettant  $\beta_k = V_k^{\frac{1}{2}} \alpha_k$  et  $\beta_k^{(r)} = V_k^{\frac{1}{2}} \alpha_k^{(r)}$ , on a :

$$\begin{aligned} \rho(\alpha) &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle V_k^{-\frac{1}{2}} \beta_k, V_{kl} V_l^{-\frac{1}{2}} \beta_l \rangle_k \\ &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle \beta_k, V_k^{-\frac{1}{2}} V_{kl} V_l^{-\frac{1}{2}} \beta_l \rangle_k \end{aligned}$$

(7)

Où

$$\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_K \end{pmatrix} \in \mathcal{X}.$$

Puisque  $V_k = V_k^{\frac{1}{2}} V_k^{\frac{1}{2}}$ , avoir  $\alpha \in C_j$  est équivalent à avoir  $\beta \in C'_j$ , où :

$$C'_1 = \{\beta \in \mathcal{X} / \sum_{k=1}^K \|\beta\|_k^2 = 1\} = \{\beta \in \mathcal{X} / \|\beta\|_k^2 = 1\}, \quad (8)$$

Et pour  $j \geq 2$  :

$$C'_j = \{\beta \in C_1 / \sum_{k=1}^K \langle \beta_k^{(r)}, \beta_k \rangle_k = 0, \forall r \in \{1, \dots, j-1\}\} = \{\beta \in C_1 / \langle \beta^{(r)}, \beta \rangle_{\mathcal{X}} = 0, \forall r \in \{1, \dots, j-1\}\}. \quad (9)$$

De plus pour tout  $\beta \in \mathcal{X}$ , nous avons :

$$\Psi \Phi^{-\frac{1}{2}} \beta = \sum_{k=1}^K \sum_{l=1, l \neq k}^K \sum_{j=1}^K \tau_k^* V_{kl} \tau_l \tau_j^* V_j^{-\frac{1}{2}} \tau_j \beta$$

$$\begin{aligned}
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \sum_{j=1}^K \delta_{lj} \tau_k^* V_{kl} V_j^{-\frac{1}{2}} \tau_j \beta \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* V_{kl} V_l^{-\frac{1}{2}} \tau_l \beta,
\end{aligned}$$

Et

$$\begin{aligned}
\Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}} \beta &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \sum_{j=1}^K \tau_j^* V_j^{-\frac{1}{2}} \tau_j \tau_k^* V_{kl} V_l^{-\frac{1}{2}} \tau_l \beta \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \sum_{j=1}^K \delta_{jk} \tau_j^* V_j^{-\frac{1}{2}} V_{kl} V_l^{-\frac{1}{2}} \tau_l \beta \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* V_k^{-\frac{1}{2}} V_{kl} V_l^{-\frac{1}{2}} \tau_l \beta \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* V_k^{-\frac{1}{2}} V_{kl} V_l^{-\frac{1}{2}} \beta_l
\end{aligned}$$

Ainsi

$$\begin{aligned}
\langle \beta, \Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}} \beta \rangle_X &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle \beta, \tau_k^* V_k^{-\frac{1}{2}} V_{kl} V_l^{-\frac{1}{2}} \beta_l \rangle_X \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle \tau_k \beta, V_k^{-\frac{1}{2}} V_{kl} V_l^{-\frac{1}{2}} \beta_l \rangle_k \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle \beta_k, V_k^{-\frac{1}{2}} V_{kl} V_l^{-\frac{1}{2}} \beta_l \rangle_k \\
&= \psi(\beta),
\end{aligned}$$

■

Où  $\psi$  est défini dans (7). Alors, le problème d'optimisation ACLG se réduit à la maximisation de  $\langle \beta, \Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}} \beta \rangle_X$  sous la contrainte (8) et (9). Puisque  $T = \Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}}$  est un opérateur auto-adjoint, il s'agit d'un problème de maximisation bien connu pour lequel une solution est obtenue à partir de l'analyse spectrale de T, comme indiqué dans le théorème.

# Analyse canonique généralisée à noyau

L'analyse canonique généralisée à noyau (ACGN), parfois aussi nommé analyse à noyaux des corrélations canoniques, est l'analyse canonique ordinaire grâce à l'astuce du noyau. Il s'agit d'une méthode statistique qui généralise l'analyse canonique linéaire (ACL) au cas où plus de deux ensembles de variables sont considérées en utilisant l'astuce du noyau. L'engouement des chercheurs pour cette méthode est qu'elle consiste à transformer tout algorithme linéaire, s'exprimant sous forme de produit scalaire, en un algorithme non linéaire et de plus nécessite peu d'hypothèse sur le type de données, assure les garanties théoriques et permettent la flexibilité.

## 2.1. Analyse canonique généralisée à noyau des V.a hilbertiennes

Pour un entier  $K \geq 2$ , considérons des variables aléatoires  $X_1, \dots, X_K$  définies sur un espace probabilisé  $(\Omega, \mathcal{A}, P)$  et évaluées dans des espaces vectoriels hilbertiens  $\mathcal{X}_1, \dots, \mathcal{X}_K$  respectivement. En désignant par  $\mathbb{E}$  l'espérance mathématiques liée à  $P$ , nous supposons que , pour tout  $k \in \{1, \dots, K\}$ , nous avons :

$$\mathbb{E}(\|X_k\|_k^2) < +\infty$$

où  $\|\cdot\|_k$  indique la norme induite par le produit interne  $\langle \cdot, \cdot \rangle_k$  de  $\mathcal{X}_k$  et sans perte de généralité, nous supposons que  $\mathbb{E}(X_k) = 0$ . Chaque vecteur  $\alpha$  dans l'espace vectoriel  $\mathcal{X} : \mathcal{X}_1 \times \dots \times \mathcal{X}_K$  est défini par

$$\alpha = \begin{pmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_K \end{pmatrix},$$

et nous rappelons que  $\mathcal{X}$  est un espace vectoriel hilbertien équipé du produit interne  $\langle \cdot, \cdot \rangle_{\mathcal{X}}$  défini par :

$$\forall (\alpha, \beta) \in \mathcal{X} \times \mathcal{X}, \langle \alpha, \beta \rangle_{\mathcal{X}} = \sum_{k=1}^K \langle \alpha_k, \beta_k \rangle_k.$$

Nous désignons par  $\| \cdot \|_{\mathcal{X}}$  la norme induit par ce produit interne. En considérant la V.a à valeur dans  $\mathcal{X}$

$$X = \begin{pmatrix} X_1 \\ \cdot \\ \cdot \\ \cdot \\ X_K \end{pmatrix},$$

nous pouvons donner la définition suivante, qui adapte la définition classique de l'analyse canonique généralisée à noyau (par exemple, Gifi(1991), Gardner et al. (2006), Takane et al.(2008), Nkiet(23 septembre 2017) au contexte des V.a hilbertiennes.

Nous considérons une transformation linéaire et continue  $\phi$  de  $\mathcal{X}$  vers un autre espace  $\mathcal{H}$ , et  $k$  le noyau reproduisant associé à cette transformation. L'utilisation de l'approche ACG linéaire dans le cadre d'un problème de discrétisation non-linéaire, consiste à la mise en oeuvre de l'ACG dans l'espace  $\mathcal{H}$  sur les données transformées  $\phi(\cdot)$ .

### **Définition 2.1.1. (analyse canonique généralisée à noyau)**

L'analyse canonique généralisée à noyau (ACGN) de  $X$  est la recherche d'une séquence  $(\alpha^{(j)})_{1 \leq j \leq q}$  de vecteurs de  $E$ , où  $q = \dim(\mathcal{X})$ , satisfaisant :

$$\alpha^{(j)} = \underset{\alpha \in C_j}{\operatorname{argmax}} E(\langle \alpha, \phi(X) \rangle_{\mathcal{H}}^2), \quad (1)$$

où

$$C_1 = \{ \alpha \in \mathcal{H} / \sum_{k=1}^K \operatorname{var}(\langle \alpha_k, \phi(X_k) \rangle_k) = 1 \}, \quad (2)$$

et, pour  $j \geq 2$

$$C_j = \{\alpha \in C_1 / \sum_{k=1}^K cov(< \alpha_k^{(r)}, \phi(X_k) >_k, < \alpha_k, \phi(X_k) >_k) = 0, \forall r \in \{1, \dots, j-1\}\}, \quad (3)$$

Avec

$$\alpha = \begin{pmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_K \end{pmatrix} \quad \text{et} \quad X = \begin{pmatrix} X_1 \\ \cdot \\ \cdot \\ \cdot \\ X_K \end{pmatrix}$$

**Remarque 2.1.1.** La définition précédente peut être défini de manière similaire que l'analyse canonique généralisée (ACG) des variables aléatoires hilbertiennes. Rappelons que L'analyse canonique linéaire généralisée (ACLG) de  $X$  est la recherche d'une séquence  $(\alpha^{(j)})_{1 \leq j \leq q}$  de vecteurs de  $E$ , où  $q = \dim(X)$ , satisfaisant :

$$\alpha^{(j)} = \underset{\alpha \in C_j}{argmax} \mathbb{E}(< \alpha, X >_{\mathcal{X}}^2), \quad (a)$$

où

$$C_1 = \{\alpha \in \mathcal{X} / \sum_{k=1}^K var(< \alpha_k, X_k >_k) = 1\}, \quad (b)$$

et, pour  $j \geq 2$

$$C_j = \{\alpha \in C_1 / \sum_{k=1}^K cov(< \alpha_k^{(r)}, X_k >_k, < \alpha_k, X_k >_k) = 0, \forall r \in \{1, \dots, j-1\}\}, \quad (c)$$

Avec

$$\alpha = \begin{pmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_K \end{pmatrix} \quad \text{et} \quad X = \begin{pmatrix} X_1 \\ \cdot \\ \cdot \\ \cdot \\ X_K \end{pmatrix}$$

**Remarque 2.1.2.** 1) Les ensembles de contraintes donnés dans (2) et (3) peuvent être exprimés en utilisant des opérateurs de covariance définis pour  $(k, l) \in \{1, \dots, K\}^2$  par :

$$V_{kl}^{\phi} = E(\phi(X_l) \otimes \phi(X_k)) = V_{lk}^{\phi*} \text{ et } V_k^{\phi} = V_{kk}^{\phi},$$



où  $\otimes$  désigne le produit tensoriel défini par, pour tout  $(x, y)$ ,  $x \otimes y$  est l'application linéaire :  $h \mapsto \langle x, h \rangle y$ , et  $T^*$  désigne l'adjoint de  $T$ . En effet, il est facile de voir que, pour  $(\alpha, \beta) \in \mathcal{X}^2$ ,

$$\begin{aligned} \text{var}(\langle \alpha_k, \phi(X_k) \rangle_k) &= E(\langle \alpha_k, \phi(X_k) \rangle_k^2) \\ &= E(\langle \alpha_k, \phi(X_k) \rangle_k \langle \alpha_k, \phi(X_k) \rangle_k) \\ &= E(\langle \alpha_k, (\phi(X_k) \otimes \phi(X_k)) \alpha_k \rangle_k) \\ &= \langle \alpha_k, V_k^\phi \alpha_k \rangle_k \end{aligned}$$

et

$$\begin{aligned} \text{cov}(\langle \alpha_k, \phi(X_k) \rangle_k, \langle \beta_l, \phi(X_l) \rangle_l) &= E(\langle \alpha_k, \phi(X_k) \rangle_k \langle \beta_l, \phi(X_l) \rangle_l) \\ &= E(\langle \alpha_k, (\phi(X_l) \otimes \phi(X_k)) \beta_l \rangle_k) \\ &= \langle \alpha_k, V_{kl}^\phi \beta_l \rangle_k. \end{aligned}$$

Donc

$$C_1 = \{\alpha \in \mathcal{H} / \sum_{k=1}^K \langle \alpha_k, V_k^\phi \alpha_k \rangle_k = 1\}, \quad (4)$$

et

$$C_j = \{\alpha \in C_1 / \sum_{k=1}^K \langle \alpha_k^{(r)}, V_k^\phi \alpha_k \rangle_k = 0, \forall r \in \{1, \dots, j-1\}\}. \quad (5)$$

2) Pour tout  $\alpha \in C_1$ , on a :

$$\begin{aligned} E(\langle \alpha, \phi(X) \rangle_{\mathcal{X}}^2) &= E\left(\sum_{k=1}^K \langle \alpha_k, \phi(X_k) \rangle_k\right)^2 \\ &= \sum_{k=1}^K \sum_{l=1}^K E(\langle \alpha_k, \phi(X_k) \rangle_k \langle \alpha_l, \phi(X_l) \rangle_l) \\ &= \sum_{k=1}^K E(\langle \alpha_k, \phi(X_k) \rangle_k^2) + \sum_{k=1}^K \sum_{l=1, l \neq k}^K E(\langle \alpha_k, \phi(X_k) \rangle_k \langle \alpha_l, \phi(X_l) \rangle_l) \\ &= \sum_{k=1}^K \text{var}(\langle \alpha_k, \phi(X_k) \rangle_k) + \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle \alpha_k, V_{kl}^\phi \alpha_l \rangle_k \\ &= 1 + \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle \alpha_k, V_{kl}^\phi \alpha_l \rangle_k \\ &= 1 + \varphi(\phi(\alpha)), \end{aligned}$$

où

$$\varphi(\phi(\alpha)) = \sum_{k=1}^K \sum_{l=1, l \neq k}^K < \alpha_k, V_{kl}^\phi \alpha_l >_k \quad (6)$$

Ensuite, l'ACG à noyau de X est obtenu en minimisant  $\varphi(\phi(\alpha))$  sous les contraintes exprimées dans (4) et (5).

Pour  $k \in \{1, \dots, K\}$ , l'opérateur de covariance  $V_k^\phi$  est un opérateur auto-adjoint non-négatif. A partir de maintenant, nous supposons qu'il est inversible. Soit  $\tau_k$  la projection canonique définie comme suit :

$$\tau_k : \alpha \in \mathcal{X} \mapsto \alpha_k \in \mathcal{X}_k$$

son adjoint  $\tau_k^*$  de  $\tau_k$  est la carte donnée par :

$$\tau_k^* : t \in \mathcal{X}_k \mapsto (\underbrace{0, \dots, 0}_{k-1 \text{ fois}}, t, 0, \dots, 0)^T \in \mathcal{X},$$

où nous désignons par  $a^T$  la transposée de a. Considérons maintenant les opérateurs de  $\mathcal{L}(\mathcal{X})$  donnés par :

$$\Phi = \sum_{k=1}^K \tau_k^* V_k^\phi \tau_k \quad \text{et} \quad \Psi = \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* V_{kl}^\phi \tau_l.$$

Du fait que  $\tau_k \tau_l^* = \delta_{kl} I_k$ , où  $\delta_{kl}$  est le symbole de Kronecker habituel et  $I_k$  est l'opérateur d'identité de  $\mathcal{X}_k$ , on voit de même que  $\Phi$  est aussi un opérateur auto-adjoint inversible et non négative, avec  $\Phi^{-1} = \sum_{k=1}^K \tau_k^* (V_k^\phi)^{-1} \tau_k$  et  $\Phi^{-\frac{1}{2}} = \sum_{k=1}^K \tau_k^* (V_k^\phi)^{-\frac{1}{2}} \tau_k$ . Le théorème suivant montre comment obtenir une ACG à noyau de X. Il ne fait que répéter un résultat connu (par exemple, Gifi (1991), Takane et al. (2008), Nkiet 23 septembre 2017) dans le cadre de ce chapitre.

### **Théorème 2.1.1.**

Soit  $\{\beta^{(1)}, \dots, \beta^{(q)}\}$  une base orthonormale de  $\mathcal{X}$  telle que  $\beta^{(j)}$  soit un vecteur propre de l'opérateur  $T = \Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}}$  associé à la j-ième plus grande valeur propre  $\rho_j$  de T. Alors, la séquence  $(\alpha^{(j)})_{1 \leq j \leq q}$  donnée par :

$$\alpha^{(j)} = \Phi^{-\frac{1}{2}} \beta^{(j)} = ((V_1^\phi)^{-\frac{1}{2}} \beta_1^{(j)}, \dots, (V_K^\phi)^{-\frac{1}{2}} \beta_K^{(j)}),$$

est constitué des solutions de (1) sous les contraintes (2) et (3), et nous avons :

$$\rho_j = \langle \beta^{(j)}, T \beta^{(j)} \rangle_E = \rho(\alpha^{(j)}).$$

**Preuve.** En mettant  $\beta_k = (V_k^\phi)^{\frac{1}{2}} \alpha_k$  et  $\beta_k^{(r)} = (V_k^\phi)^{\frac{1}{2}} \alpha_k^{(r)}$ , on a :

$$\begin{aligned}
\varphi(\phi(\alpha)) &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle (V_k^\phi)^{-\frac{1}{2}} \beta_k, V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}} \beta_l \rangle_k \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle \beta_k, (V_k^\phi)^{-\frac{1}{2}} V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}} \beta_l \rangle_k \\
&= \psi(\beta),
\end{aligned}$$

(7)

Où

$$\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_K \end{pmatrix} \in \mathcal{X}.$$

Puisque  $V_k^\phi = (V_k^\phi)^{\frac{1}{2}} (V_k^\phi)^{\frac{1}{2}}$ , avoir  $\phi(\alpha) \in C_j$  est équivalent à avoir  $\beta \in C'_j$ , où :

$$C'_1 = \{\beta \in \mathcal{X} / \sum_{k=1}^K \|\beta_k\|_k^2 = 1\} = \{\beta \in \mathcal{X} / \|\beta\|_k^2 = 1\}, \quad (8)$$

Et pour  $j \geq 2$  :

$$C'_j = \{\beta \in C_1 / \sum_{k=1}^K \langle \beta_k^{(r)}, \beta_k \rangle_k = 0, \forall r \in \{1, \dots, j-1\}\} = \{\beta \in C_1 / \langle \beta^{(r)}, \beta \rangle_{\mathcal{X}} = 0, \forall r \in \{1, \dots, j-1\}\}. \quad (9)$$

De plus pour tout  $\beta \in \mathcal{X}$ , nous avons :

$$\begin{aligned}
\psi \Phi^{-\frac{1}{2}} \beta &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \sum_{j=1}^K \tau_k^* V_{kl}^\phi \tau_l^* (V_j^\phi)^{-\frac{1}{2}} \tau_j \beta \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \sum_{j=1}^K \delta_{lj} \tau_k^* V_{kl}^\phi (V_j^\phi)^{-\frac{1}{2}} \tau_j \beta \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}} \tau_l \beta,
\end{aligned}$$

Et

$$\begin{aligned}
\Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}} \beta &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \sum_{j=1}^K \tau_j^* (V_j^\phi)^{-\frac{1}{2}} \tau_j \tau_k^* V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}} \tau_l \beta \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \sum_{j=1}^K \delta_{jk} \tau_j^* (V_j^\phi)^{-\frac{1}{2}} V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}} \tau_l \beta \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* (V_k^\phi)^{-\frac{1}{2}} V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}} \tau_l \beta \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* (V_k^\phi)^{-\frac{1}{2}} V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}} \beta_l
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\langle \beta, \Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}} \beta \rangle_X &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle \beta, \tau_k^* (V_k^\phi)^{-\frac{1}{2}} V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}} \beta_l \rangle_X \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle \tau_k \beta, (V_k^\phi)^{-\frac{1}{2}} V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}} \beta_l \rangle_k \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \langle \beta_k, (V_k^\phi)^{-\frac{1}{2}} V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}} \beta_l \rangle_k \\
&= \psi(\beta),
\end{aligned}$$

■

où  $\psi$  est défini dans (7). Alors, le problème d'optimisation de ACG à noyau se réduit à la maximisation de  $\langle \beta, \Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}} \beta \rangle_X$  sous la contrainte (8) et (9). Puisque  $T = \Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}}$  est un opérateur auto-adjoint, il s'agit d'un problème de maximisation bien connu pour lequel une solution est obtenue à partir de l'analyse spectrale de T, comme indiqué dans le théorème.

### Définition 2.1.2.

Les  $\rho_j$  sont appelés les coefficients canoniques. Les  $\alpha^{(j)}$  sont appelés vecteurs de directions canoniques.

Le théorème suivant donne quelques propriétés des coefficients canoniques.

### Théorème 2.1.2.

- (i)  $\forall j \in \{1, \dots, q\}, -1 \leq \rho_j \leq K(K-1)$ .
- (ii)  $\forall j \in \{1, \dots, q\}, \rho_j = 0 \iff \forall (k, l) \in \{1, \dots, K\}^2, k \neq l, V_{kl}^\phi = 0$ .

**Preuve.** (i) Tout d'abord, en utilisant (6), nous avons pour tout  $j \in \{1, \dots, q\}$ ,

$$\rho_j = \rho(\alpha^{(j)}) = E(< \alpha^{(j)}, \phi(X) >^2) - 1 \geq -1.$$

D'autre part, on a :

Puisque, pour tout  $k \in \{1, \dots, K\}$ , on a :

$$\begin{aligned} E(< \alpha_k^j, \phi(X_k) >^2) &= \text{var}(< \alpha_k^j, \phi(X_k) >) \\ &\leq \sum_{l=1}^K \text{var}(< \alpha_l^j, \phi(X_l) >) \\ &= 1. \end{aligned}$$

Il s'ensuit que

$$\rho_j \leq \sum_{k=1}^K \sum_{l=1, l \neq k}^K 1 = K(K-1).$$

(ii) Puisque les  $\rho_j$  sont les valeurs propres de T, on a :

$$\forall j \in \{1, \dots, q\}, \rho_j = 0 \iff T = 0 \iff \Psi = 0 \iff \forall (k, l) \in \{1, \dots, K\}^2, k \neq l, V_{kl}^\phi = 0. \quad \blacksquare$$

**Remarque 2.1.3.** 1) Lorsque  $K=2$ , on a  $\Phi = \tau_1^* V_1^\phi \tau_1 + \tau_2^* V_2^\phi \tau_2$  et  $\Psi = \tau_1^* V_{12}^\phi \tau_2 + \tau_2^* V_{21}^\phi \tau_1$ . Il est alors facile de vérifier que  $T = \tau_1^* S \tau_2 + \tau_2^* S^* \tau_1$ , où  $S = (V_1^\phi)^{-\frac{1}{2}} V_{12}^\phi (V_2^\phi)^{-\frac{1}{2}}$ . Soit  $x$  un vecteur propre de T associé à une valeur propre  $\rho \neq 0$ . Nous avons  $Tx = \rho x$ , ce qui est équivalent à avoir :

$$\tau_1^*(S \tau_2 x - \rho \tau_1 x) = -\tau_2^*(S^* \tau_1 x - \rho \tau_2 x).$$

Ceci implique :

$$\begin{cases} S \tau_2 x = \rho \tau_1 x \\ S^* \tau_1 x = \rho \tau_2 x \end{cases}$$

et, en mettant  $x_1 = \tau_1 x$  et  $x_2 = \tau_2 x$ , on obtient

$$x_2 = \rho^{-1} S^* x_1 \quad R x_1 = \rho^2 x_1, \quad (10)$$

où

$$R = S S^* = (V_1^\phi)^{-\frac{1}{2}} V_{12}^\phi (V_2^\phi)^{-1} V_{21}^\phi (V_2^\phi)^{-\frac{1}{2}}.$$

Inversement, si (10) tient alors, en mettant  $x = \tau_1^* x_1 + \tau_2^* x_2$ , on a :

$$\begin{aligned}
Tx &= \tau_1^* S \tau_2 x + \tau_2^* S^* \tau_1 x \\
&= \tau_1^* S x_2 + \tau_2^* S^* x_1 \\
&= \rho \tau_1^* x_1 + \rho \tau_2^* x_2 \\
&= \rho(\tau_1^* x_1 + \tau_2^* x_2) \\
&= \rho x.
\end{aligned}$$

*De plus, puisque*

$$\begin{aligned}
\|x_2\|_2 &= \rho^{-1} \|S^* x_1\|_1 \\
&= \rho^{-1} \sqrt{\langle S^* x_1, S^* x_1 \rangle_2} \\
&= \rho^{-1} \sqrt{\langle S S^* x_1, x_1 \rangle_1} \\
&= \|x_1\|_1
\end{aligned}$$

Et

$$\|x\|_X^2 = \|x_1\|_1^2 + \|x_2\|_2^2$$

Il s'ensuit que

$$\|x_1\|_1 = \|x_2\|_2 = \frac{1}{\sqrt{2}} \|x\|_X.$$

La remarque précédente montre l'équivalence entre l'ACG à noyau et l'analyse canonique linéaire (ACL) lorsque  $K=2$ . Rappelons que l'ACL de  $X_1$  et  $X_2$  est obtenue à partir de l'analyse spectrale de  $R$  (voir, par exemple, Dauxois et Pousse (1975), Pousse (1992), Fine (2000)). Plus précisément  $\{\beta^{(j)}, \rho_j\}_{1 \leq j \leq q}$  est défini comme dans le Théorème 2.1.1 si, et seulement si,  $\{u_1^{(j)}, u_2^{(j)}, \rho_j^2\}_{1 \leq j \leq q}$ , où  $u_l^j = \frac{1}{\sqrt{2}} \tau_l \beta^{(j)} (l \in \{1, 2\})$ , est un ACL de  $X_1$  et  $X_2$ .

## 2.2. Étude asymptotique de l'analyse canonique généralisée à noyau

Dans cette section, nous traitons de l'estimateur de l'analyse canonique généralisée (ACGN) à noyau. Pour  $k \in \{1, \dots, K\}$ , soit  $\{X_k^{(i)}\}_{1 \leq i \leq n}$  un échantillon iid de  $X_k$ . Nous utilisons des opérateurs de covariance empiriques pour obtenir les estimateurs des éléments ACGN à noyau. Ensuite,

la cohérence et la normalité asymptotique sont obtenues pour les estimateurs résultants des vecteurs de directions canoniques et des coefficients canoniques.

## 2.2.1. Estimations et convergence presque sûre des operateurs, des vecteurs de directions et des coefficients canoniques.

Pour  $(k, l) \in \{1, \dots, K\}^2$ , considérons les moyennes et les opérateurs de covariance :

$$\phi(\bar{X}_{k,n}) = \frac{1}{n} \sum_{i=1}^n \phi(X_k^i), \widehat{V}_{kl,n}^\phi = \frac{1}{n} \sum_{i=1}^n (\phi(X_k^i) - \phi(\bar{X}_{k,n})) \otimes (\phi(X_l^i) - \phi(\bar{X}_{l,n})), \widehat{V}_{k,n} = \widehat{V}_{kk,n}.$$

Et les opérateurs aléatoires évalués dans  $\mathcal{L}(X)$  définis par

$$\widehat{\Phi}_n = \sum_{k=1}^K \tau_k^* \widehat{V}_{k,n}^\phi \tau_k \quad \text{et} \quad \widehat{\Psi}_n = \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* \widehat{V}_{kl,n}^\phi \tau_l.$$

Ensuite, nous estimons T par

$$\widehat{T}_n = \widehat{\Phi}_n^{-\frac{1}{2}} \widehat{\Psi}_n \widehat{\Phi}_n^{-\frac{1}{2}}.$$

Considérant les valeurs propres  $\widehat{\rho}_{1,n} \geq \widehat{\rho}_{2,n} \dots \geq \widehat{\rho}_{q,n}$  de  $\widehat{T}_n$ , et  $\{\widehat{\beta}_n^{(1)}, \dots, \widehat{\beta}_n^{(q)}\}$  une base orthonormale de  $X$  telle que  $\widehat{\beta}_n^{(j)}$  est un vecteur propre de  $\widehat{T}_n$  associé à  $\widehat{\rho}_{j,n}$ . Alors, on estime  $\rho_j$  par  $\widehat{\rho}_{j,n}$ , et  $\beta^{(j)}$  par  $\widehat{\beta}_n^{(j)}$ . Le théorème suivant établit la consistance forte de ces estimateurs.

### **Théorème 2.2.1.**

Pour tout entier  $j \in \{1, \dots, q\}$  :

- (i)  $\widehat{\rho}_{j,n}$  converge presque sûrement,  $n \rightarrow +\infty$ , vers  $\rho_j$ .
- (ii)  $\text{sign}(\langle \widehat{\beta}_n^{(j)}, \beta^{(j)} \rangle_X) \widehat{\beta}_n^{(j)}$  converge presque sûrement vers  $\beta^{(j)}$  lorsque  $n \rightarrow +\infty$

**Preuve.** A partir d'applications évidentes de la loi forte des grands nombres, il est facilement vu que  $\widehat{T}_n$  converge presque sûrement dans  $\mathcal{L}(X)$ , lorsque  $n \rightarrow +\infty$ , vers T. Alors en utilisant le lemme 1 de Ferré et Yao (2003), nous obtenons l'inégalité  $|\widehat{\rho}_{j,n} - \rho_j| \leq \|\widehat{T}_n - T\|$  de laquelle on déduit (i). Clairement, chaque  $\beta^{(j)} \otimes \beta^{(j)}$  est un projecteur sur un espace propre. Par conséquent, en utilisant la proposition 3 de Dossou-Gbete et pousse (1991), nous déduisons que  $\widehat{\beta}_n^{(j)} \otimes \widehat{\beta}_n^{(j)}$  converge presque sûrement dans  $\mathcal{L}(X)$  vers  $\beta^{(j)} \otimes \beta^{(j)}$ , lorsque  $n \rightarrow +\infty$ . En utilisant à nouveau le lemme 1 de Ferré et Yao (2003), on obtient l'inégalité suivante

$$\| \text{sign}(\langle \widehat{\beta}_n^{(j)}, \beta^{(j)} \rangle_X) \widehat{\beta}_n^{(j)} - \beta^{(j)} \|_X \leq 2\sqrt{2} \| \beta_n^{(j)} \otimes \beta_n^{(j)} - \beta^{(j)} \otimes \beta^{(j)} \|$$

De ce nous déduisons (ii). ■

## 2.2.2. Distribution asymptotique des opérateurs de covariances empiriques

Dans cette section, nous supposons que, pour  $k \in \{1, \dots, K\}$ , nous avons  $\mathbb{E}(\|\phi(X_k)\|_k^4) < +\infty$  et  $V_k^\phi = I_k$ , où  $I_k$  désigne l'opérateur d'identité de  $\mathcal{X}_k$ . Nous dérivons d'abord une distribution asymptotique pour  $\widehat{T}_n$ , puis nous obtenons celle des coefficients canoniques.

### **Théorème 2.2.2.**

$\sqrt{n}(\widehat{T}_n - T)$  converge en distribution, lorsque  $n \rightarrow +\infty$ , vers une V.a U ayant une distribution normale dans  $\mathcal{L}(\mathcal{X})$ , avec une moyenne 0 et opérateur de covariance  $\Gamma$  égal à celui de l'opérateur aléatoire :

$$Z = \sum_{k=1}^K \sum_{l=1, l \neq k}^K -\frac{1}{2}(\tau_k^*(\phi(X_k) \otimes \phi(X_k))V_{kl}^\phi \tau_l + \tau_l^* V_{lk}^\phi (\phi(X_k) \otimes \phi(X_k))\tau_k) + \tau_k^*(\phi(X_k) \otimes \phi(X_k))\tau_l.$$

**Preuve.** Sous les hypothèses précédentes,

$$\Phi = \sum_{k=1}^K \tau_k^* V_k^\phi \tau_k = \sum_{k=1}^K \tau_k^* \tau_k = I_{\mathcal{X}},$$

Où  $I_{\mathcal{X}}$  est l'opérateur identité de  $\mathcal{X}$ , et

$$\begin{aligned} \sqrt{n}(\widehat{T}_n - T) &= \sqrt{n}(\widehat{\Phi}_n^{-\frac{1}{2}} \widehat{\Psi}_n \widehat{\Phi}_n^{-\frac{1}{2}} - \Psi) \\ &= \sqrt{n}(\widehat{\Phi}_n^{-\frac{1}{2}} - I_{\mathcal{X}}) \widehat{\Psi}_n \widehat{\Phi}_n^{-\frac{1}{2}} + \sqrt{n}(\widehat{\Psi}_n - \Psi) \widehat{\Phi}_n^{-\frac{1}{2}} + \Psi(\sqrt{n}(\widehat{\Phi}_n^{-\frac{1}{2}} - I_{\mathcal{X}})) \\ &= -\widehat{\Phi}_n^{-1}(\sqrt{n}(\widehat{\Phi}_n - I_{\mathcal{X}}))(\widehat{\Phi}_n^{-\frac{1}{2}} + I_{\mathcal{X}})^{-1} \widehat{\Psi}_n \widehat{\Phi}_n^{-\frac{1}{2}} \\ &\quad + \sqrt{n}(\widehat{\Psi}_n - \Psi) \widehat{\Phi}_n^{-\frac{1}{2}} - \Psi \widehat{\Phi}_n^{-1}(\sqrt{n}(\widehat{\Phi}_n - I_{\mathcal{X}}))(\widehat{\Phi}_n^{-\frac{1}{2}} + I_{\mathcal{X}})^{-1}. \end{aligned} \quad (11)$$

Clairement,

$$V_{kl}^\phi = E(\tau_l(\phi(X)) \otimes \tau_k(\phi(X)))$$

où

$V^\phi = E(\phi(X) \otimes \phi(X))$ . De plus, en mettant



$$\phi(X^i) = \begin{pmatrix} \phi(X_1^i) \\ \vdots \\ \phi(X_K^i) \end{pmatrix},$$

nous avons

$$\begin{aligned} \widehat{V}_{kl,n}^\phi &= \frac{1}{n} \sum_{i=1}^n \phi(X_l^i) \otimes \phi(X_k^i) - \phi(\bar{X}_{l,n}) \otimes \phi(\bar{X}_{k,n}) \\ &= \frac{1}{n} \sum_{i=1}^n \tau_l(\phi(X^i)) \otimes \tau_k(\phi(X^i)) - \tau_l \phi(\bar{X}_n) \otimes \tau_k \phi(\bar{X}_n) \\ &= \widehat{\tau_k V_n^\phi \tau_l^*}, \end{aligned} \tag{13}$$

où  $\phi(\bar{X}_n) = n^{-1} \sum_{i=1}^n \phi(X^i)$  et

$$\widehat{V}_n^\phi = \frac{1}{n} \sum_{i=1}^n \phi(X^i) \otimes \phi(X^i) - \phi(\bar{X}_n) \otimes \phi(\bar{X}_n). \tag{14}$$

Par conséquent, en utilisant (12) et (13), on obtient

$$\sqrt{n}(\widehat{\Psi}_n - \Psi) = \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* \tau_k \widehat{H}_n \tau_l^* \tau_l = f(\widehat{H}_n), \tag{15}$$

où  $\widehat{H}_n = \sqrt{n}(\widehat{V}_n^\phi - V^\phi)$  et  $f$  est l'opérateur défini comme suit

$$f : A \in \mathcal{L}(\mathcal{X}) \mapsto \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* \tau_k A \tau_l^* \tau_l \in \mathcal{L}(\mathcal{X}).$$

De plus, puisque  $I_{\mathcal{X}} = \sum_{k=1}^K \tau_k^* \tau_k$ , on obtient

$$\sqrt{n}(\widehat{\Phi}_n - I_{\mathcal{X}}) = \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* \tau_k \widehat{H}_n \tau_k^* \tau_k = g(\widehat{H}_n), \tag{16}$$

où  $g$  est l'opérateur :  $A \in \mathcal{L}(\mathcal{X}) \mapsto \sum_{k=1}^K \tau_k^* \tau_k A \tau_k^* \tau_k \in \mathcal{L}(\mathcal{X})$ . Alors, en utilisant (11), (15) et (16), nous obtenons  $\sqrt{n}(\widehat{T}_n - T) = \widehat{\varphi}_n(\widehat{H}_n)$ , où  $\widehat{\varphi}_n$  est l'opérateur aléatoire de  $\mathcal{L}(\mathcal{X})$  à lui-même défini par

$$\widehat{\varphi}_n(A) = -(\widehat{\Phi}_n^{-\frac{1}{2}} + I_{\mathcal{X}})^{-1} g(A) \widehat{\Phi}_n^{-1} \widehat{\Psi}_n \widehat{\Phi}_n^{-\frac{1}{2}} + f(A) \widehat{\Phi}_n^{-\frac{1}{2}} - \Psi(\widehat{\Phi}_n^{-\frac{1}{2}} + I_{\mathcal{X}})^{-1} g(A) \widehat{\Phi}_n^{-1}.$$

En considérant l'opérateur

$$\varphi : A \in \mathcal{L}(\mathcal{X}) \mapsto -\frac{1}{2} g(A) \Psi + f(A) - \frac{1}{2} \Psi g(A) \in \mathcal{L}(\mathcal{X}),$$

et en désignant par  $\| \cdot \|_\infty$  (resp  $\| \cdot \|_{\infty\infty}$  la norme de  $\mathcal{L}(X)$  (resp.  $\mathcal{L}(\mathcal{L}(X))$  définie par  $\| A \|_\infty = \sup_{x \in X - \{0\}} \| Ax \|_X / \| x \|_X$  (resp.  $\| h \|_{\infty\infty} = \sup_{B \in \mathcal{L}(X) - \{0\}} \| h(B) \|_\infty / \| B \|_\infty$  pour tout  $A$  (resp.  $h$ ) dans  $\mathcal{L}(X)$  (resp  $\mathcal{L}(\mathcal{L}(X))$ ), nous avons

$$\begin{aligned}
\| \widehat{\varphi}_n(\widehat{H}_n) - \varphi(\widehat{H}_n) \|_\infty &= \| -((\widehat{\Phi}_n^{-\frac{1}{2}} + I_X)^{-1} - \frac{1}{2}I_X)g(\widehat{H}_n)\widehat{\Phi}_n^{-1}\widehat{\Psi}_n\widehat{\Phi}_n^{-\frac{1}{2}} - \frac{1}{2}g(\widehat{H}_n)(\widehat{\Phi}_n^{-1}\widehat{\Psi}_n\widehat{\Phi}_n^{-\frac{1}{2}} - \Psi) \\
&\quad + f(\widehat{H}_n)(\widehat{\Phi}_n^{-\frac{1}{2}} - I_X) - \Psi((\widehat{\Phi}_n^{-\frac{1}{2}} + I_X)^{-1} - \frac{1}{2}I_X)g(\widehat{H}_n)\widehat{\Phi}_n^{-1} - \frac{1}{2}\Psi(\widehat{\Phi}_n^{-1} - I_X) \|_\infty \\
&\leq \| (\widehat{\Phi}_n^{-\frac{1}{2}} + I_X)^{-1} - \frac{1}{2}I_X \|_\infty \| g(\widehat{H}_n) \|_\infty \| \widehat{\Phi}_n^{-1}\widehat{\Psi}_n\widehat{\Phi}_n^{-\frac{1}{2}} \|_\infty \\
&\quad + \frac{1}{2} \| g(\widehat{H}_n) \|_\infty \| \widehat{\Phi}_n^{-1}\widehat{\Psi}_n\widehat{\Phi}_n^{-\frac{1}{2}} - \Psi \|_\infty + \| f(\widehat{H}_n) \|_\infty \| \widehat{\Phi}_n^{-\frac{1}{2}} - I_X \|_\infty \\
&\quad + \| \Psi \|_\infty \| \widehat{\Phi}_n^{-\frac{1}{2}} + I_X)^{-1} - \frac{1}{2}I_X \|_\infty \| g(\widehat{H}_n) \|_\infty \| \widehat{\Phi}_n^{-1} \|_\infty \\
&\quad + \frac{1}{2} \| \Psi \|_\infty \| g(\widehat{H}_n) \|_\infty \| \widehat{\Phi}_n^{-1} - I_X \|_\infty \\
&\leq (\| (\widehat{\Phi}_n^{-\frac{1}{2}} + I_X)^{-1} - \frac{1}{2}I_X \|_\infty \| g \|_{\infty\infty} \| \widehat{\Phi}_n^{-1}\widehat{\Psi}_n\widehat{\Phi}_n^{-\frac{1}{2}} \|_\infty \\
&\quad + \frac{1}{2} \| g \|_{\infty\infty} \| \widehat{\Phi}_n^{-1}\widehat{\Psi}_n\widehat{\Phi}_n^{-\frac{1}{2}} - \Psi \|_\infty + \| f \|_{\infty\infty} \| \widehat{\Phi}_n^{-\frac{1}{2}} - I_X \|_\infty \\
&\quad + \| \Psi \|_\infty \| \widehat{\Phi}_n^{-\frac{1}{2}} + I_X)^{-1} - \frac{1}{2}I_X \|_\infty \| g \|_{\infty\infty} \| \widehat{\Phi}_n^{-1} \|_\infty \\
&\quad + \frac{1}{2} \| \Psi \|_\infty \| g \|_{\infty\infty} \| \widehat{\Phi}_n^{-1} - I_X \|_\infty) \| \widehat{H}_n \|_\infty .
\end{aligned} \tag{17}$$

En utilisant la loi forte des grands nombres, il est facile de vérifier que, pour tout  $(k, l) \in (\{1, \dots, K\}^2)$  avec  $k \neq l$ ,  $\widehat{V}_{kl,n}^\phi$  (resp.  $\widehat{V}_{k,n}^\phi$ ) converge presque sûrement vers  $\widehat{V}_{kl}^\phi$  (resp  $\widehat{V}_k^\phi$ ), lorsque  $n \rightarrow +\infty$ . Par conséquent,  $\widehat{\Phi}_n$  (resp  $\widehat{\Psi}_n$ ) converge presque sûrement vers  $\Phi = I_X$  (resp  $\Psi$ ), lorsque  $n \rightarrow +\infty$ . Ceci implique la convergence presque sûre de  $(\widehat{\Phi}_n^{-\frac{1}{2}} + I_X)^{-1}$  (resp  $\widehat{\Phi}_n^{-1}\widehat{\Psi}_n\widehat{\Phi}_n^{-\frac{1}{2}}$ ; resp  $\widehat{\Phi}_n^{-1}$  resp  $\widehat{\Phi}_n^{-\frac{1}{2}}$ ) vers  $\frac{1}{2}I_X$  (resp.  $\Psi$ ); resp  $I_X$  resp  $I_X$ ; lorsque  $n \rightarrow +\infty$ . En outre en désignant par  $\| \cdot \|$  la norme de  $\mathcal{L}(X)$  définie par  $\| A \| = \sqrt{\text{tr}(AA^*)}$  et en utilisant les propriétés  $(a \otimes b)(c \otimes d) = \langle a, d \rangle c \otimes b$  et  $\text{tr}(a \otimes b) = \langle a, b \rangle$  du produit tensoriel (voir Dauxois et al. (1994)), on a :

$$\begin{aligned}
\mathbb{E}(\| \phi(X) \otimes \phi(X) \|^2) &= \mathbb{E}(\text{tr}((\phi(X) \otimes \phi(X))(\phi(X) \otimes \phi(X))) \\
&= \mathbb{E}(\| \phi(X) \|_X^4) \\
&= \mathbb{E}((\sum_{k=1}^K \| \phi(X) \|_k^2)^2) \\
&= \sum_{k=1}^K \mathbb{E}(\| \phi(X) \|_k^4) + \sum_{k=1}^K \sum_{l=1, l \neq k}^K \mathbb{E}(\| \phi(X_k) \|_k^2 \| \phi(X_l) \|^2)
\end{aligned}$$

$$\leq \sum_{k=1}^K \mathbb{E}(\|\phi(X)\|_k^4) + \sum_{k=1}^K \sum_{l=1, l \neq k}^K \sqrt{\mathbb{E}(\|\phi(X_k)\|_k^4)} \sqrt{\mathbb{E}(\|\phi(X_l)\|_l^4)} < +\infty.$$

Alors, le théorème de limite centrale peut être utilisé. Il donne la convergence en distribution, quand  $n \rightarrow +\infty$ , de  $\sqrt{n}(n^{-1} \sum_{i=1}^n \phi(X^{(i)}) \otimes \phi(X^{(i)}) - V^\phi)$  vers une V.a H ayant la distribution normale dans  $\mathcal{L}(X)$  avec une moyenne égale à 0 et un opérateur de covariance égale à celui de  $\phi(X) \otimes \phi(X)$ . Puisque, par le théorème central limite encore,  $\sqrt{n}\phi(\bar{X}_n)$  converge en distribution, lorsque  $n \rightarrow +\infty$ , vers une V.a ayant une distribution normale dans  $\mathcal{X}$  avec une moyenne égale à 0 et un opérateur de covariance égal à  $V^\phi$ , on déduit de l'égalité  $\sqrt{n}(\phi(\bar{X}_n) \otimes \phi(\bar{X}_n) = n^{-\frac{1}{2}}[\sqrt{n}\phi(\bar{X}_n)] \otimes (\sqrt{n}\phi(\bar{X}_n))$  que  $\sqrt{n}(\phi(\bar{X}_n) \otimes \phi(\bar{X}_n))$  converge en probabilité vers 0, lorsque  $n \rightarrow +\infty$ . Par conséquent, d'après (14) et du théorème de Slutsky, on déduit que  $\widehat{H}_n$  converge en distribution, lorsque  $n \rightarrow +\infty$  vers H. Puis, à partir de (17) nous concluons que  $\widehat{\varphi}_n(\widehat{H}_n) - \varphi(\widehat{H}_n)$  converge en probabilité vers 0, lorsque  $n \rightarrow +\infty$ . Puis, en utilisant à nouveau le théorème de Slutsky, nous déduisons que la distribution de  $\widehat{\varphi}_n(\widehat{H}_n)$  et  $\varphi(\widehat{H}_n)$  converge vers la même distribution. Puisque  $\varphi$  est une carte linéaire (et est, par conséquent, continue), cette distribution est simplement celle de la V.a  $U = \varphi(H)$ , ie la distribution normale dans  $\mathcal{L}(X)$  avec pour moyenne 0 et opérateur de covariance égal à celui de  $Z = \varphi(\phi(X) \otimes \phi(X))$ . Il est clair que,

$$\begin{aligned} g(\phi(X) \otimes \phi(X)) &= \sum_{k=1}^K \tau_k^* \tau_k (\phi(X) \otimes \phi(X)) \tau_k^* \tau_k \\ &= \sum_{k=1}^K \tau_k^* ((\tau_k \phi(X))) \otimes (\tau_k \phi(X)) \tau_k \\ &= \sum_{k=1}^K \tau_k^* (\phi(X_k) \otimes \phi(X_k)) \tau_k, \end{aligned}$$

et

$$\begin{aligned} f(\phi(X) \otimes \phi(X)) &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* \tau_k (\phi(X) \otimes \phi(X)) \tau_l^* \tau_l \\ &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* (\phi(X_l) \otimes \phi(X_k)) \tau_l. \end{aligned}$$

Puis, puisque  $\tau_k \tau_j^* = \delta_{kj} I_k$ , il s'ensuit que

$$\begin{aligned} g(\phi(X) \otimes \phi(X)) \Psi &= \sum_{k=1}^K \sum_{j=1}^K \sum_{l=1, l \neq j}^K \tau_k^* (\phi(X_k) \otimes \phi(X_k)) \tau_k \tau_j^* V_{jl}^\phi \tau_l \\ &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* (\phi(X_k) \otimes \phi(X_k)) V_{kl}^\phi \tau_l. \end{aligned}$$

et

$$\begin{aligned}
\Psi g(\phi(X) \otimes \phi(X)) &= \sum_{k=1}^K \sum_{j=1}^K \sum_{l=1, l \neq j}^K \tau_k^* V_{kl}^\phi \tau_l \tau_j^* (\phi(X_j) \otimes \phi(X_j)) \tau_j \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_k^* V_{kl}^\phi (\phi(X_l) \otimes \phi(X_l)) \tau_l \\
&= \sum_{k=1}^K \sum_{l=1, l \neq k}^K \tau_l^* V_{lk}^\phi (\phi(X_k) \otimes \phi(X_k)) \tau_k.
\end{aligned}$$

Ainsi,

$$Z = \sum_{k=1}^K \sum_{l=1, l \neq k}^K -\frac{1}{2} (\tau_k^* (\phi(X_k) \otimes \phi(X_k)) V_{kl}^\phi \tau_l) + \tau_l^* V_{lk}^\phi (\phi(X_k) \otimes \phi(X_k)) \tau_k + \tau_k^* (\phi(X_l) \otimes \phi(X_k)) \tau_l$$

■

En utilisant le théorème précédent et les résultats de Eaton et Tyler (19991,1994), nous pouvons maintenant donner des distributions asymptotiques pour les coefficients canoniques. Nous désignons, par  $(\rho'_j)_{1 \leq j \leq r}$  (avec  $r \in N^*$ ) la suite des valeurs propres distinctes de T en ordre décroissante, ie  $\rho'_1 > \dots > \rho'_r$ . En mettant  $m_0 = 0$ , en désignant par  $m_j$  la multiplicité de  $\rho'_j$  et en mettant  $v_j = \sum_{k=0}^{j-1} m_k$  pour tout  $j \in \{1, \dots, r\}$ , il est clair que pour tout  $i \in \{v_{j-1} + 1, \dots, v_j\}$ , on a  $\rho_i = \rho'_j$ . De plus, en considérant l'espace propre  $E_j = \ker(T - \rho'_j I)$ , on a la décomposition suivante en somme directe orthogonale :  $X = E_1 \oplus \dots \oplus E_r$ . On désigne par  $\Pi_j$  le projecteur orthogonal de  $X$  sur  $E_j$ , et par  $\Delta$  l'application continue qui associe à chaque opérateur auto-adjoint A le vecteur  $\Delta(A)$  de ses valeurs propres en ordre non croissant. Pour  $j \in \{1, \dots, r\}$ , on considère un vecteur à  $m_j$  de dimensions donné par  $v_j = \rho'_j \mathcal{J}_{m_j}$ , où  $\mathcal{J}_q$  désigne le vecteur à q dimensions donc les éléments sont tous égaux à 1, et le vecteur aléatoire à valeur  $R^{m_j}$  :

$$\widehat{v}_j^n = \begin{pmatrix} \widehat{\rho} v_{j-1} + 1.n \\ \cdot \\ \cdot \\ \cdot \\ \widehat{\rho} v_{j,n} \end{pmatrix},$$

Alors, en mettant

$$\widehat{\Lambda}_n = \begin{pmatrix} \widehat{v}_1^n \\ \cdot \\ \cdot \\ \cdot \\ \widehat{v}_r^n \end{pmatrix} \text{ et } \Lambda = \begin{pmatrix} v_1 \\ \cdot \\ \cdot \\ \cdot \\ v_r \end{pmatrix}$$

nous avons

### **Théorème 2.2.3.**

$\sqrt{n}(\widehat{\Lambda}_n - \Lambda)$  converge en distribution,  $n \in \{1, \dots, +\infty\}$ , vers le vecteur aléatoire à valeur  $\mathbb{R}^p$

$$\zeta = \begin{pmatrix} \Delta(\Pi_1 W \Pi_1) \\ \cdot \\ \cdot \\ \cdot \\ \Delta(\Pi_r W \Pi_r) \end{pmatrix}, \quad (18)$$

où  $W$  est une V.a ayant une distribution normale dans  $\mathcal{L}(X)$ , avec moyenne 0 et un opérateur de covariance  $\Theta$  donnée par :

$$\Theta = \sum_{1 \leq m, r, s, t \leq p} C(m, r, s, t) (e_m \otimes e_r) \widetilde{\otimes} (e_s \otimes e_t)$$

avec

$$C(m, r, s, t) =$$

$$\sum_{k=1}^K \sum_{j=1}^K \sum_{l=1, l \neq k}^K \sum_{q=1, q \neq j}^K (\gamma_{kljq}^{m, r, s, t} + \gamma_{kljq}^{m, r, t, s} + \gamma_{kljq}^{r, m, s, t} + \gamma_{kljq}^{r, m, t, s} - \theta_{kljq}^{m, r, s, t} - \theta_{kljq}^{r, m, s, t} - \theta_{kljq}^{s, t, m, r} - \theta_{kljq}^{t, s, m, r} + \lambda_{kljq}^{m, r, s, t}),$$

$$\gamma_{kljq}^{a, b, c, d} = \frac{1}{4} E(\langle \phi(X_k), \tau_k \beta^{(a)} \rangle_k \langle \phi(X_k), V_{kl}^\phi \tau_l \beta^{(b)} \rangle_k \langle \phi(X_j), \tau_j \beta^{(c)} \rangle_j \langle \phi(X_j), V_{jq}^\phi \tau_q \beta^{(d)} \rangle_j),$$

$$\theta_{kljq}^{a, b, c, d} = \frac{1}{2} E(\langle \phi(X_k), \tau_k \beta^{(a)} \rangle_k \langle \phi(X_k), V_{kl}^\phi \tau_l \beta^{(b)} \rangle_k \langle \phi(X_j), \tau_j \beta^{(c)} \rangle_j \langle \phi(X_q), \tau_q \beta^{(d)} \rangle_q),$$

$$\gamma_{kljq}^{a, b, c, d} = E(\langle \phi(X_k), \tau_k \beta^{(a)} \rangle_k \langle \phi(X_l), \tau_l \beta^{(b)} \rangle_l \langle \phi(X_j), \tau_j \beta^{(c)} \rangle_j \langle \phi(X_q), \tau_q \beta^{(d)} \rangle_q).$$

et

$(e_i)_{1 \leq i \leq p}$  base canonique de  $\mathbb{R}^p$

**Preuve.** Puisque  $\Delta(\widehat{T}_n) = \widehat{\Lambda}_n$  et  $\Delta(T) = \Lambda$ , on déduit du théorème 3.2 et du théorème 2.1 de Eaton et Tyler (1994) que  $\sqrt{n}(\widehat{\Lambda}_n - \Lambda)$  converge en distribution, lorsque  $n \rightarrow +\infty$ , vers la V.a donnée dans (18) avec  $W = P^* U P$ , où  $P = \sum_{l=1}^p e_l \otimes \beta^{(l)}$ . clairement,  $W$  a une distribution

normale avec pour moyenne 0 de même que  $P^*ZP$ . Afin de donner une expression de  $\Theta$ , notons d'abord que :

$$\begin{aligned} P^*ZP &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K -\frac{1}{2} (P^*\tau_k^*(\phi(X_k) \otimes \phi(X_k))V_{kl}^\phi \tau_l P + P^*\tau_l^*V_{lk}^\phi(\phi(X_k) \otimes \phi(X_k))\tau_k P) + \\ P^*\tau_k^*(\phi(X_l) \otimes \phi(X_k))\tau_l P &= \sum_{k=1}^K \sum_{l=1, l \neq k}^K -\frac{1}{2} ((P^*\tau_l^*V_{lk}^\phi(\phi(X_k))) \otimes (P^*\tau_k^*\phi(X_k))) + (P^*\tau_k^*\phi(X_k)) \otimes \\ &+ (P^*\tau_l^*V_{lk}^\phi(\phi(X_k))) + (P^*\tau_k^*\phi(X_l)) \otimes (P^*\tau_k^*\phi(X_k)). \end{aligned}$$

Puisque

$$\begin{aligned} P^*\tau_l^*V_{lk}^\phi X_k &= \left( \sum_{m=1}^p \beta^{(m)} \otimes e_m \right) \tau_l^*V_{lk}^\phi \phi(X_k) \\ &= \sum_{m=1}^p \langle \beta^{(m)}, \tau_l^*V_{lk}^\phi \phi(X_k) \rangle_X e_m \\ &= \sum_{m=1}^p \langle \tau_l \beta^{(m)}, V_{lk}^\phi \phi(X_k) \rangle e_m \end{aligned}$$

et, de même,  $P^*\tau_l^*\phi(X_k) = \sum_{m=1}^p \langle \tau_l \beta^{(m)}, \phi(X_k) \rangle_k e_m$ , il s'ensuit :

$$\begin{aligned} P^*ZP &= \sum_{m=1}^p \sum_{r=1}^p \left[ \sum_{k=1}^K \sum_{l=1, l \neq k}^K -\frac{1}{2} \langle \tau_l \beta^{(m)}, V_{lk}^\phi \phi(X_k) \rangle_l \langle \tau_k \beta^{(m)}, \phi(X_k) \rangle_k + \langle \right. \\ &\left. \tau_l \beta^{(r)}, V_{lk}^\phi \phi(X_k) \rangle_l \langle \tau_k \beta^{(m)}, \phi(X_k) \rangle_k \right] + \langle \tau_l \beta^{(m)}, \phi(X_l) \rangle_l \langle \tau_k \beta^{(r)}, \phi(X_k) \rangle_k e_m \otimes e_r. \end{aligned}$$

De :

$$\begin{aligned} E(\langle \tau_l \beta^{(m)}, V_{lk}^\phi \phi(X_k) \rangle_l \langle \tau_k \beta^{(m)}, \phi(X_k) \rangle_k) &= E(\langle (\phi(X_k) \otimes \phi(X_k))(\tau_k \beta^{(r)}), V_{kl}^\phi \tau_l \beta^{(m)} \rangle_k) \\ &= \langle E(\phi(X_k) \otimes \phi(X_k))(\tau_k \beta^{(r)}), V_{kl}^\phi \tau_l \beta^{(m)} \rangle_k \\ &= \langle V_k^\phi \tau_k \beta^{(r)}, V_{kl}^\phi \tau_l \beta^{(m)} \rangle_k \\ &= \langle \tau_k \beta^{(r)}, V_{kl}^\phi \tau_l \beta^{(m)} \rangle_k, \end{aligned}$$

$$E(\langle \tau_l \beta^{(m)}, V_{lk}^\phi \phi(X_k) \rangle_l \langle \tau_k \beta^{(m)}, \phi(X_k) \rangle_k) = \langle \tau_k \beta^{(m)}, V_{kl}^\phi \tau_l \beta^{(r)} \rangle_k$$

et

$$\begin{aligned} E(\langle \tau_l \beta^{(m)}, \phi(X_l) \rangle_l \langle \tau_k \beta^{(r)}, \phi(X_k) \rangle_k) &= E(\langle (\phi(X_l) \otimes \phi(X_k))(\tau_l \beta^{(m)}), \tau_k \beta^{(r)} \rangle_k) \\ &= \langle E(\phi(X_l) \otimes \phi(X_k))(\tau_l \beta^{(m)}), \tau_k \beta^{(r)} \rangle_k \\ &= \langle V_{kl}^\phi \tau_l \beta^{(m)}, \tau_k \beta^{(r)} \rangle_k, \end{aligned}$$

on en déduit que  $E(P^*ZP) = 0$ . Donc,

$$\Theta = E((P^*ZP) \widetilde{\otimes} (P^*ZP)) = \sum_{1 \leq m, r, s, t \leq p} C(m, r, s, t) (e_m \otimes e_r) \widetilde{\otimes} (e_s \otimes e_t), \text{ où}$$

$$C(m, r, s, t) = \sum_{k=1}^K \sum_{l=1, l \neq k}^K \sum_{q=1, q \neq j}^K E(Y_{kl}^{m,r} Y_{jq}^{s,q})$$

avec

$$Y_{kl}^{m,r} = -\frac{1}{2}(\langle \tau_l \beta^{(m)}, V_{lk}^\phi \phi(X_k) \rangle_l \langle \tau_k \beta^{(r)}, \phi(X_k) \rangle + \langle \tau_l \beta^{(r)}, V_{lk}^\phi \phi(X_k) \rangle_l \langle \tau_k \beta^{(m)}, \phi(X_k) \rangle_k) + \langle \tau_l \beta^{(m)}, \phi(X_l) \rangle_l \langle \tau_k \beta^{(r)}, \phi(X_k) \rangle_k.$$

D'autres calculs donnent

$$E(Y_{kl}^{m,r} Y_{jq}^{s,q}) = \gamma_{kljq}^{m,r,s,t} + \gamma_{kljq}^{m,r,t,s} + \gamma_{kljq}^{r,m,s,t} + \gamma_{kljq}^{r,m,t,s} - \theta_{kljq}^{m,r,s,t} - \theta_{kljq}^{r,m,s,t} - \theta_{kljq}^{s,t,m,r} - \theta_{kljq}^{m,r,s,t}.$$

■

Lorsque T a des valeurs propres simples, ie  $\rho_1 > \dots > \rho_q$ , le théorème précédent est plus simple. Nous avons :

### Corollaire

Lorsque les valeurs propres de T sont simples,  $\sqrt{n}(\widehat{\Lambda}_n - \Lambda)$  converge en distribution, lorsque  $n \rightarrow +\infty$ , vers une V.a ayant une distribution normale dans  $R^p$  ayant pour moyenne 0 et matrice de covariance  $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq p}$  avec :

$$\sigma_{ij} = \sum_{1 \leq m, r, s, t \leq p} \beta_m^{(i)} \beta_r^{(i)} \beta_s^{(j)} \beta_t^{(j)} C(m, r, s, t).$$

**Preuve.** : Dans ce cas,  $m_1 = \dots = m_p = 1$  et, pour tout  $j \in \{1, \dots, p\}$ ,  $\Pi_j = \beta^{(j)} \otimes \beta^{(j)}$ . Ainsi

$$\begin{aligned} \Pi_j W \Pi_j &= ((\beta^{(j)} \otimes \beta^{(j)}) W (\beta^{(j)} \otimes \beta^{(j)})) \\ &= (\beta^{(j)} \otimes \beta^{(j)}) (\beta^{(j)} \otimes (W \beta^{(j)})) \\ &= \langle \beta^{(j)}, W \beta^{(j)} \rangle_X \beta^{(j)} \otimes \beta^{(j)}, \end{aligned}$$

et, par conséquent,  $\Delta(\Pi_j W \Pi_j) = \langle \beta^{(j)}, W \beta^{(j)} \rangle_X$ . Alors,  $\zeta$  est une fonction linéaire de W et, par conséquent, elle a une distribution normale de moyenne 0 et de matrice de covariance  $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq p}$  avec  $\sigma_{ij} = E(\langle \beta^{(i)}, W \beta^{(i)} \rangle_X \langle \beta^{(j)}, W \beta^{(j)} \rangle_X)$ . En notant par  $\langle \cdot, \cdot \rangle$  le produit interne des opérateurs définis par  $\langle A, B \rangle = \text{tr}(AB^*)$ , on a :

$$\begin{aligned} \langle W, \beta^{(j)} \otimes \beta^{(j)} \rangle &= \text{tr}(W(\beta^{(j)} \otimes \beta^{(j)})) \\ &= \text{tr}(\beta^{(j)} \otimes (W \beta^{(j)})) \\ &= \langle \beta^{(j)}, W \beta^{(j)} \rangle_X, \end{aligned}$$

il s'ensuit que

$$\begin{aligned}
\sigma_{ij} &= E(< \beta^{(i)}, W\beta^{(i)} >_X < \beta^{(j)}, W\beta^{(j)} >_X) \\
&= E(< W, \beta^{(i)} \otimes \beta^{(i)} > < W, \beta^{(j)} \otimes \beta^{(j)} >) \\
&= E(< W \otimes W, (\beta^{(i)} \otimes \beta^{(i)}), \beta^{(j)} \otimes \beta^{(j)} >) \\
&= < E(W \otimes W)(\beta^{(i)} \otimes \beta^{(i)}), \beta^{(j)} \otimes \beta^{(j)} > \\
&= < \Theta(\beta^{(i)} \otimes \beta^{(i)}), \beta^{(j)} \otimes \beta^{(j)} > \\
&= \sum_{1 \leq m, r, s, t \leq p} C(m, r, s, t) < ((e_m \otimes e_r) \otimes (e_s \otimes e_t))(\beta^{(i)} \otimes \beta^{(i)}), \beta^{(j)} \otimes \beta^{(j)} > \\
&= \sum_{1 \leq m, r, s, t \leq p} C(m, r, s, t) < e_m \otimes e_r, \beta^{(i)} \otimes \beta^{(i)} > < e_s \otimes e_t, \beta^{(j)} \otimes \beta^{(j)} > .
\end{aligned}$$

Alors, le résultat requis est obtenu par .

$$\begin{aligned}
< e_m \otimes e_r, \beta^{(i)} \otimes \beta^{(i)} > &= tr((e_m \otimes e_r)(\beta^{(i)} \otimes \beta^{(i)})) \\
&= tr(< e_m, \beta^{(i)} >_X \beta^{(i)} \otimes e_r) \\
&= < e_m, \beta^{(i)} >_X < e_r, \beta^{(i)} >_{(X)} \\
&= \beta_m^{(i)} \beta_r^{(i)}
\end{aligned}$$

$$\text{et } < e_s \otimes e_t, \beta^{(j)} \otimes \beta^{(j)} > = \beta_s^{(j)} \beta_t^{(j)} .$$

■



# SIMULATIONS ET APPLICATIONS

## 3.1. Introduction

Ce chapitre est consacré au test de non-corrélation dans le cas multidimensionnel. Une application dans le cas particulier  $k = 2$  sur les variables fonctionnelles est ainsi réalisé pour tester la performance de nos travaux. L'utilisation de ce type de variables est récent et peu abordé dans la littérature. Un aperçu de ce test est abordé dans [1]. Pour fixer les idées, il utilise les projections sur un sous espace de dimension finie réalisant une approximation de l'espace de Hilbert dans lequel  $X$  prend ses valeurs. Nous proposons par la suite la statistique de test et déterminons sa loi limite sous l'hypothèse nulle, ce qui nous permettra d'élaborer un algorithme correspondant à ce test.

## 3.2. Test de non-corrélation mutuelle multidimensionnelle

Dans cette section, nous considérons le problème du test de non-corrélation mutuelle entre  $X_1, \dots, X_K$ , c'est-à-dire tester la non-corrélation entre toute paire  $(X_k, X_l)$  pour  $(k, l) \in \{1, \dots, K\}^2$ . L'hypothèse nulle est

$$\mathcal{H}_0 : \forall (k, l) \in \{1, \dots, K\}^2, k \neq l, V_{kl}^\phi = 0$$

et l'alternative est donnée par

$$\mathcal{H}_1 : \exists (k, l) \in \{1, \dots, K\}^2, k \neq l, V_{kl}^\phi \neq 0$$

c'est-à-dire tester la non-corrélation entre toute paire  $(X_k, X_l)$  de variable. Nous allons d'abord introduire une statistique de test pour ce problème, puis nous dériverons sa distribution asymptotique sous l'hypothèse nulle, dans le cas général en utilisant les résultats de la section précédente.

### 3.2.1. Une statistique de test

Pour  $(k, l) \in \{1, \dots, K\}^2$ , on désigne par  $\pi_{kl}$  l'opérateur

$$\pi_{kl} : A \in \mathcal{L}(\mathcal{X}) \mapsto \tau_k A \tau_l^* \in \mathcal{L}(\mathcal{X}_l, \mathcal{X}_k),$$

on a pour  $k \neq l$  :

$$\begin{aligned} \pi_{kl}(T) &= \pi_{kl}(\Phi^{-\frac{1}{2}} \Psi \Phi^{-\frac{1}{2}}) \\ &= \sum_{i=1}^K \sum_{j=1, j \neq i}^K \tau_k \tau_i^* (V_i^\phi)^{-\frac{1}{2}} V_{ij}^\phi (V_j^\phi)^{-\frac{1}{2}} \tau_j \tau_l^* \\ &= \sum_{i=1}^K \sum_{j=1, j \neq i}^K \delta_{ki} \delta_{jl} V_{jl}^\phi (V_i^\phi)^{-\frac{1}{2}} V_{ij}^\phi (V_j^\phi)^{-\frac{1}{2}} \\ &= (V_k^\phi)^{-\frac{1}{2}} V_{kl}^\phi (V_l^\phi)^{-\frac{1}{2}}. \end{aligned}$$

Par conséquent,  $\mathcal{H}_0$  équivaut à avoir  $\sum_{k=2}^K \sum_{l=1}^{k-1} \text{tr}(\pi_{kl}(T) \pi_{kl}(T)^*) = 0$ . Ceci nous conduit à prendre comme statistique de test la V.a  $\widehat{S}_n$  donnée par :

$$\widehat{S}_n = \sum_{k=2}^K \sum_{l=1}^{k-1} \text{tr}(\pi_{kl}(\widehat{T}_n) \pi_{kl}(\widehat{T}_n)^*).$$

### 3.2.2. Distribution asymptotique sous l'hypothèse nulle

Pour  $k \in \{1, \dots, K\}$ , nous désignons par  $\{e_j^{(k)}; j = 1 \dots p_k\}$  une base orthonormale de  $\mathcal{X}_k$ , où  $p_k = \dim(\mathcal{X}_k)$ , et nous considérons la matrice :

$$\Gamma = \begin{bmatrix} \Gamma_{21,21} & \Gamma_{21,31} & \Gamma_{21,32} & \dots & \Gamma_{21,KK-1} \\ \Gamma_{31,21} & \Gamma_{31,31} & \Gamma_{31,32} & \dots & \Gamma_{31,KK-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Gamma_{KK-1,21} & \Gamma_{KK-1,31} & \Gamma_{KK-1,32} & \dots & \Gamma_{KK-1,KK-1} \end{bmatrix}, \quad (19)$$

où  $\Gamma_{kl,rs}$  est la matrice  $p_k p_l \times p_r p_s$  donnée par

$$\Gamma = \begin{bmatrix} \gamma_{1111}^{kl,rs} & \gamma_{1121}^{kl,rs} & \cdots & \gamma_{11p_r1}^{kl,rs} & \cdots & \gamma_{111p_s}^{kl,rs} & \gamma_{112p_s}^{kl,rs} & \vdots & \gamma_{111p_r p_s}^{kl,rs} \\ \gamma_{2111}^{kl,rs} & \gamma_{2121}^{kl,rs} & \cdots & \gamma_{21p_r1}^{kl,rs} & \cdots & \gamma_{211p_s}^{kl,rs} & \gamma_{212p_s}^{kl,rs} & \vdots & \gamma_{21p_r p_s}^{kl,rs} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{p_k111}^{kl,rs} & \gamma_{p_k121}^{kl,rs} & \cdots & \gamma_{p_k1p_r1}^{kl,rs} & \cdots & \gamma_{p_k11p_s}^{kl,rs} & \gamma_{p_k12p_s}^{kl,rs} & \vdots & \gamma_{p_k1p_r p_s}^{kl,rs} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{1p_l11}^{kl,rs} & \gamma_{1p_l21}^{kl,rs} & \cdots & \gamma_{1p_l1p_r}^{kl,rs} & \cdots & \gamma_{1p_l1p_s}^{kl,rs} & \gamma_{1p_l2p_s}^{kl,rs} & \vdots & \gamma_{1p_l1p_r p_s}^{kl,rs} \\ \gamma_{2p_l11}^{kl,rs} & \gamma_{2p_l21}^{kl,rs} & \cdots & \gamma_{2p_l1p_r1}^{kl,rs} & \cdots & \gamma_{2p_l1p_s}^{kl,rs} & \gamma_{2p_l2p_s}^{kl,rs} & \vdots & \gamma_{2p_l1p_r p_s}^{kl,rs} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{p_k p_l11}^{kl,rs} & \gamma_{p_k p_l21}^{kl,rs} & \cdots & \gamma_{p_k p_l1p_r1}^{kl,rs} & \cdots & \gamma_{p_k p_l1p_s}^{kl,rs} & \gamma_{p_k p_l2p_s}^{kl,rs} & \vdots & \gamma_{p_k p_l1p_r p_s}^{kl,rs} \end{bmatrix}, \quad (20)$$

avec

$$\gamma_{ijpq}^{kl,rs} = E(< \phi(X_{(k)}), e_i^k >_k < \phi(X_r), e_p^{(r)} >_k < \phi(X_l), e_j^{(l)} >_l < \phi(X_s), e_q^{(s)} >_s).$$

Alors, en posant  $d = \sum_{k=1}^K \sum_{l=1}^{k-1} p_k p_l$ , on a :

### **Théorème 3.2.1.**

sous  $\widehat{H}_0$ , la séquence  $n\widehat{S}_n$  converge en distribution, lorsque  $n \longrightarrow +\infty$ , vers  $\widehat{Q} = \mathcal{W}^T \mathcal{W}$ , où  $\mathcal{W}$  est une V.a ayant une distribution normale centrée dans  $R^d$  avec une matrice de covariance  $\Gamma$ .

**Preuve.** D'abord, notons que sur  $\mathcal{H}_0$  nous avons  $\Psi = 0$  et, par conséquent,  $T = 0$ . Ensuite, la distribution asymptotique de  $\widehat{S}_n$  sous  $\mathcal{H}_0$  peut être obtenue Théorème 2.1.1 sur l'estimation et l'étude asymptotique de l'analyse canonique généralisée à noyau qui montre que  $\sqrt{n}\widehat{T}_n$  converge en distribution, lorsque  $n \longrightarrow +\infty$ , vers une v.a. ayant une distribution normale dans  $\mathcal{L}(\mathcal{X})$  de moyenne 0 et de covariance égale à celle de  $Z$ . Puis, puisque la carte  $A \in \mathcal{L}(\mathcal{X}) \mapsto \sum_{k=2}^K \sum_{l=1}^{k-1} \text{tr}(\pi_{kl}(A)\pi_{kl}(A)^*) \in R$  est continue, on en déduit que  $n\widehat{S}_n$  converge en distribution, lorsque  $n \longrightarrow +\infty$ ,  $Q = \sum_{k=2}^K \sum_{l=1}^{k-1} \text{tr}(W_{kl}W_{kl}^*)$ , où  $W_{kl} = \pi_{kl}(U)$ . On peut écrire  $W_{kl} = \sum_{i=1}^{p_k} \sum_{j=1}^{p_l} W_{kl}^{ij} e_j^{(l)} \otimes e_i^{(k)}$  où  $W_{kl}^{ij} = < W_{kl}, e_j^{(l)} \otimes e_i^{(k)} >$ , et en utilisant les égalités  $(u \otimes v)(w \otimes z) = < u, z > w \otimes v$  et  $\text{tr}(u \otimes v) = < u, v >$  (cf. Dauxois et al. (1994)), on a :

$$\begin{aligned} \text{tr}(W_{kl}W_{kl}^*) &= \sum_{i=1}^{p_k} \sum_{j=1}^{p_l} \sum_{r=1}^{p_k} \sum_{s=1}^{p_l} \delta_{js} \delta_{ir} W_{kl}^{ij} W_{kl}^{rs} \\ &= \sum_{i=1}^{p_k} \sum_{j=1}^{p_l} (W_{kl}^{ij})^2 \end{aligned}$$

$$= \mathcal{W}_{k,l}^T \mathcal{W}_{k,l}$$

où  $\delta$  désigne le symbole de kronecker habituel et  $\mathcal{W}_{k,l} = (W_{kl}^{11}, W_{kl}^{21}, \dots, W_{kl}^{1p_l}, W_{kl}^{2p_l}, \dots, W_{kl}^{p_k p_l})$ . Par conséquent,  $Q = \mathcal{W}^T \mathcal{W}$ , où  $\mathcal{W} = (\mathcal{W}_{2,1}^T, \mathcal{W}_{3,1}^T, \mathcal{W}_{3,2}^T, \dots, \mathcal{W}_{K,1}^T, \mathcal{W}_{K,K-1}^T)^T$ .

Puisque  $\mathcal{W}$  est une fonction linéaire de  $U$ , elle est aussi normalement distribuée avec une moyenne 0 et la matrice de covariance  $\Gamma$  définie dans (19) et (20) avec  $\gamma_{ijpq}^{kl,rs} = E(W_{kl}^{ij} W_{rs}^{pq})$ . En outre,

$$\begin{aligned} \gamma_{ijpq}^{kl,rs} &= E(\langle W_{kl}, e_j^{(l)} \otimes e_i^{(k)} \rangle \langle W_{rs}, e_q^{(s)} \otimes e_p^{(r)} \rangle) \\ &= E(\langle (W_{kl} \widetilde{\otimes} W_{rs})(e_j^{(l)} \otimes e_i^{(k)}), e_q^{(s)} \otimes e_p^{(r)} \rangle) \\ &= \langle E(W_{kl} \widetilde{\otimes} W_{rs})(e_j^{(l)} \otimes e_i^{(k)}), e_q^{(s)} \otimes e_p^{(r)} \rangle, \end{aligned}$$

où  $\otimes$  désigne le produit tensoriel lié au produit interne des opérateurs  $\langle A, B \rangle = \text{tr}(AB^*)$ . En rappelant que  $U$  a le même opérateur de covariance que  $\varphi(X \otimes X)$ , c'est-à-dire  $E(U \widetilde{\otimes} U) = E((\varphi(X \otimes X)) \widetilde{\otimes} (\varphi(X \otimes X)))$ , on obtient

$$\begin{aligned} E(W_{kl} \widetilde{\otimes} W_{rs}) &= E((\pi_{kl}(U)) \widetilde{\otimes} (\pi_{rs}(U))) \\ &= \pi_{rs} E(U \widetilde{\otimes} U) \pi_{kl}^* \\ &= \pi_{rs} E((\varphi(\phi(X) \otimes \phi(X))) \widetilde{\otimes} (\varphi(\phi(X) \otimes \phi(X)))) \pi_{kl}^* \\ &= E((\pi_{kl}(\varphi((X) \otimes \phi(X))) \widetilde{\otimes} (\pi_{rs}(\varphi(\phi(X) \otimes \phi(X)))) \\ &= E((\phi(X_l) \otimes \phi(X_k)) \widetilde{\otimes} (\phi(X_s) \otimes \phi(X_r))). \end{aligned}$$

Ainsi,

$$\begin{aligned} E(\mathcal{W}_{kl} \mathcal{W}_{rq}^T) &= E(\langle \phi(X_l) \otimes \phi(X_k), e_j^{(l)} \otimes e_i^{(k)} \rangle \langle \phi(X_s) \otimes \phi(X_r), e_q^{(s)} \otimes e_p^{(r)} \rangle) \\ \gamma_{ijpq}^{kl,rs} &= E(\langle \phi(X_l) \otimes \phi(X_k), e_j^{(l)} \otimes e_i^{(k)} \rangle \langle \phi(X_s) \otimes \phi(X_r), e_q^{(s)} \otimes e_p^{(r)} \rangle) \\ &= E(\langle \phi(X_k), e_i^k \rangle \langle \phi(X_r), e_p^r \rangle \langle \phi(X_l), e_j^l \rangle \langle \phi(X_s), e_q^s \rangle). \end{aligned}$$

■

### 3.3. Test de non-corrélation de variables aleatoire hilbertiennes pour K=2

#### 3.3.1. Hypothèse de test

Soient  $X : (\Omega, \mathcal{A}) \longrightarrow (H_1, B_{H_1})$  et  $Y : (\Omega, \mathcal{A}) \longrightarrow (H_2, B_{H_2})$  deux variables aléatoires hilbertiennes centrées, où  $H_1$  et  $H_2$  sont des espaces de Hilbert. Pour  $l \in \{1, 2\}$ , nous notons  $\langle \cdot, \cdot \rangle_l$  le produit scalaire de  $H_l$  et  $\|\cdot\|_l$  la norme associée. Nous supposons dans ce chapitre que :

$$\mathbb{E}(\|X\|_1^4) < +\infty \quad \text{et} \quad \mathbb{E}(\|Y\|_2^4) < +\infty.$$

Ainsi l'opérateur de covariance croisé de  $X$  et  $Y$  donné par :

$$V_{12}^\phi = \mathbb{E}(\phi(Y) \otimes \phi(X)) = V_{21}^\phi$$

où

$\phi$  une transformation linéaire et continue de  $\mathcal{X}$  vers un autre espace  $\mathcal{H}$ ,  $k$  le noyau reproduisant associé à cette transformation et  $\otimes$  désigne le produit tensoriel défini par :

pour tout  $(u, v) \in H_1 \times H_2$ ,

$$\begin{aligned} v \otimes u : H_2 &\longrightarrow H_1 \\ h &\longmapsto (\langle v, h \rangle_2)u \end{aligned}$$

Nous nous intéressons dans ce chapitre, au test de non-corrélation de  $X$  et  $Y$  dans le cadre du noyaux, c'est-à-dire au test donc l'hypothèse nulle est :

$$\mathcal{H}_0^{a1} : V_{12}^\phi = 0$$

Pour construire ce test, nous allons utiliser l'analyse canonique sur les données fonctionnelles.

Considérons les opérateurs  $V_1^\phi = \mathbb{E}(\phi(X) \otimes \phi(X))$  et  $V_2^\phi = \mathbb{E}(\phi(Y) \otimes \phi(Y))$ . Dans le classique, *Dauxois et Nkiet (1997)* (voir [6]) montre que cette analyse se ramène à l'analyse spectrale de l'opérateur

$$R = (V_1^\phi)^\dagger V_{12}^\phi (V_2^\phi)^\dagger V_{21}^\phi$$

où  $T^\dagger$  est l'inverse généralisée de Moore-Penrose de l'opérateur  $T$

Nous allons considérer cet opérateur dans le cas fonctionnel. Mais *Leurgans et al. (1993)* montre que dans le cas fonctionnel, il faut procéder à une régularisation. Ainsi on va remplacer  $V_1^\phi$  par  $V_l^\phi(\alpha) = V_l^\phi + \alpha I_l$  où  $I_l$  est l'opérateur identité dans  $H_l$ ,  $l \in \{1, 2\}$  et  $\alpha > 0$ . Nous allons considérer l'opérateur

$$R(\alpha) = (V_1^\phi)^\dagger(\alpha) V_{12}^\phi (V_2^\phi)^\dagger(\alpha) V_{21}^\phi$$

La non-corrélation entre  $X$  et  $Y$  est équivalent à la limite de tous les coefficients canoniques obtenus grâce à l'analyse spectrale de l'opérateur  $R(\alpha)$ . Nous avons donc à tester l'hypothèse nulle

$$\mathcal{H}_0 : T(\alpha) = 0$$

contre l'hypothèse alternative

$$\mathcal{H}_1 : T(\alpha) \neq 0,$$

où

$$T(\alpha) = \text{tr}(R(\alpha)).$$

Il nous reste à définir la statistique de test et de déterminer sa loi asymptotique sous l'hypothèse nulle.

### 3.3.2. Statistique de test

Soit  $\{(\phi(X^{(i)}), \phi(Y^{(i)}))\}_{1 \leq i \leq n}$  un échantillon indépendant et identiquement distribué (i.i.d.) de  $(\phi(X), \phi(Y))$ , à partir duquel on peut estimer  $R(\alpha)$  et définir la statistique de test. On pose :

$$\begin{aligned} (V_{12}^\phi)^{(n)} &= \frac{1}{n} \sum_{i=1}^n (\phi(Y^{(i)}) \otimes \phi(X^{(i)})) - \phi(\bar{Y}^n) \otimes \phi(\bar{X}^n) = (V_{21}^\phi)^{(n)*}; \\ (V_1^\phi)^{(n)} &= \frac{1}{n} \sum_{i=1}^n (\phi(X^{(i)}) \otimes \phi(X^{(i)})) - \phi(\bar{X}^n) \otimes \phi(\bar{X}^n), \\ (V_2^\phi)^{(n)} &= \frac{1}{n} \sum_{i=1}^n (\phi(Y^{(i)}) \otimes \phi(Y^{(i)})) - \phi(\bar{Y}^n) \otimes \phi(\bar{Y}^n) \end{aligned}$$

Ainsi, nous estimons  $R(\alpha)$  par l'opérateur aléatoire

$$\widehat{R}^{(n)}(\alpha) = ((V_1^\phi)^{(n)})^\dagger (V_{12}^\phi)^{(n)} ((V_2^\phi)^{(n)})^\dagger (V_{21}^\phi)^{(n)}$$

où  $(V_l^\phi)^{(n)}(\alpha) = (V_l^\phi)^{(n)} + \alpha I_l$ .

Et nous définissons comme statistique de test, la variable aléatoire

$$\widehat{T}^{(n)}(\alpha) = \text{tr}(\widehat{R}^{(n)}(\alpha))$$

### 3.3.3. Loi limite de la statistique de test sous hypothèse nulle

L'on veut obtenir la loi limite de  $\widehat{T}^{(n)}(\alpha)$  sous  $\mathcal{H}_0$  et pour cela, il est nécessaire que celle de  $\widehat{R}^{(n)}(\alpha)$  soit obtenue.

### 3.3.4. Propriétés asymptotiques de $\widehat{R}^{(n)}(\alpha)$

Supposons que nous avons la propriété suivante :

**Pr 1** : Pour  $k = 1, 2$ ,  $\text{rg}((V_k^\phi)^{(n)})$  converge presque sûrement vers  $\text{rg}(V_k^\phi)$ , lorsque  $n \rightarrow +\infty$ .

On a alors :

**Proposition 3.3.1.** La suite  $\widehat{R}^{(n)}(\alpha)$  converge presque sûrement vers  $\widehat{R}(\alpha)$ , lorsque  $n \rightarrow +\infty$ .

**Preuve.** La loi forte de grands nombres donne la convergence presque sûre dans  $H_1$  (resp.  $H_2$ ) ; resp.  $\mathcal{L}(H_2, H_1)$ ,  $\mathcal{L}(H_1)$  ; resp.  $\mathcal{L}(H_2)$  de  $\bar{X}^n$  (resp.  $\bar{Y}^n$  ; resp.  $\frac{1}{n} \sum_{i=1}^n Y^{(i)} \otimes X^{(i)}$  ; resp.  $\frac{1}{n} \sum_{i=1}^n X^{(i)} \otimes X^{(i)}$ ,  $\frac{1}{n} \sum_{i=1}^n Y^{(i)} \otimes Y^{(i)}$  ; vers  $\mathbb{E}(X) = 0$  (resp.  $\mathbb{E}(Y) = 0$  ; resp.  $\mathbb{E}(X \otimes X) = 0$  ;  $\mathbb{E}(Y \otimes Y) = 0$ ). Donc  $(\bar{X}^n \text{ resp. } \bar{Y}^n)$  converge presque sûrement vers 0 (resp. 0). Nous obtenons de ce fait la convergence presque sûre dans  $\mathcal{L}(H_2, H_1)$  ( resp.  $\mathcal{L}(H_1)$  ; resp.  $\mathcal{L}(H_2)$  de  $V_{12}^n$  (resp.  $V_1^n$  ; resp.  $V_2^n$ ) vers  $V_{12}$  (resp.  $V_1$  ; resp.  $V_2$ ). On déduit donc de Nasched (1976) que  $V_1^{(n)\dagger}$  (resp.  $V_2^{(n)\dagger}$ ) converge presque sûrement vers  $V_1^\dagger$  (resp.  $V_2^\dagger$ ), ce qui termine notre preuve.

**Conclusion** :  $\widehat{R}^{(n)}(\alpha) \rightarrow R(\alpha)$  ■

Supposons que  $H_1$  et  $H_2$  sont deux sous espaces d'un espace de Hilbert  $H$  tel que :

$$H = H_1 \oplus H_2 \quad \text{et} \quad H_1 \perp H_2$$

Notons  $\pi_l$ ,  $l \in \{1, 2\}$  le projecteur orthogonal de  $H$  dans  $H_l$ , et  $Z$  un vecteur aléatoire tel que  $X = \pi_1 Z$  et  $Y = \pi_2 Z$

En posant  $V^\phi = \mathbb{E}(\phi(Z) \otimes \phi(Z))$ , on a

$$V_1^\phi = \pi_1 V^\phi \pi_1, \quad V_2^\phi = \pi_2 V^\phi \pi_2, \quad V_{12}^\phi = \pi_1 V^\phi \pi_2$$

avec

$$A_{11} \in \mathcal{L}(\mathbb{R}^p), \quad A_{12} \in \mathcal{L}(\mathbb{R}^q, \mathbb{R}^p), \quad A_{21} \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q), \quad A_{22} \in \mathcal{L}(\mathbb{R}^q).$$

On peut alors considérer les projections canoniques :

$$\pi_{11} : A \in \mathcal{L}(\mathbb{R}^p \times \mathbb{R}^q) \mapsto A_{11} \in \mathcal{L}(\mathbb{R}^p)$$

$$\pi_{12} : A \in \mathcal{L}(\mathbb{R}^p \times \mathbb{R}^q) \mapsto A_{12} \in \mathcal{L}(\mathbb{R}^q \times \mathbb{R}^p)$$

$$\pi_{21} : A \in \mathcal{L}(\mathbb{R}^p \times \mathbb{R}^q) \mapsto A_{21} \in \mathcal{L}(\mathbb{R}^p \times \mathbb{R}^q)$$

$$\pi_{22} : A \in \mathcal{L}(\mathbb{R}^p \times \mathbb{R}^q) \mapsto A_{22} \in \mathcal{L}(\mathbb{R}^q)$$

l'application  $\varphi_1$  de  $(\mathbb{R}^p \times \mathbb{R}^q)$  vers  $\mathcal{L}(\mathbb{R}^p)$  définie par :

$$\begin{aligned} \varphi_1(H) &= -(V_1^\phi)^\dagger \pi_{11}(H)(V_1^\phi)^\dagger V_{12}^\phi (V_2^\phi)^\dagger V_{21}^\phi - \pi_{\ker(V_1^\phi)} \pi_{11}(H)(V_1^\phi)^\dagger V_{12}^\phi (V_2^\phi)^\dagger V_{21}^\phi \\ &\quad + (V_1^\phi)^\dagger \pi_{12}(H)(V_2^\phi)^\dagger V_{21}^\phi - (V_1^\phi)^\dagger V_{12}^\phi (V_2^\phi)^\dagger \pi_{22}(H)(V_2^\phi)^\dagger V_{21}^\phi + (V_1^\phi)^\dagger V_{12}^\phi (V_2^\phi)^\dagger \pi_{21}(H) \end{aligned}$$

et la variable aléatoire à valeurs dans  $\mathbb{R}^p \times \mathbb{R}^q$  définie par :

$$Z = \begin{bmatrix} \phi(X) \\ \phi(Y) \end{bmatrix}$$

On a alors, le résultat suivant :

**Proposition 3.3.2.** lorsque  $n \longrightarrow +\infty$ , la variable aléatoire  $\sqrt{n}(\widehat{R}^{(n)}(\alpha) - R(\alpha))$  converge en loi, dans  $\mathcal{L}(\mathbb{R}^p)$ , vers une variable aléatoire de loi normale centrée et ayant même opérateur de covariance que  $\varphi_1(Z \otimes Z)$ .

**Preuve.** Posons

$$\begin{aligned} S_{11}(\alpha) &= \mathcal{N}_{12} \varphi_\alpha(V_2^\phi) V_{21}^\phi \varphi_\alpha(V_1^\phi) \\ S_{12}(\alpha) &= V_{12}^\phi \varphi'_\alpha(\mathcal{N}_2) V_{21}^\phi \varphi_\alpha(V_1^\phi) \\ S_{13}(\alpha) &= V_{12}^\phi \varphi_\alpha(V_2^\phi) \mathcal{N}_{21} \varphi_\alpha(V_1^\phi) \\ S_{14}(\alpha) &= V_{12}^\phi \varphi_\alpha(V_2^\phi) V_{21}^\phi \varphi'_\alpha(\mathcal{N}_1) \\ S_1(\alpha) &= \sum_{k=1}^4 S_{1k}(\alpha) \end{aligned}$$

où

$$\begin{aligned} \varphi_\alpha &: \mathbb{C} \longrightarrow \mathbb{C} \\ z &\longmapsto \varphi_\alpha(z) = (z - \alpha)^{-1} \end{aligned}$$

avec  $\varphi_\alpha((V_j^\phi)^{(n)})$  qui converge en probabilité vers  $\varphi_\alpha(V_j^\phi)$ ,  $j = 1, 2$ .

Et pour tout  $i, j = 1, 2$ ,  $\mathcal{N}_{ij}$  est un opérateur gaussien de  $L^2(E_i)$  vers  $L^2(E_j)$ , avec  $E_1 = \mathbb{R}^p$  et  $E_2 = \mathbb{R}^q$ ,  $\forall i, j = 1, 2$  et

$$\varphi'_\alpha(\mathcal{N}_j) = - \sum_{k=1}^{\text{rang}(V_j^\phi)} (\lambda_{jk} + \alpha)^{-2} P_{jk} \mathcal{N}_j P_{jk} + \sum_{j \neq k} \overline{(\lambda_{jk} + \alpha)(\lambda_{ji} + \alpha)} P_{ji} \mathcal{N}_j P_{jk}$$



avec  $\{\lambda_{jk}, P_{jk}\}_{k=1}^{rang(V_j^\phi)}$  représentant les valeurs propres et les projecteurs propres de  $V_j^\phi$ ,  $j = 1, 2$ .

$\hookrightarrow$  Montrons que  $\sqrt{n}(\widehat{R}^n(\alpha) - R(\alpha)) \xrightarrow{d} S_1(\alpha)$  où  $S_1(\alpha)$  est défini ci-dessus.

Premièrement, posons

$$\begin{aligned}\widehat{A}_{11}(\alpha) &= [(V_{12}^\phi)^{(n)} - V_{12}^\phi] \varphi_\alpha((V_2^\phi)^{(n)}) V_{21}^{\phi(n)} \varphi_\alpha(V_1^{\phi(n)}) \\ \widehat{A}_{12}(\alpha) &= V_{12}^{\phi(n)} [\varphi_\alpha(V_2^{\phi(n)}) - \varphi_\alpha(V_2^\phi)] V_{21}^{\phi(n)} \varphi_\alpha(V_1^{\phi(n)}) \\ \widehat{A}_{13}(\alpha) &= V_{12}^{\phi(n)} \varphi_\alpha(V_2^{\phi(n)}) [V_{21}^{\phi(n)} - V_{21}^\phi] \varphi_\alpha(V_1^{\phi(n)}) \\ \widehat{A}_{14}(\alpha) &= V_{12}^{\phi(n)} \varphi_\alpha(V_2^{\phi(n)}) V_{21}^{\phi(n)} [\varphi_\alpha(V_1^{\phi(n)}) - \varphi_\alpha(V_1^\phi)]\end{aligned}$$

et

$$\begin{aligned}\widehat{B}_{11}(\alpha) &= V_{12}^\phi [\varphi_\alpha(V_2^{\phi(n)}) V_{21}^{\phi(n)} \varphi_\alpha(V_1^{\phi(n)}) - \varphi_\alpha(V_2^\phi) V_{21}^\phi \varphi_\alpha(V_1^\phi)] \\ \widehat{B}_{12}(\alpha) &= [V_{12}^{\phi(n)} - V_{12}^\phi] \varphi_\alpha(V_2^\phi) [V_{21}^{\phi(n)} \varphi_\alpha(V_1^{\phi(n)}) - V_{21}^\phi \varphi_\alpha(V_1^\phi)] \\ \widehat{B}_{13}(\alpha) &= [V_{12}^{\phi(n)} \varphi_\alpha(V_2^{\phi(n)}) - V_{12}^\phi \varphi_\alpha(V_2^\phi)] V_{21}^\phi [\varphi_\alpha(V_1^{\phi(n)}) - \varphi_\alpha(V_1^\phi)] \\ \widehat{B}_{14}(\alpha) &= [V_{12}^{\phi(n)} \varphi_\alpha(V_2^{\phi(n)}) V_{21}^{\phi(n)} - V_{12}^\phi \varphi_\alpha(V_2^\phi) V_{21}^\phi] \varphi_\alpha(V_1^\phi)\end{aligned}$$

Notons que la différence  $\sqrt{n}(V_1^{\phi(n)}(\alpha) - V_1^\phi(\alpha))$  peut s'exprimer par :

$$\sqrt{n}(V_1^{\phi(n)}(\alpha) - V_1^\phi(\alpha)) = \sqrt{n} \left[ \sum_{j=1}^4 \widehat{A}_{1j}(\alpha) \right] + \sqrt{n} \left[ \sum_{j=1}^4 \widehat{B}_{1j}(\alpha) \right]$$

Or le théorème de continuité et le théorème 2.5.9 énoncé dans la thèse de David King publié en 2009 assure que :

$$\sqrt{n} \left[ \sum_{j=1}^4 \widehat{B}_{1j}(\alpha) \right] = o_p(n^{-1})$$

Par exemple, on a :

$$\widehat{B}_{11}(\alpha) = V_{12}^\phi [\varphi_\alpha(V_2^{\phi(n)}) V_{21}^{\phi(n)} \varphi_\alpha(V_1^{\phi(n)}) - \varphi_\alpha(V_2^\phi) V_{21}^\phi \varphi_\alpha(V_1^\phi)] \xrightarrow{P} S_1(\alpha) - S_1(\alpha) = 0$$

où la vitesse de convergence est  $o_p(n^{-3/2})$ . Les applications des théorèmes 2.5.8, 2.5.9, 3.9 de la même thèse de David King (voir[8]) et du théorème de Slutsky montre que :

$$\sqrt{n} \left[ \sum_{j=1}^4 \widehat{A}_{1j}(\alpha) \right] \xrightarrow{d} S_1(\alpha)$$

Par exemple,  $\widehat{A}_{11}(\alpha)$  est constitué des facteurs  $\sqrt{n}(V_{12}^{\phi(n)} - V_{12}^\phi) \xrightarrow{d} \mathcal{N}_{12}$  multiplié par le facteur

$$\varphi_\alpha(V_2^{\phi(n)}) V_{21}^{\phi(n)} \varphi_\alpha(V_1^{\phi(n)}) \xrightarrow{P} \varphi_\alpha(V_2^\phi) V_{21}^\phi \varphi_\alpha(V_1^\phi)$$

de vitesse de convergence  $0_p(n^{-3/2})$ . Ce qui achève la preuve.

### Conclusion

$$\sqrt{n}(\widehat{R}^n(\alpha) - R(\alpha)) \xrightarrow{d} S_1(\alpha)$$

■

**Remarque 3.3.1.** Posons  $\bar{Z}^n = \frac{1}{n} \sum_{i=1}^n Z^{(i)} = \begin{pmatrix} \phi(\bar{X}^n) \\ \phi(\bar{Y}^n) \end{pmatrix}$ .

On a :

$$\begin{aligned} V_Z^{\phi(n)} &= \frac{1}{n} \sum_{i=1}^n (Z^{(i)} - \bar{Z}^n) \otimes (Z^{(i)} - \bar{Z}^n) \\ &= \begin{pmatrix} V_1^{\phi(n)} & V_{12}^{\phi(n)} \\ V_{21}^{\phi(n)} & V_2^{\phi(n)} \end{pmatrix}. \end{aligned}$$

Posons :  $k^{\phi(n)} = \sqrt{n}(V_Z^{\phi(n)} - V_Z^\phi)$ . D'après la proposition 5 dans Dauxois et al. (voir [6]), lorsque  $h \rightarrow +\infty$ , la suite  $k^{\phi(n)}$  converge en loi, dans  $\mathcal{L}(\mathbb{R}^p \times \mathbb{R}^q)$  vers une V.a.  $k^\phi$ , de la loi normale centrée, avec pour opérateur de covariance :

$$\mathbb{E}((\phi(Z) \otimes \phi(Z) - V_Z^\phi) \otimes (\phi(Z) \otimes \phi(Z) - V_Z^\phi))$$

Sous l'hypothèse nulle,  $V_{12}^\phi = 0$ , ce qui entraîne que  $\varphi(Z \otimes Z) = 0$ . Donc l'opération de covariance de  $\varphi(k^\phi)$  est nulle. Ceci rend inutilisable cette proposition pour une éventuelle inférence statistique. Il nous faut donc une autre loi asymptotique pour  $\hat{R}^{(n)}(\alpha)$ .

**Proposition 3.3.3.** Soit  $\mathcal{H}_0$ , la variable aléatoire  $n\hat{R}^{(n)}(\alpha)$  converge en loi, lorsque  $n \rightarrow +\infty$ , dans  $\mathcal{L}$  vers  $U_1 U_2$ , où  $U_1$  (resp.  $U_2$ ) est une variable aléatoire centrée gaussienne ayant le même opérateur de covariance que  $V_1^{\phi^\dagger}(Y \otimes X)$  (resp.  $V_2^{\phi^\dagger}(X \otimes Y)$ ).

**Preuve.** Posons  $S_1^{\phi(n)} = V_1^{\phi(n)\dagger}(\alpha) V_{12}^{\phi(n)}$  et  $S_2^{\phi(n)} = V_2^{\phi(n)\dagger}(\alpha) V_{21}^{\phi(n)}$ , on a :  $\hat{R}_1^{(n)}(\alpha) = (\sqrt{n} S_1^{\phi(n)}) (\sqrt{n} S_2^{\phi(n)})$ .

Comme  $V_{12}^\phi = 0$ , nous pouvons écrire :

$$\sqrt{n} S_1^{\phi(n)} = V_1^{\phi(n)\dagger}(\alpha) [\sqrt{n} (V_{12}^{\phi(n)} - V_{12}^\phi)] = V_1^{(n)\dagger}(\alpha) \tilde{K}_2^{(n)}$$

où  $\tilde{K}_2^{(n)} = [\sqrt{n} (V_{12}^{\phi(n)} - V_{12}^\phi)]$ . Comme  $(V_1^{\phi^\dagger})^{(n)}(\alpha)$  converge en probabilité vers  $V_1^{\phi^\dagger}(\alpha)$ , lorsque  $n \rightarrow +\infty$ , et d'après le théorème central limite,  $\tilde{K}_2^{(n)}$  converge en loi vers une variable aléatoire normale centrée  $K_2^\phi$  qui a même opérateur de covariance que  $(Y \otimes X)$ . D'après Slutsky,  $\sqrt{n} S_1^{\phi(n)}$  converge en loi vers une variable aléatoire centrée de loi normale  $U_1 = V_1^{\phi^\dagger}(\alpha) K_2^\phi$  qui a même opérateur de covariance que  $V_1^{\phi^\dagger}(\alpha)(Y \otimes X)$ . On montre de la même manière que  $\sqrt{n} S_2^{\phi(n)}$

converge en loi vers une variable aléatoire centrée de loi normale  $U_2 = V_2^{\phi^\dagger}(\alpha)K_2^{\phi*}$  qui a même opérateur de covariance que  $V_2^{\phi^\dagger}(\alpha)(X \otimes Y)$ .

$$\begin{aligned} n\hat{R}^{(n)}(\alpha) - U_1 U_2 &= \left( \sqrt{n}S_1^{\phi(n)} \right) \left( \sqrt{n}S_2^{\phi(n)} \right) - U_1 U_2 \\ &= \left( \left( \sqrt{n}S_1^{\phi(n)} \right) - U_1 \right) \left( \sqrt{n}S_2^{\phi(n)} \right) + U_1 \left( \sqrt{n}S_2^{\phi(n)} - U_2 \right) \end{aligned}$$

$\left( \left( \sqrt{n}S_1^{\phi(n)} \right) - U_1 \right)$  et  $\left( \sqrt{n}S_2^{\phi(n)} - U_2 \right)$  convergent vers 0, on conclut que la variable aléatoire  $n\hat{R}_1^{(n)}(\alpha)$  converge vers  $U_1 U_2$ . ■

### 3.3.5. Loi limite de $\widehat{T}^{(n)}(\alpha)$ sous $\mathcal{H}_0$

Posons  $\Gamma_1 = V_1^{\phi^\dagger} \otimes^K V_2^{\phi^\dagger}$  où  $\otimes^K$  désigne le produit de Kronecker. Considérons l'application linéaire  $\Phi_1 : \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q) \longrightarrow \mathcal{L}(\mathbb{R}^{pq})$  définie par  $\Phi_1(A) = \text{vec}(A)$  où  $\text{vec}$  est l'opérateur de vectorisation d'une matrice ; alors, en notant  $x^T$  la transposée de  $x$ , nous avons :

**Proposition 3.3.4.** Sous  $\mathcal{H}_0$ ,  $n\widehat{T}_1^{(n)}(\alpha)$  converge en loi, lorsque  $n \longrightarrow +\infty$ , vers  $Q_1 = M_1^T \Gamma_1 M_1$  où  $M_1$  est un vecteur aléatoire de loi normale centrée, dans  $\mathbb{R}^{pq}$ , et d'opérateur de covariance égal à celui de  $\Phi_1(Y \otimes X)$ .

**Preuve.** D'après la proposition 3.3.2, nous déduisons que  $n\widehat{T}_1^{(n)}(\alpha)$  converge en loi vers  $Q_1 = \text{tr}(U_1 U_2)$ , car l'opérateur trace est continu sur l'ensemble des opérateurs de rang fini. Il nous reste à prouver que  $Q_1$  a la forme requise. Soient  $\{e_1, \dots, e_p\}$  et  $\{f_1, \dots, f_q\}$  les bases canoniques de  $\mathbb{R}^p$  et  $\mathbb{R}^q$  respectivement, on peut écrire  $K_2$  sous la forme :

$$K_2 = \sum_{i=1}^p \sum_{j=1}^q K_{2,ij} f_j \otimes e_i$$

ce qui entraîne

$$U_1 = \sum_{i=1}^p \sum_{j=1}^q K_{2,ij} f_j \otimes V_1^{\phi^\dagger} e_i \quad \text{et} \quad U_2 = \sum_{i=1}^p \sum_{j=1}^q K_{2,ij} e_i \otimes V_2^{\phi^\dagger} f_j$$

donc

$$\begin{aligned}
Q_1 &= tr(U_1 U_2) \\
&= tr \left( \sum_{i=1}^p \sum_{j=1}^q K_{2,ij} f_j \otimes V_1^{\phi^\dagger} e_i \sum_{k=1}^p \sum_{l=1}^q K_{2,kl} e_k \otimes V_2^{\phi^\dagger} f_l \right) \\
&= tr \left( \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^p \sum_{l=1}^q K_{2,ij} K_{2,kl} (f_j \otimes V_1^{\phi^\dagger} e_i) (e_k \otimes V_2^{\phi^\dagger} f_l) \right) \quad \text{où} \\
&= \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^p \sum_{l=1}^q K_{2,ij} K_{2,kl} tr \left( \langle f_j, V_2^{\phi^\dagger} f_l \rangle_{\mathbb{R}^q} (e_k \otimes V_1^{\phi^\dagger} e_i) \right) \\
&= \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^p \sum_{l=1}^q K_{2,ij} K_{2,kl} \langle f_j, V_2^{\phi^\dagger} f_l \rangle_{\mathbb{R}^q} \langle e_k, V_1^{\phi^\dagger} e_i \rangle_{\mathbb{R}^p} \\
&= M_1^T \Gamma_1 M_1
\end{aligned}$$

$$M_1 = \begin{pmatrix} K_{2,11} \\ K_{2,12} \\ \vdots \\ K_{2,1q} \\ K_{2,21} \\ \vdots \\ K_{2,pq} \end{pmatrix} = vec(K_2) = \Phi_1(K_2).$$

Comme  $\Phi_1$  est linéaire et comme  $K_2$  est une gaussienne centrée de même opérateur de covariance que  $Y \otimes X$ , alors  $M_1$  est un vecteur gaussien centré qui a le même opérateur de covariance que  $\Phi_1(Y \otimes X)$ .

■

### 3.3.6. Construction du test

Soit  $\mathbb{F}_{Q_1}$  la fonction de répartition de  $Q_1$  ; d'après le **paragraphe 1.4** et la *proposition* (3.3.4),  $\mathbb{F}_{Q_1}$  est déterminée à partir des opérateurs  $\Sigma_1$  et  $\Gamma_1$  où  $\Sigma_1 = cov(M_1)$ . Mais, comme en pratique,  $\Sigma_1$  et  $\Gamma_1$  sont inconnus, il est nécessaire de procéder à leur estimation.

Pour l'estimation de  $\Gamma_1$ , nous considérons l'opérateur aléatoire

$$\widehat{\Gamma}_1^{(n)} = \left( V_1^{\phi(n)} \right)^\dagger \otimes^K \left( V_2^{\phi(n)} \right)^\dagger.$$

La convergence presque sûre de  $\widehat{\Gamma}_1^{(n)}$ , lorsque  $n \rightarrow +\infty$ , est facilement déduite du fait que  $W_k^{(n)}$  converge presque sûrement vers  $W_k$ , ( $k = 1, 2$ ) lorsque **Pr1** est satisfaite.

Un estimateur convergent de  $\Sigma_1$  est donné par :

$$\widehat{\Sigma}_1^{(n)} = \frac{1}{n} \sum_{i=1}^n \Phi_1(Z_a^{(i)}) \otimes \Phi_1(Z_a^{(i)}) - \Phi_1(\bar{Z}_a^n) \otimes \Phi_1(\bar{Z}_a^n)$$

$$\text{où } Z_a^{(i)} = \phi(Y^{(i)}) \otimes \phi(X^{(i)}) \quad \text{et} \quad \bar{Z}_a^n = \frac{1}{n} \sum_{i=1}^n Z_a^{(i)}$$

Sa convergence presque sûre provient d'une application évidente de la loi forte des grands nombres.

**Remarque 3.3.2.** Dans tout ce qui précède, nous avons utilisé des notations tensorielles et des opérateurs. Toutefois, la transcription de ces notations en notations matricielles, utile pour une mise en oeuvre pratique de la méthode proposée, est facile à obtenir en utilisant les résultats de Dauxois et al. (1994) dans [6]. Plus précisément, la matrice de l'opérateur  $x \otimes y$  est donnée par  $YX^T$  où  $X = (x_1, \dots, x_p)^T$  (resp.  $Y = (y_1, \dots, y_q)^T$ ) est la représentation matricielle du vecteur  $x$  (resp.  $y$ ) dans la base canonique de  $\mathbb{R}^p$  (resp.  $\mathbb{R}^q$ ).

D'après cette remarque, identifiant les opérateurs et les matrices qui les représentent par rapport aux bases canoniques, on peut écrire :

$$\begin{aligned} (V_{12}^\phi)^{(n)} &= \frac{1}{n} \sum_{i=1}^n (\phi(X^{(i)}) - \phi(\bar{X}^n)) (\phi(Y^{(i)}) - \phi(\bar{Y}^n))^T = ((V_{21}^\phi)^{(n)})^T; \\ (V_1^\phi)^{(n)} &= \frac{1}{n} \sum_{i=1}^n (\phi(X^{(i)}) - \phi(\bar{X}^n)) (\phi(X^{(i)}) - \phi(\bar{X}^n))^T; \\ (V_2^\phi)^{(n)} &= \frac{1}{n} \sum_{i=1}^n (\phi(Y^{(i)}) - \phi(\bar{Y}^n)) (\phi(Y^{(i)}) - \phi(\bar{Y}^n))^T; \\ Z_a^{(i)} &= (\phi(X^{(i)})) (\phi(Y^{(i)}))^T; \\ \widehat{\Sigma}_1^{(n)} &= \frac{1}{n} \sum_{i=1}^n (\Phi_1(Z_a^{(i)}) - \Phi_1(\bar{Z}_a^n)) (\Phi_1(Z_a^{(i)}) - \Phi_1(\bar{Z}_a^n))^T \end{aligned}$$

La mise en oeuvre pratique du test proposé est obtenue par l'algorithme suivant :

**1** Calculer les vecteurs

$$\phi(x_i) = \phi(X^{(i)}) = (\langle \phi_1^{(1)}, \phi(X^{(i)}) \rangle_1, \dots, \langle \phi_p^{(1)}, \phi(X^{(i)}) \rangle_1), \quad i = 1, \dots, n \quad \text{et} \quad \phi(\bar{x}^{(n)}) = \frac{1}{n} \sum_{i=1}^n \phi(x_i)$$

de même que les vecteurs

$$\phi(y_i) = \phi(Y^{(i)}) = (\langle \phi_1^{(2)}, \phi(Y^{(i)}) \rangle_2, \dots, \langle \phi_q^{(2)}, \phi(Y^{(i)}) \rangle_2), \quad i = 1, \dots, n \quad \text{et} \quad \phi(\bar{y}^{(n)}) = \frac{1}{n} \sum_{i=1}^n \phi(y_i)$$

**2** Calculer les matrices

$$\begin{aligned} V_{12}^{\phi(n)} &= \frac{1}{n} \sum_{i=1}^n \left( \phi(x_i) - \phi(\bar{x}^{(n)}) \right) \left( \phi(y_i) - \phi(\bar{y}^{(n)}) \right)^T = V_{21}^{\phi(n)T}; \\ V_1^{\phi(n)} &= \frac{1}{n} \sum_{i=1}^n \left( \phi(x_i) - \phi(\bar{x}^{(n)}) \right) \left( \phi(x_i) - \phi(\bar{x}^{(n)}) \right)^T; \\ V_2^{\phi(n)} &= \frac{1}{n} \sum_{i=1}^n \left( \phi(y_i) - \phi(\bar{y}^{(n)}) \right) \left( \phi(y_i) - \phi(\bar{y}^{(n)}) \right)^T; \end{aligned}$$

**3** Calculer la matrice  $\widehat{R}_1^{\phi(n)}(\alpha) = V_1^{\phi(n)\dagger} V_{12}^{\phi(n)} V_2^{\phi(n)\dagger} V_{21}^{\phi(n)}$  et la statistique de test

$$\widehat{T}_1^{(n)}(\alpha) = \text{tr}(\widehat{R}_1^{(n)}(\alpha)).$$

**4** Calculer la matrice  $\widehat{\Gamma}_1^{(n)} = \left( V_1^{\phi(n)} \right)^\dagger \otimes^K \left( V_2^{\phi(n)} \right)^\dagger$ .

**5** Calculer  $Z_a^{(i)} = \phi(x_i) \phi(y_i)^T$ ,  $z_i = \Phi_1(Z_a^{(i)})$ ,  $\bar{z}^{(n)} = \frac{1}{n} \sum_{i=1}^n z_i$ ,

$$\text{et } \widehat{\Sigma}_1^{(n)} = \frac{1}{n} \sum_{i=1}^n \left( z_i - \bar{z}^{(n)} \right) \left( z_i - \bar{z}^{(n)} \right)^T$$

**6** Calculer la matrice  $\left( \widehat{\Sigma}_1^{(n)} \right)^{1/2} \widehat{\Gamma}_1^{(n)} \left( \widehat{\Sigma}_1^{(n)} \right)^{1/2}$  et déterminer ses valeurs propres  $\varepsilon_1, \dots, \varepsilon_{pq}$ .

**7** Définir, comme au *paragraphe*

**8** Calculer la probabilité critique  $p_c = 1 - \mathbb{F}_{Q_1}(\widehat{T}_1^{(n)}(\alpha))$ . Pour un niveau  $\alpha \in ]0, 1[$ , l'hypothèse de non-corrélation sera rejetée si  $p_c < \alpha$ .

**Remarque 3.3.3.** Il est clair, au vu de cet algorithme, que la nature hilbertienne des  $\phi(X^{(i)})$  et  $\phi(Y^{(i)})$  n'intervient qu'à la première étape lors du calcul des produits scalaires  $\langle \phi_k^{(1)}, \phi(X^{(i)}) \rangle_1$  et  $\langle \phi_l^{(2)}, \phi(Y^{(i)}) \rangle_2$ ,  $k = 1, \dots, p$ ,  $l = 1, \dots, q$ ,  $i = 1, \dots, n$ .

### 3.3.7. Convergence du test

Le résultat suivant assure la convergence du test

**Proposition 3.3.5.** Sous  $\mathcal{H}_1$ , i.e  $T_1(\alpha) \neq 0$  et pour tout  $\beta \in ]0, 1[$ , on a :

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left[ \mathbb{F}_{Q_1}(\widehat{T}_1^{(n)}(\alpha)) > 1 - \beta \right] = 1$$

**Preuve.** Sous l'hypothèse  $\mathcal{H}_1$ ,  $T_1(\alpha) > 0$ , il est possible de considérer un réel  $\epsilon$  tel que  $0 < \epsilon < T_1(\alpha)$ . De l'inégalité

$$\mathbb{P} \left( \left| \widehat{T}_1^{(n)}(\alpha) \right| < \epsilon \right) \leq \mathbb{P} \left( \widehat{T}_1^{(n)}(\alpha) > T_1(\alpha) - \epsilon \right)$$

et du fait que  $\widehat{T}_1^{(n)}(\alpha)$  converge en probabilité, lorsque  $n \rightarrow +\infty$ , vers  $T_1(\alpha)$ , nous déduisons que

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left( \widehat{T}_1^{(n)}(\alpha) > T_1(\alpha) - \epsilon \right) = 1 \quad (3.1)$$

Soit  $n_0$  un entier tel que  $\forall n \geq n_0$ ,  $n^{-1} \mathbb{F}_{Q_1}^{-1}(1 - \beta) < T_1(\alpha) - \epsilon$ , alors pour tout  $n \geq n_0$ , on a

$$\mathbb{P} \left( \widehat{T}_1^{(n)}(\alpha) > T_1(\alpha) - \epsilon \right) \leq \mathbb{P} \left( \widehat{T}_1^{(n)}(\alpha) > n^{-1} \mathbb{F}_{Q_1}^{-1}(1 - \beta) \right)$$

et de 3.1, nous déduisons que

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left( \mathbb{F}_{Q_1} \left( n \widehat{T}_1^{(n)}(\alpha) \right) > 1 - \beta \right) = 1$$

■

**Conclusion :** On vient de construire un test permettant de prouver la non-corrélation entre deux variables  $X$  et  $Y$  dans un espace de Hilbert. La mise en place de ce test est basée entièrement sur l'analyse canonique à noyaux.

### 3.4. APPLICATION AUX DONNÉES FONCTIONNELLES

En guise d'application, on consacre cette section à montrer comment les différentes procédures développées précédemment s'appliquent sur les données fonctionnelles. Soit  $I = [a, b]$  un intervalle de  $\mathbb{R}$ , nous considérons le processus  $\phi(X) = (X_t)_{t \in I}$ , ou  $\phi(X)$  est définie à partir de la transformation  $\phi : X \mapsto \phi(X)$  de  $\mathcal{H}$  vers l'espace transformé  $\mathcal{H}$  dont nous voulons tester la non corrélation. Soit  $(X^{(i)})_{1 \leq i}$  un échantillon i.i.d de  $X$ ; comme c'est souvent le cas en pratique, nous supposons que les courbes  $(X^{(i)})$  sont observées en des points  $t_1, t_2, \dots, t_L$  tel que  $a = t_1 < t_2 < \dots < t_L$ . Nous pouvons donc ranger ces données dans des matrices  $X = (X^{(i)}(t))$  qui contient toutes les observations. Ainsi,  $\phi(X^{(i)})$  sont observés en des points  $t_1, t_2, \dots, t_L$  tel que  $a = t_1 < t_2 < \dots < t_L$ . Ces données sont rangées dans des matrices  $\phi(X) = \phi(X^{(i)}(t))$  qui contient toutes les observations. L'application de notre méthode introduite dans le chapitre précédent à ce type de données est subordonné aux choix effectués pour les espaces de Hilbert  $\mathcal{H}$  dans lequel  $X$  est supposé prendre ses valeurs. En nous référant des choix classiques utilisés dans la littérature, nous nous limitons au cas d'espace de Sobolev pour notre méthode.

---

### 3.4.1. Application de la méthode

Comme le montre notre algorithme, la nature fonctionnelle de  $\phi(X)$  n'intervient qu'à la première étape permettant la mise en oeuvre pratique, en calculant  $\langle \beta_k, \phi(X^i) \rangle$ . Ce calcul nécessite que soient précisés les espaces de Hilbert de travail. Bien qu'il soit moins fréquent dans la littérature de considérer que des données fonctionnelles émanant de variables aléatoires à valeur dans des espaces de Sobolev, la pertinence d'une telle approche a été montrée, par exemple, par [6], ( Champely Deleuze 1995 ) dans la modélisation de données fonctionnelles provient d'une étude en foresterie . Dans ce cas nous utiliserons des projections sur des sous espaces engendrés par des fonctions B-Splines.

### 3.4.2. Cas de données dans l'espace de Sobolev

Lorsque l'on suppose que l'espace de Sobolev  $H_1 := H^{r-1}(I)$  avec  $r \in \mathcal{N}$ . Nous considérons des projections sur les sous-espaces engendrés par les fonctions B-Splines d'ordre  $r$  considérés dans  $I$ . Plus précisément, on a  $p = r + L$  et on pose  $\beta_k = \beta_k^r$  ( $k = 1, 2, \dots, p$ ) où les  $\beta_k^r$  sont les fonctions B-Splines évoqué ci-avant . Alors on a :

$$\langle \beta_k^1, \phi(X^i) \rangle_{H^p(I)} = \sum_{k=0}^p \int_a^b \frac{d^l}{dt^l} \beta_k^1(t) \frac{d^l}{dt^l} \phi(X^i) dt$$

Les dérivées successives des  $\phi(X^i)$  sont déterminées en calculant en pratique par les formules classiques d'approximation sur la base des observations en  $t_1, \dots, t_L$  , contrairement à celles des fonctions Spline qui sont programmées dans une librairie de R. Ceci offre la possibilité d'obtenir le calcul des intégrales précédentes par des formules classiques de quadrature.

## 3.5. Simulations

On débute cette partie par une simulation faite sur le test de non-correlation dans le paragraphe ci-haut. Nos programmes sont réalisés en utilisant le langage R. Dans cette section, nous présenterons les résultats obtenus. Nous présentons en prélude les résultats sur le niveau empirique, la puissance et enfin de compte la simulation des données réelles par la méthode de l'AC.

### 3.5.1. Le modèle utilisé

Dans le cadre de notre étude, nous présentons le modèle suivant : Soit  $I = [0, 1]$ ,  $X(t) = m(t) + e(t)$  où  $e(t) \rightsquigarrow N(0, 1)$  et  $m(t) = \cos(t + Y\Pi_\gamma/q)$  avec  $q$  le nombre de groupes considérés



et  $\gamma \in [0, 1]$ . soit  $\phi$  une application associée à  $X(t)$ , à tout  $Y \in \mathbb{Y}$ . L'espace image de  $\phi$  est alors défini par  $\phi(X(t)) = m'(t) + e(t)$  où  $m'(t) = \cos(t + \phi(Y)\Pi_\gamma/q)$ . Dans ce qui suit, on génère un échantillon de courbe  $\{(X(i), Y(i))\}_{1 \leq i \leq n}$  i.i.d de  $\{(X, Y)\}$  suivant notre modèle. Nous observons par la suite ces courbes en  $r_0 = 50$  dans l'intervalle  $I$ . Les données de cet intervalle sont stockées dans une matrice  $X$  de taille  $(n, 50)$  où  $n$  est la taille de l'échantillon. On se contente de distinguer dans notre étude, les tailles  $n = 25; 50; 100; 200; 300; 400; 500; 1000$ .

Il apparait que la variable aléatoire  $\phi(X)$  telle que définie en 3.4 est un élément de  $H^p(I)$ ,  $p \in \mathbb{N}$  et  $L^2(I)$ . Nous appliquons dans notre cadre d'étude de test de non-correlation de l'analyse canonique. Nous utilisons les fonctions B-spline d'ordre 2 avec 6 noeuds sur l'intervalle  $I$  et y incluant les bornes. Afin de simplifier le calcul, le produit scalaire est calculé en utilisant des formules classiques pour l'approximation de la dérivée et la méthode de Trapèze pour l'intégrale.

### 3.5.2. Données météorologiques utilisées

Les données météorologiques que nous disposons dans le cadre de notre études sont celles du Cameroun. Elles ont été prises sur le site [https : //crudata.uea.ac.uk/cru/data/hrg/cruts4.04/cruts.2004151855.v4.04/pre](https://crudata.uea.ac.uk/cru/data/hrg/cruts4.04/cruts.2004151855.v4.04/pre) pour les

données de précipitations et

[https : //crudata.uea.ac.uk/cru/data/hrg/cruts4.04/cruts.2004151855.v4.04/tmp](https://crudata.uea.ac.uk/cru/data/hrg/cruts4.04/cruts.2004151855.v4.04/tmp) pour les données de température.

Elles représentent les variations de températures et précipitation de la période 1901-2019 et s'étendent sur toute l'année(janvier-décembre). Ces données sont assez homogènes, de bonne qualité et représentatives. Les différentes stations que nous avons à notre actif pour notre étude sont : Abong-Mbang, Akonolinga, Bafia, Batouri, Bertoua, Douala, Eséka, Garoua, Kaélé, Koundja, Kribi, Lomié, Maroua, Nanga Eboko, Ngambé, Ngaoundéré, Nkongsamba, Poli, Sang-mélina, Tibati, Yaoundé, Yoko. Les stations sont plus réparties au Sud et au centre du Cameroun.

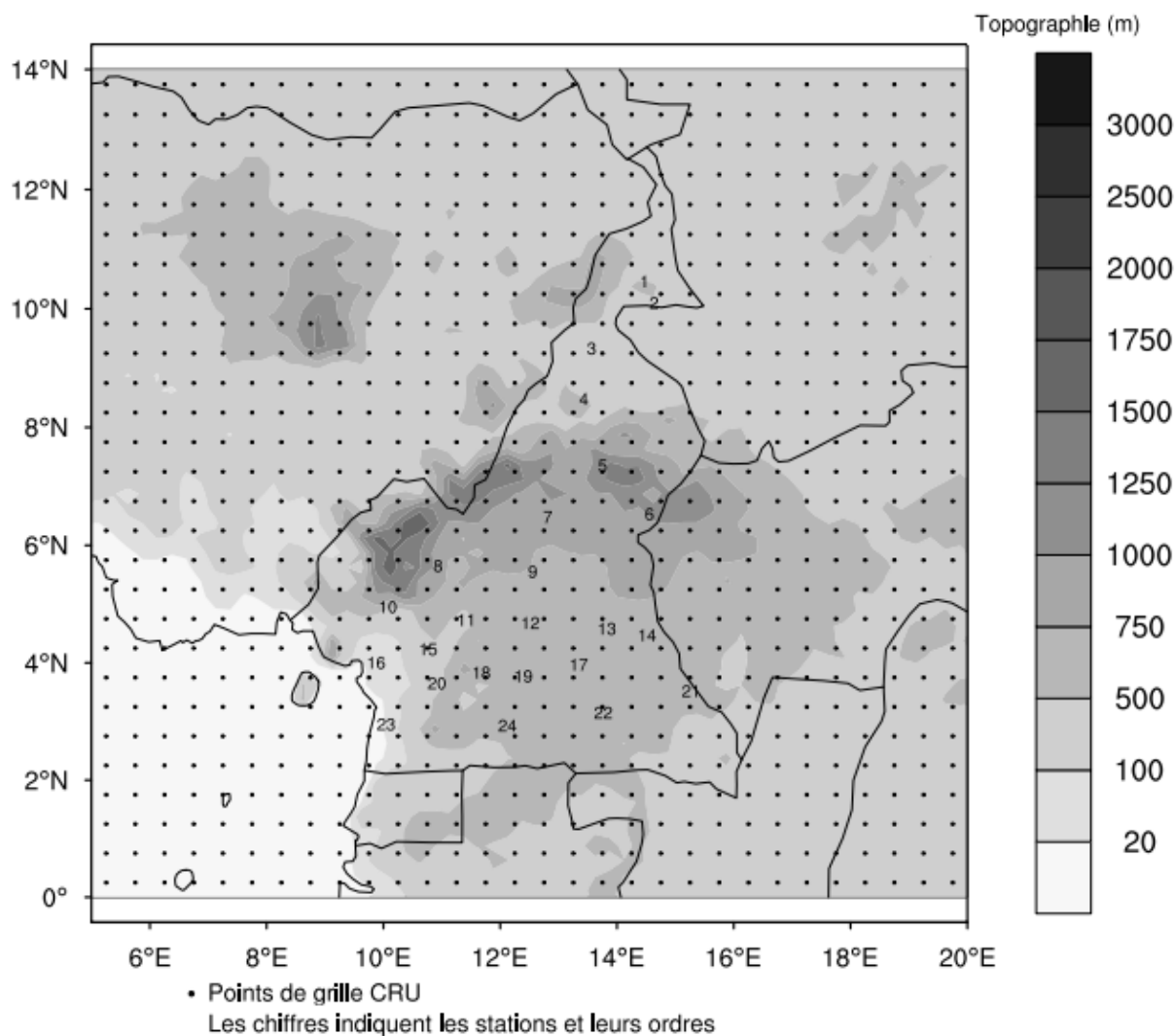


FIGURE 1.1-Position géographique des 24 stations

Sur cette carte, on a des pointillés qui représentent les points de grilles CRU, les chiffres qui représentent les stations concernées. Ainsi, on peut voir qu'on a 24 stations. Le tableau suivant associe à chaque numéro le nom de la station correspondante.

1 ="Abong-Nbang"	9 ="Kaélé"	17 ="Ngaoundéré"
2 ="Akonolinga"	10 ="Koundja"	18 ="Nkongsamba"
3 ="Bafia"	11 ="Kribi"	19 ="Poli"
4 ="Batouri"	12 ="Lomié"	20 ="Sangmélina"
5 ="Bertoua"	13 ="Maroua"	21 ="Tibati"
6 ="Douala"	14 ="Meiganga"	22 ="Yaoundé"
7 ="Eséka"	15 ="Nanga-Eboko"	23 ="Yokadouma"
8 ="Garoua"	16 ="Ngambé"	24 ="Yoko"

TABLE 1.1-Correspondance nom et numéro de stations

```
> summary(data)
```

AbongB		Akono1		Bafia		Batour	
Min.	: 0.000	Min.	: 0.000	Min.	: 0.00	Min.	: 0.00
1st Qu.:	5.827	1st Qu.:	5.747	1st Qu.:	4.53	1st Qu.:	4.98
Median	:12.525	Median	:12.680	Median	:15.30	Median	:12.94
Mean	:12.974	Mean	:13.404	Mean	:14.97	Mean	:12.57
3rd Qu.:	18.367	3rd Qu.:	19.142	3rd Qu.:	22.16	3rd Qu.:	18.00
Max.	:37.560	Max.	:38.140	Max.	:52.78	Max.	:37.15
Bertou		Douala		Es.ka		Garoua	
Min.	: 0.00	Min.	: 0.000	Length:	1428	Min.	: 0.000
1st Qu.:	5.27	1st Qu.:	7.968	Class	:character	1st Qu.:	0.000
Median	:13.16	Median	: 21.775	Mode	:character	Median	: 5.355
Mean	:12.71	Mean	: 22.813			Mean	: 8.576
3rd Qu.:	18.20	3rd Qu.:	33.782			3rd Qu.:	16.337
Max.	:37.38	Max.	:113.900			Max.	:38.510
Kaele		Koundj		Kribi		Lomi.	
Min.	: 0.000	Min.	: 0.000	Min.	: 0.00	Min.	: 0.000
1st Qu.:	0.000	1st Qu.:	3.035	1st Qu.:	11.37	1st Qu.:	6.838
Median	: 2.590	Median	:17.360	Median	:18.46	Median	:12.605
Mean	: 6.918	Mean	:16.481	Mean	:19.53	Mean	:13.118
3rd Qu.:	12.870	3rd Qu.:	25.988	3rd Qu.:	26.22	3rd Qu.:	18.073
Max.	:34.970	Max.	:54.010	Max.	:79.59	Max.	:45.580

Min. : 0.000	Min. : 0.00	Min. : 0.000	Min. : 0.000
1st Qu.: 0.000	1st Qu.: 0.70	1st Qu.: 4.643	1st Qu.: 6.205
Median : 2.590	Median :12.57	Median :12.915	Median :18.285
Mean : 6.918	Mean :12.22	Mean :13.223	Mean :18.136
3rd Qu.:12.870	3rd Qu.:21.18	3rd Qu.:19.185	3rd Qu.:26.273
Max. :34.970	Max. :39.26	Max. :37.910	Max. :64.920
Ngaoun	Nkongs	Polì	Sangm.
Min. : 0.00	Min. : 0.000	Length:1428	Min. : 0.000
1st Qu.: 0.35	1st Qu.: 6.695	Class :character	1st Qu.: 6.465
Median :13.03	Median : 20.895	Mode :character	Median :13.515
Mean :12.34	Mean : 21.640		Mean :13.679
3rd Qu.:21.65	3rd Qu.: 32.663		3rd Qu.:19.410
Max. :37.96	Max. :102.800		Max. :42.260
Tibati	Yaounde	Yokadouma	Yoko
Min. : 0.00	Min. : 0.000	Min. : 0.020	Min. : 0.00
1st Qu.: 1.10	1st Qu.: 6.105	1st Qu.: 6.558	1st Qu.: 2.37
Median :14.52	Median :13.510	Median :13.360	Median :14.51
Mean :13.17	Mean :14.163	Mean :12.779	Mean :13.47
3rd Qu.:22.69	3rd Qu.:20.282	3rd Qu.:17.550	3rd Qu.:21.33
Max. :38.32	Max. :43.130	Max. :35.870	Max. :42.52

Figure 1.2 - Caractérisation des variables météorologiques (Précipitations)

A travers la figure (FIGURE 1.3) ci-dessous, nous visualisons de manière globale la tendance des précipitations

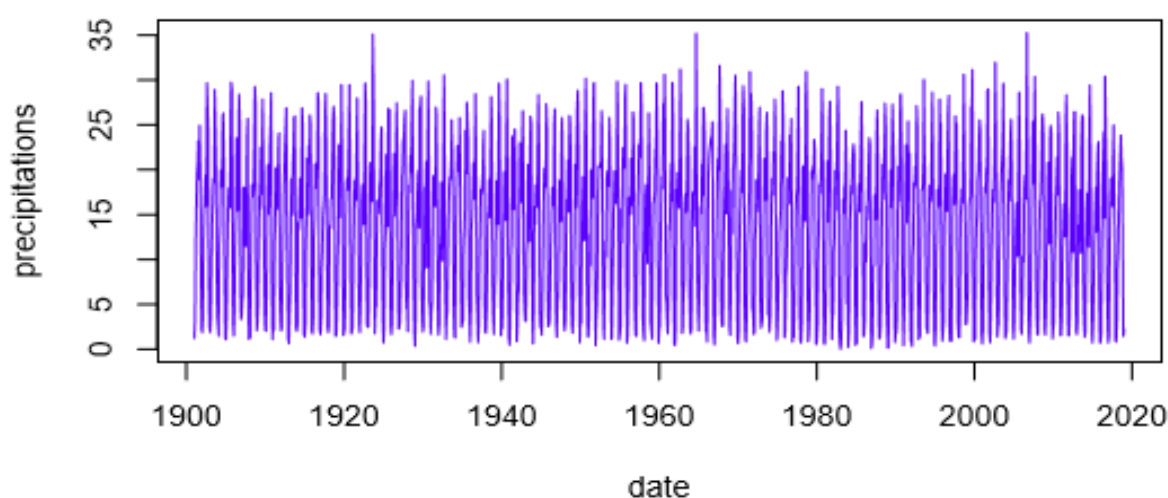


Figure 1.3 - Tendance des précipitations

On constate par cette visualisation une alternance entre les périodes de fortes précipitations et celles de faibles précipitations au fil du temps.

```
> summary(data)
```

Abong		Akono		Bafia		Batouri	
Min.	:21.60	Min.	:21.40	Min.	:21.30	Min.	:21.00
1st Qu.	:22.90	1st Qu.	:22.50	1st Qu.	:22.70	1st Qu.	:22.60
Median	:23.40	Median	:22.90	Median	:23.20	Median	:23.10
Mean	:23.74	Mean	:23.32	Mean	:23.54	Mean	:23.47
3rd Qu.	:24.60	3rd Qu.	:24.10	3rd Qu.	:24.30	3rd Qu.	:24.23
Max.	:27.10	Max.	:26.70	Max.	:26.90	Max.	:26.80

Bertoua		Douala		Es.ka		Garoua	
Min.	:21.20	Min.	:19.90	Min.	:18.30	Min.	:22.90
1st Qu.	:22.68	1st Qu.	:21.40	1st Qu.	:19.60	1st Qu.	:26.50
Median	:23.20	Median	:22.10	Median	:20.20	Median	:27.60
Mean	:23.52	Mean	:22.36	Mean	:20.51	Mean	:28.19
3rd Qu.	:24.30	3rd Qu.	:23.20	3rd Qu.	:21.30	3rd Qu.	:29.93
Max.	:26.80	Max.	:25.60	Max.	:23.70	Max.	:34.00

Ka.1.		Koundja		Kribi		Lomi.	
Min.	:21.60	Min.	:18.30	Min.	:19.90	Min.	:21.20
1st Qu.	:26.00	1st Qu.	:19.60	1st Qu.	:21.40	1st Qu.	:22.68
Median	:27.10	Median	:20.20	Median	:22.10	Median	:23.20
Mean	:27.75	Mean	:20.51	Mean	:22.36	Mean	:23.52
3rd Qu.	:29.62	3rd Qu.	:21.30	3rd Qu.	:23.20	3rd Qu.	:24.30
Max.	:34.00	Max.	:23.70	Max.	:25.60	Max.	:26.80

Maroua	Meiganga	Nanga	Ngamb.
Min. :21.60	Min. :18.20	Min. :21.40	Min. :18.30
1st Qu.:26.00	1st Qu.:20.50	1st Qu.:22.50	1st Qu.:19.60
Median :27.10	Median :21.00	Median :22.90	Median :20.20
Mean :27.75	Mean :21.41	Mean :23.32	Mean :20.51
3rd Qu.:29.62	3rd Qu.:22.12	3rd Qu.:24.10	3rd Qu.:21.30
Max. :34.00	Max. :25.20	Max. :26.70	Max. :23.70
Ngaound.r.	Nkongsamb	Poli	Sangm.lima
Min. :19.00	Min. :19.90	Min. :22.40	Min. :21.40
1st Qu.:21.30	1st Qu.:21.40	1st Qu.:25.60	1st Qu.:22.50
Median :21.90	Median :22.10	Median :26.70	Median :22.90
Mean :22.19	Mean :22.36	Mean :27.17	Mean :23.32
3rd Qu.:23.00	3rd Qu.:23.20	3rd Qu.:28.70	3rd Qu.:24.10
Max. :26.10	Max. :25.60	Max. :32.50	Max. :26.70
Tibati	Yaound.	Yokadouma	Yoko
Min. :20.80	Min. :21.40	Min. :20.50	Min. :21.40
1st Qu.:22.30	1st Qu.:22.50	1st Qu.:22.10	1st Qu.:22.50
Median :22.80	Median :23.10	Median :22.90	Median :22.90
Mean :23.21	Mean :23.44	Mean :23.07	Mean :23.32
3rd Qu.:23.90	3rd Qu.:24.30	3rd Qu.:24.00	3rd Qu.:24.10
Max. :27.00	Max. :26.80	Max. :26.50	Max. :26.70

Figure 1.4 - Caractérisation des variables météorologiques (Températures)

A travers la figure (FIGURE 1.5)) ci-dessous, nous visualisons de manière globale la tendance des températures

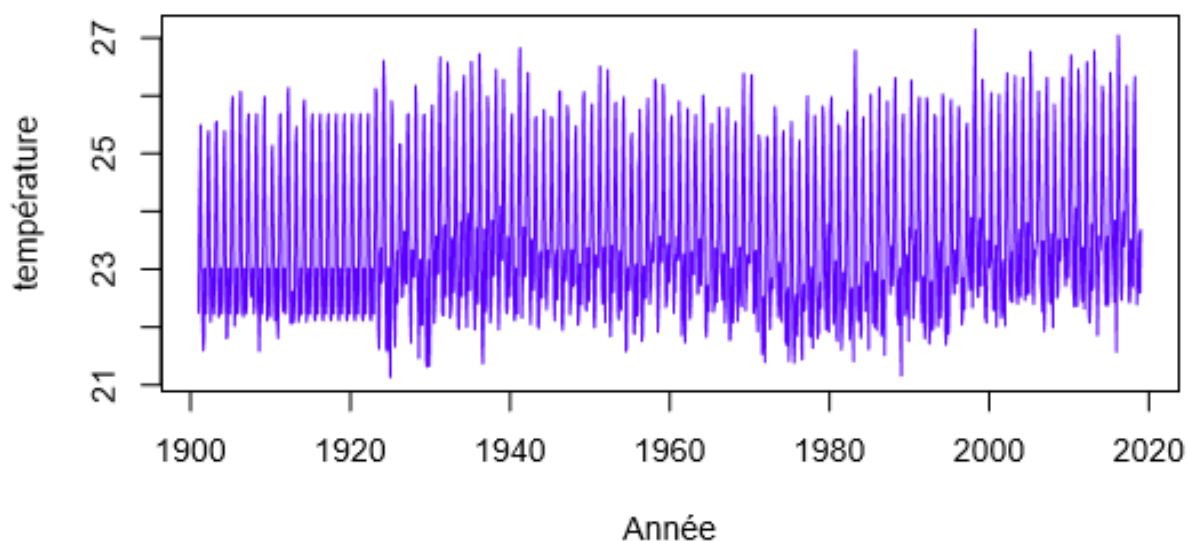


Figure 1.5 - Tendance des précipitations

---

Sur cette figure, on peut voir une alternance entre les périodes de température élevée et les périodes de faible température

### 3.6. Approche fonctionnelle

Le principe de base de l'Analyse de Données Fonctionnelles (ADF) d'après Ramsay et Silverman [17], est de considérer que les individus étudiés sont décrits par des fonctions plutôt que par des vecteurs de  $\mathbb{R}^n$ . En d'autres termes, chaque variable observée est à valeurs fonctionnelles plutôt que réelles. Ce principe est naturel dans de nombreux cas, comme l'illustrent les exemples suivants :

- l'étude du climat est un domaine dans lequel la modélisation fonctionnelle est particulièrement adaptée. En effet, les "individu" observés présentent systématiquement une variabilité temporelle et/ou spatiale. L'exemple le plus simple est celui de l'étude du climat dans un pays donné, par l'intermédiaire des variations de la température mensuelle moyenne au cours de l'année en différents points du pays. Chaque station météo est alors un individu qui est décrit par une fonction qui à chaque mois associe la température moyenne observée (cf Ramsay et Silverman [7]).
- On peut aussi observer une zone géographique avec deux échelles temporelles. On peut par exemple étudier l'évolution de diverses mesures géophysiques sur une zone donnée pendant un an. Chaque année d'observation correspond à un individu et on étudie la série temporelle des individus, comme dans la modélisation auto-régressive fonctionnelle du phénomène *El Niño* proposée dans Besse et al. [20] ;
- à partir des exemples climatiques, on constate que l'ADF est bien adaptée aux problèmes dans lesquels les individus étudiés présentent une variabilité temporelle. En dehors de la météo, on peut citer un des exemples de Ramsay et Silverman [17], à savoir l'étude de la croissance d'enfants : chaque enfant est représenté par une fonction qui à l'âge associe sa taille. Un autre exemple est donné dans Rice et Wu [21], article qui présente l'exemple de l'évolution du nombre de lymphocytes T4 dans le sang de malades atteints du SIDA. L'ADF permet souvent de résumer les données en remplaçant les observations réelles par une modélisation fonctionnelle simple, de traiter des données difficiles à soumettre à des méthodes classiques, même en négligeant les problèmes de volume ou de prise en compte de la structure.

#### Visualisation des courbes de données

Pour obtenir ces courbes de données, nous avons transformé les données de chaque station pour



avoir la courbe d'évolution de la station. L'approche fonctionnelle stipule que si  $Y_{ij} = X_i(t_{ij}) + \epsilon_{ij}$  est la valeur de la température ou des précipitations de la station  $i$  pendant le jour  $j$ , cette valeur peut être retrouvée avec la fonction d'évolution  $X_i$  appliquée au temps  $t_{ij}$  et  $\epsilon_{ij}$  est l'erreur d'approximation de cette fonction au temps  $t_{ij}$  qui suit la loi normale de moyenne 0 et de variance constante. Nous nous intéressons en ce moment aux fonctions  $X_i$  encore appelées fonctions d'évolution ou courbes lisses (courbe d'évolution de la température ou des précipitations de la station  $i$ ). on a :

$$X_i = \sum_{k=1}^K C_{ik} \phi_k(t)$$

Où :

$K$  est le nombre de base

$\phi_k$  sont les fonctions B-splines cubiques

$C_{ik}$  sont les coefficients

Ce qui est également important, c'est le choix des fonctions de base. Nous projectons les données sur les fonctions de base selon leur nature. Ici nous avons choisi les fonctions B-splines en raison de leurs bonnes caractéristiques (mise en oeuvre simple, peuvent être dérivées et intégrées). Nous utilisons deux figures (FIGURE 1.6, FIGURE 1.7 et FIGURE 1.8) pour la visualisation des données des précipitations et des données de température :

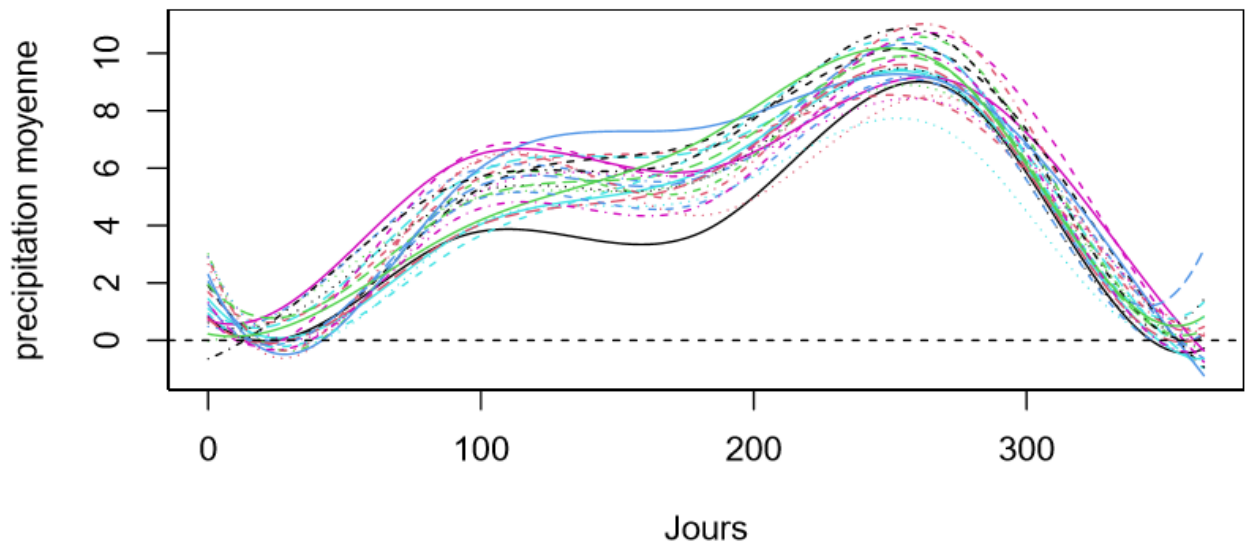


Figure 1.6-Visualisation des courbes de précipitations des 21 stations situées dans la partie sud du Cameroun.

Cette première Figure (FIGURE 1.6) nous permet de visualiser les courbes des des précipita-



tions des 21 stations situées dans la partie Sud du Cameroun. On peut bien voir à travers ces courbes que les précipitations sont bimodales. La première saison communément appelée "petite saison de pluies" allant de la moitié du mois de Mars au mois de Avril voir début du mois de Mai, et la seconde saison communément appelée "grande saison de pluies" qui va du mois de Juillet au mois d'octobre :

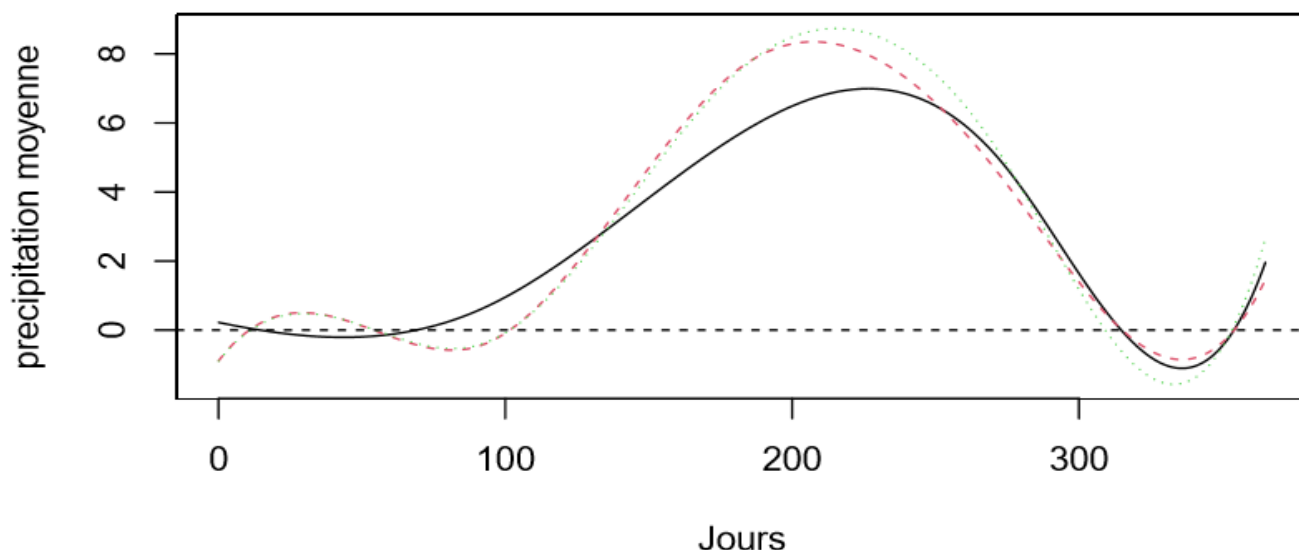


Figure 1.7 - Visualisation des courbes de précipitations des 3 stations situées dans la partie nord du Cameroun

Sur la figure (FIGURE 1.7) ci-dessus nous visualisons les courbes des précipitations des trois stations situées dans la partie septentrionale du Cameroun. On constate que dans cette partie du pays on a une seule saison de pluies.

La figure (FIGURE 1.8) ci-dessous présente une visualisation des courbes des données de température :

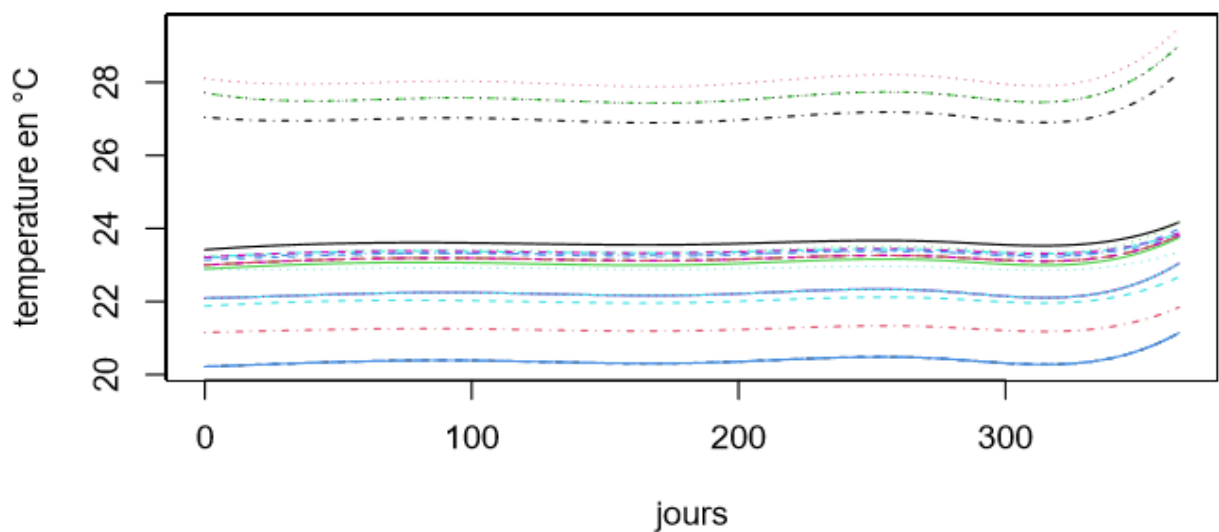


Figure 1.8 - Visualisation des courbes de température

Sur cette figure (FIGURE 1.8), on remarque que trois courbes sont placées au dessus des autres courbes. ces courbes représentent les trois stations des regions du sahel à savoir "Garoua", "Maroua" et "Kaelé", car dans cette partie du pays, les températures sont plus élevées.

## 3.7. Implémentation sur des données réelles

Dans cette partie nous appliquons notre test sur des données météorologiques décrites plus haut dans ce même chapitre. Ainsi, nous montrons à travers un test de non-corrélation l'existence d'un lien entre les précipitations et la température puis nous retrouvons le même resultat avec le test que nous avons proposé.

### 3.7.1. Test de non-corrélation

Nous effectuons un test de non-corrélation sur ces deux variables suivant les étapes ci après :

#### ♣ Choix des variables du test

A l'aide de la fonction "Summary" du logiciel R, nous obtenons les précipitations moyennes et les températures moyennes par mois. Ainsi, nous constituons deux Vecteurs : Un vecteur Pm qui contient les précipitations moyennes de chaque mois et un vecteur Tm qui contient les températures moyennes de chaque mois.

#### ♣ Formulation des hypothèses du test

Hypothèse nulle  $H_0$  : " il n'y a pas de corrélation entre les variables température et précipita-

---

tions", autrement dit, les phénomènes de température et précipitations ne sont pas liés. Cette hypothèse nulle sera opposée à l'hypothèse alternative,  $H_1$  : "il existe un lien entre les variables température et précipitations".

#### ♣ Seuil de signification

Nous effectuons au risque 5 pourcent, c'est-à-dire que l'hypothèse nulle sera rejetée si  $pe < 0.05$  où  $pe$  désigne la "p-value". Ce qui signifie que la probabilité de rejeter l'hypothèse nulle alors qu'elle est vraie est de 0.05.

#### ♣ Degré de liberté $\nu$

le degré de liberté est déterminé par la formule  $\nu = n - 2$  où  $n$  est la taille de l'échantillon. Ainsi on a :  $\nu = 1426$

#### ♣ Décision

La fonction "cor.test()" du langage de programmation R nous propose trois tests de non-corrélation à savoir : le test de non-corrélation de Pearson, le test de non-corrélation de Kendall et le test de non-corrélation de Spearman. Nous avons à l'aide de cette fonction, réalisé un test de non-corrélation sur nos variables  $P_m$  et  $T_m$  et on a obtenu de ces trois méthodes que ces variables sont liées. En utilisant la méthode de Pearson pour effectuer le test de non-corrélation, on obtient les informations suivantes :

$P - value = 9.483 * 10^{-13}$ , le nombre de degrés de liberté  $df = 1426$ , le coefficient de corrélation  $cor = -0.187351$  et la décision "alternative hypothesis true".

**Conclusion** : On peut donc conclure qu'il existe un lien entre les précipitations et la température.

### 3.7.2. Utilisation de la méthode proposée

Après avoir montré à l'aide du test d'indépendance du Chi-deux l'existence d'une liaison entre les températures et les précipitations, nous retrouvons le même résultat en utilisant le test que nous avons proposé.

#### ♣ Définition des variables

Nous calculons les précipitations moyennes et les températures moyennes de chaque année et Nous désignons par  $X$  et  $Y$  les vecteurs contenant respectivement les précipitations moyennes et les températures moyennes.

#### ♣ variables fonctionnelles

---

Nous utilisons les courbes avec 50 noeuds d'observations sur un intervalles  $I = I_1 = I_2 = [0, 1]$ . Ces courbes qui constituent nos variables fonctionnelles  $X$  et  $Y$ .

#### ♣ Formulation des hypothèses du test

Hypothèse nulle  $H_0$  : " il y a indépendance entre les variables températures et précipitations", autrement dit, les phénomènes de température et précipitations ne sont pas liés ( $V_{12} = 0$ ). Cette hypothèse nulle sera opposée à l'hypothèse alternative,  $H_1$  : "il existe un lien entre les variables température et précipitations".

#### ♣ Seuil de signification

Nous effectuons également au risque  $\alpha = 0.05$ . On va alors rejeter l'hypothèse nulle si la  $pvalue(p_e)$  est inférieur à 0.05.

#### ♣ Nombre de degrés de liberté

Comme nous avons indiqué au paragraphe(2.3), nous déterminons le nombre de degrés de liberté par la formule

$$\nu = \frac{\left( \sum_{j=1}^p \lambda_j \right)^2}{\left( \sum_{j=1}^p \lambda_j^2 \right)}$$

ainsi on obtient le résultat  $\nu = 8$ . Nous effectuons donc un test à 8 degrés de liberté.

#### ♣ La p-value

La p-value est donnée par :  $p_c = 1 - \mathbb{F}_{\mathcal{G}}(\hat{T}^n)$

comme indiqué dans l'algorithme de construction du test. Ainsi, après exécution du test nous obtenons  $p_c = 0.006724204$

#### ♣ Décision

la p-valeur trouvée est inférieure au seuil de signification( $p_c = 0.006724204$ ), on rejette l'hypothèse nulle en faveur de l'hypothèse alternative. Nous pouvons donc conclure que les phénomènes de précipitation et des températures sont liées . Nous pouvons donc conclure que le test que nous avons proposé fonctionne normalement.

**Conclusion :** La p-value étant égale à 0.006724204 d'après le résultat de nos simulations, la décision qui est le rejet de l'hypothèse nulle est donc accepté.

---

---

## ✠ Conclusion et perspective ✠

---

Dans ce mémoire, nous avons proposé une généralisation de l'analyse canonique à noyau sur des variables fonctionnelles. Cette analyse est définie à partir d'un problème de maximisation sous des contraintes spécifiques, et on montre qu'elle est obtenue à partir de l'analyse spectrale d'un opérateur approprié. Nous introduisons en premier lieu un chapitre rappelant les outils mathématiques qui sont utiles dans la suite de notre manuscrit. Dans la suite, nous abordons le problème de l'estimation de l'ACG à noyau. Plus précisément, nous introduisons des estimateurs basés sur des opérateurs de covariance empiriques. Ensuite, la cohérence des estimateurs obtenus est prouvée. En outre, nous dérivons la distribution asymptotique de l'estimateur utilisé de l'opérateur susmentionné, ainsi que celle de l'estimateur du vecteur des coefficients canoniques dans le cas général. En dernier temps, on consacre cette partie pour les simulations et les application pour évaluer la performance de nos travaux. Pour cela, on construit un test de non-corrélation mutuelle entre les variables aléatoires impliquées dans l'ACG à noyau. Les résultats obtenus pour la théorie asymptotique de l'ACG à noyau sont ensuite utilisés afin de dériver la distribution asymptotique de la statistique de test utilisée sous l'hypothèse nulle. Nous avons également dans ce dernier chapitre utilisé le test proposé pour établir l'existence d'une liaison entre les phénomènes de précipitations et de températures. Nous avons réalisé ces simulations avec le logiciel R.

Nous comptons dans la suite de nos recherches, construire un test de non-corrélation sur des données fonctionnelles en utilisant la méthode des projections dans des sous-espaces de dimension finies pour évaluer la performance de nos travaux.

---

---

## ✠ Bibliographie ✠

---

- [1] Aghoukeng. G. Inférence Statistique sur des données fonctionnelles : non-corrélation, homogénéité, indépendance. Mémoire de thèse de doctorat, Université de Yaoundé 1 - Université de sciences et technique de Masuku, 2010.
- [2] Boor de C : A pratical guide to splines, vol. 27 de Applied Mathematical Sciences. SpringerVerlag, New York, revised édn, 2001.
- [3] Besse. P. Approximation spline de l'analyse en composantes principales d'une variablealéatoirehilbertienne. AnnalesdelafacultédessciencesdeToulouse5esérie.12(2), 329-349, 1991.
- [4] Champely.S, and Deleuze Ch. Co-inertie de deux variables aléatoires hilbertiennes et approximation B-spline application en foresterie. Revue de statistique appliquée, 43, 91-107, 1991.
- [5] Cuevas, A.,Febrero,M., Fraiman. R. An anova test for functional data. Comput. Statist. Data Anal. 47, 111-122, 2004.
- [6] Dauxois, J., Guy. M.K. Canonical analysis of two Euclidean subspaces and its applications. Linear Algebra Appl. 264, 355-388, 1997.
- [7] Dauxois, J, Pousse, A. Une extension de l'analyse canonique. Quelques applications. In Annales de L'LHP Probabilités et statistiques. 1975. p. 355-379, 1975.
- [8] Ferraty, F., Vieu, Ph. Nonparametric functional data analysis : theory and practice. Springer, New York, 2006.
- [9] Guy. M.K. Asymptotic theorie of multiple-set linear canonical analysis. Mathematical Methods of Statistics, 26(3), 196-211, 2017.
- [10] Leurgans. S, Moyeed. R, et Silverman. B. Canonical correlation analysis when the data are curves. Journal of the Royal Statistical Society B, 55 (3) :725 ?740, 1993.

- [11] Maya. K. Méthodes à noyaux en reconnaissance de formes, prédiction et classification. Applications aux biosignaux. Sciences de l'ingénieur [physics]. Université de Technologie de Troyes, 2012.
- [12] Mika. S. Kernalgorithmen zur nichtlinearen signalverarbeitung in merkmalsräumen (kernel algorithms for nonlinear signal processing in feature spaces). Masters thesis, Technische Universität Berlin, Berlin, Germany, November, (ACP), 1998.
- [13] Mika. S, Rätsch. G, Weston. G, Schölkopf. B, and Müller. K.R. Fisher discriminant analysis with kernels. In Y. H. Hu, J. Larsen, E. Wilson, and S. Douglas, editors, Advances in neural networks for signal processing, pages 41,48, San Mateo, CA, USA. Morgan Kaufmann, (AD), 1999.
- [14] Mika. S. Kernel Fisher Discriminants. Phd thesis, University of Technology, Berlin, October, (AD), 2002.
- [15] Ramsay, J.O., Silverman, B.W. Functional Data Analysis. Springer, New york, 2005.
- [16] Roi. D. Analyse de corrélation canonique des données fonctionnelles. Université d'Etat de l'Arizona, 2009.
- [17] Schumaker. L.L, Spline functions : Basic Theory. Wiley, New York, 1981.
- [18] Scholkopf. B, Smola. A.J, and Müller. K.R. Nonlinear component analysis as a kernel eigenvalue problem. Neural Computation, 10(5) :1299,1319, (ACP), 1998.
- [19] Scholkopf, A.J. Smola, and Müller. K.R. Kernel principal component analysis. Advances in kernel methods : support vector learning, pages 327,352, (ACP), 1999.
- [20] Silverman. B, Ramsay . Functional Data Analysis. Springer Series in Statistics, Springer Verlag, June 1997.
- [21] Springer-Verlag. Jolliffe. I. T. Principal Component Analysis. New York, NY, USA, (ACP), 1986.
- [22] Wahba. G. Spline Models for Observationalt Data. Society for Industrial and Applied Math (SIAM), Philadelphia, 1990.