

# Rapport de troisième année de Magistère

Dimitri LEW

"Tant plus je vois que vous blâmez  
Sa noble discipline,  
plus à l'aimer vous enflammez  
ma volonté encline.  
Le ciel orné de tels flambeaux  
n'est-il point admirable ?  
La notice de corps si beaux  
n'est-elle point désirable ?  
Et que chaut-il à qui l'honore  
qu'elle soit méprisée,  
Science, de cil qui l'ignore,  
est toujours condamnée."  
J. Peletier du Mans.

# Table des matières

<b>I</b>	<b>Cursus de Magistère, domaine de recherche et apprentissage hors mur</b>	<b>5</b>
<b>1</b>	<b>Cursus de Magistère</b>	<b>7</b>
<b>2</b>	<b>Domaine de Recherche : Apprentissage et Méthodes de Régression</b>	<b>13</b>
<b>3</b>	<b>LSQ</b>	<b>25</b>
<b>II</b>	<b>TERs</b>	<b>39</b>
<b>4</b>	<b>Du théorème de GOLAB</b>	<b>41</b>
<b>5</b>	<b>Reconstruction de mesures discrètes</b>	<b>67</b>
<b>III</b>	<b>Rapport de stage de Master 2</b>	<b>81</b>



## Première partie

# Cursus de Magistère, domaine de recherche et apprentissage hors mur



Chapitre 1

Cursus de Magistère

## Avertissement

Avant de rentrer plus précisément dans la description du cursus de magistère que j'ai eu la chance de suivre ces trois dernières années, il est important de noter, je pense, la difficulté de cet exercice. En effet il conduit l'étudiant inexpérimenté (que je suis) à faire un inventaire plus ou moins exhaustif des matières suivies et de donner avec plus ou moins de précisions les conséquences sur les choix d'orientations, tout ceci menant à un texte relativement formel et proche d'un CV.

Pour comprendre les raisons qui m'ont poussé à suivre les études que j'ai suivies et l'implication que j'y ai mis, il faut voir la scolarité post-bac comme un tout. Si mon envie de faire des mathématiques (et de les comprendre) à un bon niveau ne s'est pas manifestée durant mon lycée, c'est en classe préparatoire (notamment grâce à des professeurs impliqués et désireux de ne pas se contenter des limites, très scolaires, des programmes) que j'ai pu apprécier les réflexions mathématiques et leurs conséquences. C'est certainement cet enthousiasme qui m'a donné envie de continuer à approfondir mes connaissances dans le domaine des mathématiques. Cette formation m'a aussi appris l'importance d'avoir un savoir solide dans tous les domaines des mathématiques (et des sciences en générale), cette idée m'aura suivi tout au long de la formation de magistère : si la spécialisation est nécessaire pour obtenir un niveau suffisant dans le travail de recherche, avoir une culture générale importante sur des domaines éloignés de sa spécialité n'est pas moins crucial pour espérer proposer un travail complet. De plus les méthodes de raisonnement sont souvent assez différentes entre les domaines des mathématiques et les maîtriser est une plus value inestimable pour son travail futur.

Malgré la qualité de l'enseignement et des enseignants en classe préparatoire il est dommageable de constater que dans ce cursus les mathématiques (et les sciences en générale) sont réduites à l'état de mètre étalon des performances et de la qualité des étudiants. Le bachotage se fait alors au mépris de la compréhension profonde de l'enseignement proposé. L'envie d'approfondir et d'avoir le temps d'assimiler les concepts m'ont fait préférer la voie universitaire à une troisième année de classe préparatoire ou à une école d'ingénieur. Le choix de l'université s'est faite en fonction de l'existence d'une voie de magistère et de la généralité de la formation, c'est donc naturellement que j'ai postulé à l'université Paris Sud.

## Première année

### Premier semestre

Le premier semestre de troisième année de licence est l'occasion d'un premier choix dans l'option de magistère, celle que j'ai choisie est *théorie des graphes*. Ce choix répondait à une curiosité que j'avais eue plusieurs fois à propos de cette théorie dont j'avais entendu parlé en "prépa" et eu l'occasion d'entendre cité dans des articles de vulgarisation. Les côtés formel et théorique de ce cours ont été important pour acquérir de solides bases dans ce domaine.

Le cours obligatoire de programmation m'a permis de reprendre (et de mieux comprendre) les cours déjà suivis en "prépa". C'est avant tout l'état d'esprit de



la programmation et de l'informatique qui m'ont intéressé.

## Second semestre

J'ai choisit la voie d'informatique théorique pour ce second semestre. Cette voie m'apparaissait naturelle étant donné la curiosité envers l'informatique (et la formation en algorithmique des années antérieures). Cependant les cours dispensés ont été une vraie surprise. Je ne pensais pas que l'informatique était un domaine à ce point formalisé et que les résultats énoncés dépassés de loin de cadre de l'informatique (ayant notamment des conséquences en logique).

J'ai aussi pu proposer un TER sur la démonstration du théorème de GOLAB (= la semi-continuité inférieure de la mesure de HAUSDORFF de dimension 1). Le but de ce TER était l'étude d'existence de solution pour des problèmes d'optimisation. Le théorème de GOLAB permettant de justifier l'existence de tels maximums.

Le cours de magistère (tenu par M. ULLMO) sur la topologie générale est un complément très appréciable au cours de topologie et calcul différentiel du premier semestre. Il permet de bien voir et comprendre l'évolution des notions de topologies et celles d'analyses qui s'ensuivent.

## Stage

Dans le cadre de l'apprentissage hors mur j'ai pu effectuer un stage de trois semaines dans le laboratoire de sociologie quantitative du CREST à Montrouge. Ce stage m'a permis de découvrir les sciences sociales et de constater le peu de liens qu'il pouvait y avoir entre elles et les mathématiques. En effet elles ont souvent besoin d'outils mathématiques pour travailler ou formaliser leurs théories, en revanche les mathématiciens sont assez peu présents pour développer des outils spécifiques à tel ou tel problème de science sociale.

J'ai pu voir le type de questions qui se posent en sociologie et les méthodes mises en oeuvre pour tenter d'y répondre. De plus souvent les différentes sciences sociales tissent des liens assez étroits entre elles : les analyses comportementales des sociétés permettant d'éclairer et de comprendre les comportements des agents économiques individuels.

## Deuxième année

### Premier semestre

La variété et la qualité des enseignements rend le choix des matières difficiles. Cependant ces enseignements ont pour objectif de préparer à une spécialisation, j'ai donc choisit de me tourner vers les probabilités. Ce choix s'est fait naturellement puisqu'il permet une ouverture plus évidente vers les autres sciences, favorise le travail interdisciplinaire et propose une analyse rigoureuse des cadres d'expériences dans d'autres domaines que les mathématiques.

J'ai pu aussi suivre un cours de distribution, qui m'a permis de m'ouvrir à l'analyse moderne et qui est un cours nécessaire pour l'étude des équations aux dérivées partielles. Un cours d'algèbre, sur la théorie de GALOIS qui se distingue pour au moins deux points : le début de l'algèbre moderne, l'intérêt historique

de la théorie et du personnage de GALOIS dans l'histoire mathématique et républicaine (Evariste GALOIS, la fabrication d'une icône mathématique). Un cours d'analyse générale tourné sur la notion d'intégration, construit de manière historique et proposant l'intégration de KURZWEIL et HENSTOCK à, ce cours a permis de bien comprendre les relations (et surtout les différences) entre les définitions des intégrales qui ont été données dans l'histoire mathématique. Un cours de programmation avancée, qui fait suite à mon intérêt pour la programmation.

Le cours obligatoire de magistère sur l'analyse spectrale et l'analyse harmonique dont le professeur était absolument **exceptionnel** mais qui n'a pas pris la peine de démontrer le théorème de JORDAN.

## Second semestre

J'ai continué ma spécialisation par le choix du cours de Statistique dont les liens avec les autres sciences sont nombreux.

J'ai également suivi le cours de géométrie, m'ouvrant ainsi à un nouveau domaine et à ses méthodes d'analyses. Un cours de programmation sur graphe qui complète les cours de graphes de L3 et ceux de programmations. Le cours de logique a également été important pour le formalisme mathématique et comprendre comment ce formalisme s'inscrit dans une démarche de rigueur, mais aussi les limites de ce formalisme par les théorèmes d'incomplétude de GÖDEL. Enfin le cours d'histoire des mathématiques est essentiel pour comprendre comment les concepts mathématiques sont nés et comment ils se sont inscrits dans la communauté scientifique de l'époque. Pour comprendre aussi les erreurs mathématiques, leurs conséquences, leur correction et leur statut aujourd'hui.

Le second cours obligatoire de magistère portait sur les systèmes dynamiques.

J'ai proposé un TER de Statistique sur la reconstruction de mesure discrète permettant de se soustraire de la limite de diffraction si contraignante pour les physiciens. Ce TER m'a permis de voir et de comprendre comment un théorème théorique peut permettre l'amélioration effective des méthodes d'approximations et de tirer une information à partir d'observations a priori mauvaises.

## Troisième année

Le choix du Master Probabilité et Statistique s'inscrit en droite ligne de mes spécialisations de l'année précédente.

J'ai suivi des cours de Statistique sur différents domaines de recherche, ces cours étaient très inter-dépendant et d'apporter des compléments intéressants les uns aux autres. J'ai également choisi de participer à un certain nombre de cours de Probabilité pour avoir des bases théoriques solides.

Le cours supplémentaire que j'ai validé pour le magistère est un cours d'apprentissage statistique non supervisé (clustering en anglais). J'ai choisi ce cours en complément d'un cours d'apprentissage statistique supervisé. Ce cours m'a permis de saisir les différences entre deux domaines très proches formellement mais dont les approches sont éloignées.

Pour la validation de mon année de Master 2, j'ai choisi de faire un stage d'application dans le département de Recherche et Développement d'optimisation

et simulation (OSIRIS) chez EDF. Ce stage à eu pour objectif de comprendre, de maîtriser et d'améliorer le modèle de prévision de consommation électrique utilisé par EDF.



## Chapitre 2

# Domaine de Recherche : Apprentissage et Méthodes de Régression

## Introduction

Dans une société où il faut être capable de sectoriser et d'inventorier les disciplines, il est parfois difficile de situer certains domaines de connaissances dans l'un des nombreux champs de la science moderne. Durant longtemps les probabilités et les statistiques ont fait parti de ces champs de connaissances inclassables. Ces agrégats de connaissances étaient surtout utilisés dans des cadres sociaux (LAPLACE), démocratiques (paradoxe de CONDORCET-ARROW), sociologiques (COMTE et DURKHEIM), religieux (Le pari de PASCAL). Outils donnant un cadre théorique à des affirmations empiriques ; arguments absolus justifiant la genèse de nouvelles théories scientifiques.

Durant cette période les statistiques sont restées descriptives et les probabilités un amas de résultats plus ou moins clairs. Dans ces deux domaines, les affirmations étaient plus ou moins justifiées et le cadre théorique aussi précis que la notion de limite pour LEIBNIZ ou NEWTON. En somme les statistiques et les probabilités étaient plus un ensemble de méthodes d'analyse et de réflexions que des domaines scientifiques à part entière.

Depuis le développement de la théorie de la mesure, et la formalisation de la probabilité (en tant que domaine unifié) par Andrei KOLMOGOROV, le statut des statistiques s'est précisé. Si le domaine de la probabilité est l'étude des variables aléatoires, de leur loi de probabilité, de leurs propriétés, on peut naturellement définir le sujet de la statistique comme le problème inverse de tirer des informations des variables aléatoires, de leur loi de probabilité, de leurs propriétés à partir de tirages de ces variables aléatoires. La différence entre probabilité et statistique est essentiellement l'approche qu'en font les chercheurs. La nature même du questionnement de la statistique en fait un domaine avec une forte aspiration à l'application (ne serait-ce que pour vérifier empiriquement les résultats théoriques).

Aujourd'hui le développement de ces deux domaines des mathématiques (enfin "de la mathématique" pour rester cohérent) fait partie des plus prolifiques de notre époque pour au moins trois raisons (plus ou moins nobles mathématiquement) : un effet de mode au sein de la mathématique, une demande croissante des instituts et organisations de développements (INSEE, OCDE, OMS, OMC, ONG, ...) pour posséder et exploiter les données relatives à nos sociétés et nos comportements, une explosion des demandes du secteur privé pour exploiter les masses d'informations sur les consommateurs (Data Science et Data Scientist sont les nouvelles AOC [Appellation d'Origine Contrôlée] destinées au secteur privé). La raison principale de cette explosion est sûrement la prolifération de méthodes d'analyse statistique particulières qui commencent à remplacer la sacro-sainte régression linéaire par moindres carrés.

# Table des matières

<b>1</b>	<b>Apprentissage statistique</b>	<b>3</b>
1.1	Problématique . . . . .	3
1.2	Formalisation . . . . .	3
1.3	Cas de la régression . . . . .	4
<b>2</b>	<b>Sélection de modèle</b>	<b>4</b>
2.1	Problématique . . . . .	4
2.2	Un moyen détourné . . . . .	5
<b>3</b>	<b>Méthodes de régression : GLM et GAM</b>	<b>5</b>
3.1	Généralités . . . . .	5
3.2	Maximum de vraisemblance . . . . .	6
3.3	Les modèles linéaires généralisés (GLM) . . . . .	7
3.4	Les modèles additifs généralisés (GAM) . . . . .	8
3.5	Axe de recherche . . . . .	9

# 1 Apprentissage statistique

Nous pourrions lire utilement [Arl11] et [Del] pour la partie de régression.

## 1.1 Problématique

On peut trouver plusieurs domaines de recherches en statistique, le plus en vogue est celui de "l'apprentissage statistique" qui donne naissance à divers champs de recherche (intelligence artificielle, méthodes de régression, classification, optimisation) et qui est en lien avec d'autres domaines de la statistique (sélection de modèle, mélange de prédicteurs, validation croisée, bootstrap).

Dans ces différents champs, même s'ils abordent différemment le problème de l'apprentissage statistique, se pose le problème suivant. En posant  $\mathfrak{Y}$  et  $\mathfrak{X}$  deux ensembles et étant donnée des observations  $(Y_i)_{i \leq n}$  dans  $\mathfrak{Y}^n$  qui se sont produites dans un certain environnement  $(X_i)_{i \leq n}$  dans  $\mathfrak{X}^n$  (aussi appelée variables explicatives), pourra-t-on prévoir la nouvelle sortie  $Y_{n+1}$  dans  $\mathfrak{Y}$  pour de nouvelles données  $X_{n+1}$  dans  $\mathfrak{X}$  ?

Pour la reconnaissance de motifs ("Pattern Recognition") dans des images, cela revient à trouver les voitures, les pommes, les mangoustes dans une nouvelle image (le problème est qu'un objet n'est pas nécessairement toujours strictement identique au précédent, il faut donc être capable de reconnaître les caractères du motif plus que la forme exacte : une voiture peut être grande, longue, large, bleue, à trois portes, décapotable, ...). Pour l'intelligence artificielle, un robot peut apprendre à visser avec un tourne-vis plat et saura le faire avec un tourne-vis cruciforme sans l'avoir appris spécifiquement auparavant.

## 1.2 Formalisation

Soient  $(Y_i, X_i)_{i \leq n}$  dans  $(\mathfrak{Y} \times \mathfrak{X})^n$ ; sous l'hypothèse que les variables d'observations sont des fonctions de l'environnement, nous aimerions trouver une fonction mesurable de  $\mathfrak{X}$  dans  $\mathfrak{Y}$  (souvent appelé un classifieur)  $t$ , dans l'ensemble des fonctions mesurable  $\mathfrak{T}$ , telle que  $t(X_{n+1})$  soit "proche" de  $Y_{n+1}$ .

Commençons la répartition en plusieurs groupes :

- $\mathfrak{Y} = \{0, 1\}$ . Classification (binaire),
- $\mathfrak{Y} = \{0, \dots, m\}$ . Classification multivariée,
- $\mathfrak{Y} = \mathbb{R}$ . Autre modèle d'apprentissage statistique, notamment celui de la régression.

Dans la suite, nous serons amenés à considérer surtout le dernier modèle dans le cadre particulier de la régression.

Il arrive parfois que nous n'ayons pas entièrement accès aux valeurs de  $Y_i$ , la problématique est alors très différente en fonction de la connaissance partielle, totale ou nulle des valeurs prises par les  $Y_i$ . Nous pouvons proposer une nouvelle répartition :

- classification non supervisée (clustering en anglais). Nous n'avons pas accès aux observations  $(Y_i)_{i \leq n}$ . Dans ce cas :  $\mathfrak{Y} = \{0, \dots, m\}$  et l'objectif est de répartir les observations en différents groupes,
- classification semi-supervisée. Nous n'avons accès qu'à une partie des  $(Y_i)_{i \leq n}$ ,
- classification supervisée. Nous avons accès à toutes les observations  $(Y_i)_{i \leq n}$ .

Dans notre futur modèle de régression, nous serons dans le cas de la classification supervisée.



Pour juger de la performance d'un estimateur  $t$ , nous définissons la notion de fonction de perte (ou de contraste)  $\ell : \mathfrak{Y}^2 \mapsto \mathbb{R}$  qui doit mesurer à quel point  $t(X_{n+1})$  et  $Y_{n+1}$  sont "proches" :  $\ell(t(X_{n+1}), Y_{n+1})$ . Nous définissons le meilleur des classifieur

$$t^* \in \operatorname{argmin}_{t \in \mathfrak{T}} \{\mathbb{E}(\ell(t(X), Y))\}.$$

Notre objectif est de nous approcher de la perte de ce classifieur. Pour cela définissons un estimateur comme une fonction mesurable

$$\hat{s}_n : (\mathfrak{Y} \times \mathfrak{X})^n \mapsto \mathfrak{T},$$

nous ferons l'amalgame entre la fonction  $\hat{s}_n$  et son image  $\hat{t}_n$ . Il est courant de choisir l'estimateur  $\hat{t}_n$  de manière à ce qu'il minimise la fonction de contraste

$$\hat{s}_n((Y_i, X_i)_{i \leq n}) = \hat{t}_n \in \operatorname{argmin}_{t \in ?} \left\{ \sum_{i=1}^n \ell(t(X_i), Y_i) \right\}.$$

**Remarque.** L'espace  $?$  sur lequel a lieu la minimisation est volontairement non spécifié. Le choix de cet espace répond à plusieurs critères : calcul effectif de l'estimateur, intérêt de l'espace pour la prévision. Gardons à l'esprit que  $?$  est un espace inclus dans  $\mathfrak{T}$ .

L'objectif d'un tel estimateur  $\hat{t}_n$  est d'obtenir le résultat asymptotique

$$\lim_{n \rightarrow +\infty} \ell(t^*, \hat{t}_n) = 0 \quad P^{\otimes n} p.s.$$

### 1.3 Cas de la régression

Le cas particulier de la régression, qui s'inscrit dans le cadre de l'apprentissage supervisé, donne la situation suivante. Soit  $f$  une fonction d'un certain espace fonctionnel  $\mathcal{F}$ , on fait l'hypothèse (très forte) suivante

$$Y = f(X) + \varepsilon,$$

où  $\varepsilon$  suit une certaine loi. Soit  $\ell$  une fonction de contraste, alors on choisit l'estimateur de la fonction  $f$  comme suit

$$\hat{f} \in \operatorname{argmin}_{f \in \mathcal{F}} \left\{ \sum_{i=1}^n \ell(f(X_i), Y_i) \right\}.$$

## 2 Sélection de modèle

Nous pourrions regarder la partie sur la sélection de modèle de [Arl11] ou encore le cours [Mas03] pour un point de vue plus complet.

### 2.1 Problématique

L'idée de la sélection de modèle est de choisir, à partir d'une famille de modèles (ou plus concrètement d'estimateurs), le modèle ayant "la meilleure performance".

Donnons nous un ensemble de modèles  $(\mathcal{M}_p)_p$  et la famille d'estimateurs associés à chacun des modèles  $(\hat{f}_p)_p$  (par exemple  $p$  peut représenter le nombre de variables explicatives d'une régression, la taille de la base de projection de l'estimateur ou un certain paramètre comme cela

sera le cas dans la partie suivante). Etant donné ces estimateurs, nous aimerions pouvoir choisir le meilleur d'entre eux (celui qui minimise une certaine fonction de perte). Cependant il reste une difficulté : la non connaissance de la fonction  $f$  dont les  $\hat{f}_p$  sont des estimateurs ne nous permet pas de pouvoir choisir le meilleur des estimateurs de manière mesurable.

## 2.2 Un moyen détourné

Nous pouvons néanmoins espérer faire des sélections d'estimateurs suffisamment pertinentes pour ne pas être beaucoup plus mauvais que le meilleur estimateur. Si nous choisissons un certain critère de sélection : par exemple la minimisation d'une certaine quantité  $\text{crit}(\hat{f}_p)$  en prenant

$$\hat{p} \in \underset{p}{\operatorname{argmin}} \left\{ \text{crit}(\hat{f}_p) \right\},$$

sélectionnons l'estimateur  $\hat{f}_{\hat{p}}$ . L'objectif est de comparer sa performance  $\ell(f, \hat{f}_{\hat{p}})$  à la meilleure performance possible  $\inf_p \left\{ \ell(f, \hat{f}_p) \right\}$ . Lorsqu'un tel critère est choisi, nous voulons pouvoir trouver deux types de résultats de nature différente et complémentaire :

- Asymptotique, en assurant que notre méthode de sélection donne une erreur qui tend vers l'erreur du meilleur des modèles

$$\hat{p} \in \underset{p}{\operatorname{argmin}} \left\{ \text{crit}(\hat{f}_p) \right\} \implies \lim_{n \rightarrow \infty} \frac{\mathbb{E}(\ell(f, \hat{f}_{\hat{p}}))}{\inf_p \mathbb{E}(\ell(f, \hat{f}_p))} = 1,$$

- Non asymptotique (inégalité oracle), en assurant que le risque du modèle que nous avons choisi reste de l'ordre de grandeur du meilleur des modèles

$$\exists C > 0, \forall n \in \mathbb{N}, \exists R_n \in \mathbb{R}, \hat{p} \in \underset{p}{\operatorname{argmin}} \left\{ \text{crit}(\hat{f}_p) \right\} \implies \ell(f, \hat{f}_{\hat{p}}) \leq C \inf_p \left\{ \ell(f, \hat{f}_p) + R_n \right\},$$

les propriétés de la suite  $(R_n)_{n \in \mathbb{N}}$  font aussi partie des difficultés des preuves de tels résultats.

Le premier résultat est essentiel pour justifier l'utilisation d'un tel critère, la seconde est souvent plus délicate à obtenir, mais non moins intéressante.

## 3 Méthodes de régression : GLM et GAM

[Woo06] et [Del] pour les modèles GLM et GAM. Pour la partie sur la vraisemblance nous pourrions lire les cours de [Gas] et [RS12]. Par curiosité sur la validation croisée [AC10].

### 3.1 Généralités

Le développement de la méthode des moindres carrés par LEGENDRE et GAUSS au XIX<sup>e</sup> siècle, permettant de prévoir la trajectoire d'objets célestes, a placé les mathématiciens dans un monde où la recherche de solutions exactes (souvent complexes à obtenir) était moins intéressante que l'étude de solutions approchées.

Soit  $(Y, X)$ , nous supposons qu'il existe une fonction de lien  $f : \mathfrak{X} \mapsto \mathfrak{Y}$  ; la méthode des moindres carrés consiste à choisir la fonction  $\hat{f}$  pour minimiser la quantité

$$\sum_{i=1}^n (Y_i - f(X_i))^2.$$

Le choix de  $\hat{f}$ , et plus généralement le choix d'espace fonctionnel auquel appartient  $\hat{f}$ , est un problème subtil (un choix entre complexité de l'espace fonctionnel et précision de la modélisation).  
Finalement nous pouvons proposer, dans le cadre de la régression, la modélisation

$$Y = f(X) + \varepsilon,$$

$$\hat{f}_n \in \underset{f \in \mathfrak{F}}{\operatorname{argmin}} \left\{ \sum_{i=1}^n (Y_i - f(X_i))^2 \right\}.$$

### 3.2 Maximum de vraisemblance

Pour la suite, il est nécessaire de donner quelques informations sur une méthode d'estimation classique en statistique et qui se justifie dans un certain nombre de cas. Nous nous contenterons de donner une idée intuitive de cette méthode.

Soit une variable aléatoire  $X$  admettant une densité  $f$  par rapport à une mesure  $\lambda$  ( $d\mathbb{P}(x) = f(x)d\lambda(x)$ ), nous appelons  $f(X(\omega))$  la vraisemblance de  $X$  en  $\omega$  (nous n'écrivons plus l'aléa  $\omega$  dans la suite). Nous pouvons aussi considérer la log-vraisemblance :  $L(X) = \log(f(X))$ .

Prenons  $f_\mu$  la densité d'une loi normale centrée en un réel  $\mu$  et réduite. Donnons nous deux réels  $\mu_1 \ll \mu_2$  et une variable aléatoire  $X$  de densité  $f_{\mu_1}$  alors la quantité  $f_{\mu_1}(X)$  vaudra, avec forte probabilité,  $\frac{1}{\sqrt{2\pi}}$  tandis que  $f_{\mu_2}(X)$  sera, avec forte probabilité, proche de 0. Nous prendrions pour estimateur de  $\mu_1$

$$\hat{\mu} \in \underset{\mu \in \mathbb{R}}{\operatorname{argmax}} \{f_\mu(X)\}.$$

Naturellement choisir la valeur d'un estimateur sur une seule valeur d'une variable aléatoire n'est pas conseillé. Si nous nous donnons un échantillon  $(X_i)_{i \leq n}$  alors le nouvel estimateur est

$$\hat{\mu}_n \in \underset{\mu \in \mathbb{R}}{\operatorname{argmax}} \left\{ \prod_{i=1}^n f_\mu(X_i) \right\}.$$

Cet estimateur est appelé l'estimateur du maximum de vraisemblance et l'espoir (peut être fou) que nous plaçons en lui est qu'il converge vers la valeur  $\mu_1$ .

Toutes ces définitions se généralisent à n'importe quelle distribution  $f_\mu$ , et l'espoir de voir le maximum de vraisemblance converger vers la valeur du vrai paramètre reste intact.

**Exemple.** *Plaçons nous dans le contexte d'une régression linéaire avec un bruit gaussien ( $Y = {}^t(Y_1 \cdots Y_n) \in \mathbb{R}^n$ ,  $X = {}^t(X_1 \cdots X_n) \in (\mathbb{R}^d)^n$ ,  $\beta \in \mathbb{R}^d$ )*

$$Y = X\beta + \varepsilon, \quad \text{où } \varepsilon \sim \mathcal{N}(0, \sigma^2 I).$$

*Si  $(Y_i, X_i)_{i \leq n}$  est un ensemble de couples vérifiant cette relation alors la log-vraisemblance s'écrit*

$$L((Y_i, X_i)_{i \leq n}, \beta, \sigma^2) = -n \log(\sqrt{2\pi}\sigma) - \frac{1}{2\pi\sigma^2} \sum_{i=1}^n (Y_i - X_i\beta)^2.$$

*La vraisemblance dépendant de deux paramètres,  $\beta$  et  $\sigma$ , la recherche du maximum de vraisemblance en  $\beta$  revient à minimiser  $\sum_{i=1}^n (Y_i - X_i\beta)^2$ . On a donc équivalence entre la minimisation par moindres carrés et maximum de vraisemblance dans le cas particulier du bruit gaussien.*

### 3.3 Les modèles linéaires généralisés (GLM)

Nous élargissons le cadre de régression défini précédemment. Cette fois nous prenons  $(Y_i)_{i \leq n}$  des observations indépendantes suivant une loi de la famille exponentielle de paramètre réel  $\theta_i$ , c'est-à-dire admettant une densité par rapport à une mesure  $\lambda$  de la forme suivante

$$f_{\theta_i}(y) = \exp((y\theta_i - b_i(\theta_i))/a_i(\phi) + c_i(y, \phi)),$$

où les fonctions  $a_i : \mathbb{R} \rightarrow \mathbb{R}$ ,  $b_i : \mathbb{R} \rightarrow \mathbb{R}$  et  $c_i : \mathbb{R}^2 \rightarrow \mathbb{R}$ . Nous supposons la fonction  $a_i$  linéaire ( $a_i(\varphi) = \varphi/\omega_i$ ).

Nous supposons qu'il existe alors un lien entre les observations  $(Y_i)_{i \leq n}$  et des variables explicatives  $(X_i)_{i \leq n}$  dans  $\mathbb{R}^d$  (supposées déterministes)

$$\mathbb{E}(Y_i) = X_i\beta,$$

où  $\beta$  appartient à  $\mathbb{R}^d$ . Plus généralement, si nous prenons une fonction dite de lien  $g : \mathbb{R} \rightarrow \mathbb{R}$

$$g(\mathbb{E}(Y_i)) = X_i\beta.$$

L'estimation du paramètre  $\beta$  (seule inconnue du système) se fait par maximisation de la vraisemblance (l'exemple précédent nous a permis de voir que ceci est, dans certain cas, équivalent à faire une minimisation par moindres carrés)

$$\hat{\beta} \in \operatorname{argmax}_{\beta \in \mathbb{R}^d} \{L((Y_i, X_i)_{i \leq n}, \beta)\}$$

Dans le cadre des familles exponentielles nous avons les relations suivantes :  $\mathbb{E}(Y_i) = b'_i(\theta_i)$  et  $\operatorname{Var}(Y_i) = -b''_i(\theta_i)/a_i(\varphi)$ .

Nous calculons le maximum de vraisemblance (en réalité de la log-vraisemblance). Pour cela nous dérivons dans chacune des directions et demandons à annuler la dérivée (les paramètres ne sont pas nécessairement les mêmes pour chaque observation)

$$L((Y_i, X_i)_{i \leq n}, \beta) = \sum_{i=1}^n \omega_i (Y_i \theta_i - b_i(\theta_i)) / \varphi + c_i(\varphi, Y_i),$$

$$\frac{\partial L}{\partial \beta_j} = \frac{1}{\varphi} \sum_{i=1}^n \omega_i \left( Y_i \frac{\partial \theta_i}{\partial \beta_j} - \frac{\partial \theta_i}{\partial \beta_j} b'_i(\theta_i) \right).$$

Ecrivons cette dernière égalité avec des quantités que nous connaissons (faire disparaître  $\theta_i$  quitte à faire apparaître de nouvelles variables du modèle). Pour cela posons  $\mu_i = \mathbb{E}(Y_i) = b'_i(\theta)$  donc  $\frac{d\mu_i}{d\theta_i} = b''_i(\theta)$ . Nous avons

$$\frac{\partial \theta_i}{\partial \beta_j} = \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \beta_j},$$

et

$$\frac{\partial \mu_i}{\partial \theta_i} = b''_i(\theta_i) \implies \frac{\partial \theta_i}{\partial \mu_i} = \frac{1}{b''_i(\theta_i)}.$$

Remplaçons ces quantités dans l'expression de la vraisemblance

$$\frac{\partial L}{\partial \beta_j} = \frac{1}{\varphi} \sum_{i=1}^n \frac{Y_i - b'_i(\theta_i)}{b''_i(\theta_i)/\omega_i} \frac{\partial \mu_i}{\partial \beta_j}.$$

Posons  $V(\mu_i) = b_i''(\theta_i)/\omega_i = \frac{\partial \mu_i}{\partial \theta_i}/\omega_i$ . Finalement il nous reste à trouver  $\beta$  tel que

$$\forall j, \sum_{i=1}^n \frac{Y_i - \mu_i}{V(\mu_i)} \frac{\partial \mu_i}{\partial \beta_j} = 0.$$

En supposant la fonction  $V$  indépendante de  $\beta$ , nous avons équivalence avec la minimisation de

$$S = \sum_{i=1}^n \frac{(Y_i - \mu_i)^2}{V(\mu_i)}.$$

### 3.4 Les modèles additifs généralisés (GAM)

Ces modèles sont du même type que les modèles linéaires généralisés, en se permettant d'ajouter une partie fonctionnelle au lien qui lie les observations des variables explicatives

$$g(\mathbb{E}(Y_i)) = X_i\beta + f(X_i),$$

Le choix de l'espace fonctionnel  $\mathfrak{F}$  sur lequel nous opérons la régression est aussi sensible. Nous le faisons sur l'espace  $\mathfrak{F}$  des bases de splines. Plusieurs raisons nous permettent de choisir cet espace : les fonctions polynômes ont de bonnes propriétés (simplicité des calculs, approximation d'une fonction continue, ...), l'espace des fonctions splines possède des propriétés géométriques intéressantes permettant le calcul effectif de la fonction de régression  $f$  (c'est un espace de Hilbert, voir le rapport de Master 2), différentes bases de splines peuvent être utilisées ce qui donne de la souplesse à la méthode.

Une autre évolution est de faire apparaître une perte différente des moindres carrés : la perte Ridge. Pour bien comprendre comment intervient cette perte dans le processus de régression, regardons un exemple.

**Exemple.** Prenons  $g = Id$  et  $Y \sim \mathcal{N}(X\beta, \sigma^2)$ , nous écrivons la régression GAM comme suit

$$Y = X\beta + f(X) + \varepsilon,$$

avec la fonction de perte Ridge

$$(\hat{\beta}, \hat{f})_\lambda \in \underset{\beta \in \mathbb{R}^d, f \in \mathfrak{F}}{\operatorname{argmin}} \left\{ \|Y - X\beta - f(X)\|^2 + \lambda \int f''(t)^2 dt \right\}.$$

On peut voir cette fonction de perte comme une "pénalisation" sur la partie fonctionnelle. Les fonctions avec une intégrale importante seront moins susceptibles d'être choisies pour la régression. Pour une discussion plus précise sur ce type de perte voir le rapport de stage de Master 2. En écrivant la fonction  $f$  comme une combinaison linéaire de fonctions splines de base ( $f(x) = \sum \alpha_i g_i(x)$ ) alors il existe une matrice symétrique positive  $S$  tel que  $\int f''(t)^2 dt = {}^t \alpha S \alpha$ .

Nous obtenons ici une famille d'estimateurs (indexée par  $\lambda$ ), il faut donc trouver un critère pour en sélectionner un seul. Ceci se fait par la méthode de validation croisée généralisée (de même voir le rapport de Master 2 pour plus de précision et [CW79], [CG05], [CG06])

$$\hat{\lambda}_{GCV} \in \underset{\lambda > 0}{\operatorname{argmin}} \frac{\|(I - M_\lambda)Y\|^2}{(1 - \operatorname{tr}(M_\lambda/n))^2},$$

où  $M = X({}^t X X + \lambda S)^{-1} {}^t X$ . Cette méthode de sélection nous assure les deux types de résultats que nous avons énoncés dans la partie sur la sélection de modèles (asymptotique et non asymptotique).

L'optimisation par maximum de vraisemblance et perte Ridge donne

$$(\hat{\beta}, \hat{\alpha}) \in \operatorname{argmax}_{\beta, \alpha} \{L((Y_i, X_i)_{i \leq n}, \beta \in \mathbb{R}^d, \alpha \in \mathbb{R}^n) - \lambda^t \alpha S \alpha\}.$$

En faisant la même suite de calculs et en supposant que le paramètre  $V$  ne dépend pas des  $\beta$ , nous obtenons la minimisation de la quantité

$$S = \sum_{i=1}^n \frac{(Y_i - \mu_i)^2}{V(\mu_i)} + \lambda^t \alpha S \alpha.$$

### 3.5 Axe de recherche

Pour arriver à ce niveau il aura fallu le temps de développer plusieurs théories annexes permettant une estimation efficace de ces modèles; essentiellement il aura fallu attendre des avancées dans trois domaines : les RKHS (voir Annexe du rapport de Master 2), les méthodes de sélection et de validations croisées, la régression avec perte Ridge.

Finalement c'est à partir des années 2000, qu'une estimation efficace a pu être mise en place, avant cela les méthodes d'estimations avaient plusieurs défauts : complexité des méthodes de calculs qui rendait les peu commodes, difficultés d'obtenir des fonctions splines qui représentaient correctement les données.

Depuis, ces méthodes ont connu des améliorations plus ou moins importantes. La principale est le développement d'une méthode pour traiter les très grands volumes de données. L'algorithme repose sur une décomposition  $QR$  et la possibilité de construire la matrice  $Q$  par récurrence sur le nombre d'observations ; cela a un avantage majeur (outre de traiter de gros volume de données) : le calcul récursif permet une mise à jour rapide dans le cas où de nouvelles observations arriveraient.

Des généralisations de ces méthodes sont possibles et sont des axes de recherches : la mise en place de régression GAM pour des observations  $Y$  ne suivant pas nécessairement une loi de la famille exponentielle.

## Références

- [AC10] S. Arlot and A. Celisse. A survey of cross-validation procedures for model selection. *Statistics surveys*, 2010.
- [Arl11] S. Arlot. Sélection de modèles et sélection d'estimateurs pour l'apprentissage statistique. <http://www.di.ens.fr/~arlot/peccot.htm>, 2011. Cours Peccot.
- [CG05] Y. Cao and Y. Golubev. On oracle inequalities related to a polynomial fitting. *mathematical methods of statistics*, 2005.
- [CG06] Y. Cao and Y. Golubev. On oracle inequalities related to smoothing splines. *mathematical methods of statistics*, 2006.
- [CW79] P. Craven and G. Wahba. Smoothing noisy data with spine functions, estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 1979.
- [Del] B. Delyon. Régression, cours de deuxième année de master. <https://perso.univ-rennes1.fr/bernard.delyon/regression.pdf>. Université Rennes I.
- [Gas] E. Gassiat. Statistiques, notes de cours, m2. <http://www.math.u-psud.fr/~gassiat/notesdecours-M2Stat.pdf>. Université Paris Sud.

- [Mas03] P. Massart. *Concentration Inequalities and Model Selection*. Springer, 2003.
- [RS12] V. Rivoirard and G. Stoltz. *Statistiques en action*. Vuibert, seconde edition, 2012.
- [Woo06] S.N. Wood. *Generalized Additive Models, an introduction with R*. Chapman & Hall, 2006.





## Chapitre 3

# Immersion dans le Laboratoire de Sociologie Quantitative (CREST)

# Introduction

Ce rapport présente le travail effectué entre le lundi 3 juin et le vendredi 28 juin dans le Laboratoire de Sociologie Quantitative (LSQ) du CREST.

Suite à la troisième année de licence et dans le cadre de l'apprentissage "hors mur" lié au magistère de mathématique de l'université d'Orsay, j'ai choisi le laboratoire de sociologie quantitative du CREST (rattaché à l'ENSAE [Ecole Nationale de Statistiques Administratives et Economiques]) pour effectuer un stage de quatre semaines durant le mois de juin.

J'ai voulu faire ce stage en sociologie car c'est une science rarement appréhendée dans les études mathématiques, et qu'elle utilise un certain nombre d'outils mathématiques concrets qui sont peu étudiés dans les premières années du cursus (statistiques et probabilités). J'espérais aussi pouvoir apporter un peu de mon savoir, ou du moins mieux comprendre les aspects théoriques grâce à ma formation. Ce stage m'a également permis de voir comment s'organise la vie d'un laboratoire de recherche.

Durant ces semaines de travail j'ai ainsi proposé une contribution à deux travaux (à la fois sur des aspects théoriques d'analyses des données et sur des interprétations de ces données), assisté à plusieurs conférences, à des séances de critiques et corrections de travaux.

J'ai aussi appris un nouveau langage de programmation (SAS) ce qui a enrichi encore ma connaissance en algorithmique et m'a permis d'appréhender de nouvelles applications à la programmation.

Je vais donc présenter ces quelques semaines de travail en commençant par décrire le cadre général dans lequel j'ai pu travailler, ensuite chacun des travaux auxquels j'ai participé et comment s'est effectué mon travail.

## Table des matières

<b>1</b>	<b>La vie dans le Laboratoire</b>	<b>2</b>
1.1	Le LSQ . . . . .	2
1.2	Ma place au sein du laboratoire . . . . .	3
<b>2</b>	<b>Une contribution de thèse</b>	<b>5</b>
2.1	Préparation des données et premiers problèmes . . . . .	5
2.2	Premières études et premiers résultats . . . . .	7
<b>3</b>	<b>Contribution à un projet d'étude sociologique d'un site internet humoristique</b>	<b>9</b>
3.1	Construire une nomenclature et choisir les VDM . . . . .	9
3.2	Analyses et traitements des données . . . . .	10
<b>4</b>	<b>Conclusion</b>	<b>12</b>

# 1 La vie dans le Laboratoire

La description du fonctionnement d'un laboratoire est une chose assez "simple" (si l'on contente de décrire le calendrier du laboratoire). En revanche décrire les relations et les rapports entre les hommes et femmes d'un tel groupe est beaucoup plus complexe. Cette difficulté est encore augmentée lorsqu'on ne possède pas les connaissances pour le faire, ce qui est mon cas puisque je n'ai pu me pencher sur l'ethnologie et l'anthropologie durant ce stage. Ainsi je me contenterai de décrire le fonctionnement habituel du LSQ et d'essayer de voir quelle place j'ai pu avoir dans cette organisation.

## 1.1 Le LSQ

Le Laboratoire de Sociologie Quantitative (LSQ) a été créé pour favoriser la recherche et l'enseignement de la sociologie en France. Aujourd'hui le LSQ est dirigé par Michel Gollac. Ce laboratoire est encadré par le CREST (Centre de Recherche en Economie et Statistique). Organisme qui est lui même dépendant du GENES (Groupe des Ecoles Nationales d'Economie et Statistique) de l'INSEE dont l'ENSAE est l'un des membres, le laboratoire s'inscrit ainsi à la fois dans un objectif de recherche mais aussi d'enseignement par ses liens avec les Ecoles Nationales faisant parties du GENES.

Le laboratoire est constitué d'une quinzaine de doctorants et d'une vingtaine de chercheurs. Par le nombre important de ses membres le LSQ peut couvrir un grand nombre de domaines de recherche (structures sociales et comportements individuels [groupes sociaux et professions, mobilité sociale, sociologie du genre, représentations et valeurs], intégration [intégration professionnelle, travail et marché du travail, intégration scolaire, intégration des immigrés, discrimination nationale ou ethnique]), ce qui permet une grande diversité dans les études sociologiques au sein du laboratoire et donne ainsi une complémentarité nécessaire à un travail de qualité.

Le LSQ a vocation à utiliser l'ensemble des outils (mathématiques, tests internationaux standardisés) à sa disposition pour fournir des études reconnues. En effet le qualificatif "quantitatif" ne renvoie pas uniquement à une expression médiatique destiné à rendre cet organisme plus attrayant, mais il s'inscrit dans une volonté de faire évoluer la sociologie. Il ne s'agit plus seulement de développer une pensée à partir d'entretiens (en nombres restreints) mais d'utiliser l'ensemble des bases de données des tests internationaux et l'ensemble des outils statistiques qui ont pu être développés ou qui le sont à l'occasion d'études approfondies. Les chercheurs du laboratoire appuient donc leurs conclusions sur des bases solides, propices à la discussion et dont l'objectivité est assurée par le sérieux des bases de données utilisées. Le laboratoire a donc pour objectif de fournir des analyses sérieuses, objectives et critiquables.

Au sein du laboratoire nombre de chercheurs sont affiliés à plusieurs laboratoires, ainsi le LSQ se remplit et se vide au grès des différents emplois du temps. Cela n'empêche pas l'organisation de journée de séminaire (une fois par mois environ) très suivies. Durant cette journée se sont deux ou trois séminaires qui ont lieu. Dans un premier temps c'est un séminaire "interne" qui est présenté, à cette occasion l'un des membres du laboratoire propose un exposé d'environ une heure sur un travail ou un projet qui est partiellement fini. Ceci permet d'avoir

l'avis de l'ensemble des chercheurs présent mais également de savoir vers quelles pistes approfondir son travail (c'est d'ailleurs suite à un tel séminaire que mon travail sur le site VDM a été défini de manière plus précise). C'est ensuite à un intervenant extérieur de venir présenter un travail (aussi complet que possible cette fois ci). Ces journées de séminaires permettent de garder un contact avec la recherche extérieure au laboratoire et d'éveiller la curiosité des chercheurs dans des domaines qu'ils connaissent moins bien. Ces séminaires sont également très intéressants puisqu'ils permettent une critique et une discussion longue sur les travaux présentés (le format habituel de ces présentations est d'une heure d'exposé pour une heure de discussion, de plus le public n'est composé que d'une quinzaine de personnes). Ce sont donc des séminaires très ouverts.

En dehors de ces séminaires chaque chercheur assiste à de nombreux colloques ou exposés. J'ai ainsi pu assister à l'intervention, à l'INSEE, d'un employé du ministère de l'éducation, qui a travaillé avec Mme. Le Donné, sur la réussite scolaire (en rapport direct avec mon premier travail). Cette intervention m'a donc permis d'avoir un aspect historique et de mieux comprendre les évolutions de mes propres résultats par rapport aux résultats que cet intervenant nous a énoncés.

De manière complémentaire l'ensemble des doctorants du LSQ participent régulièrement à des séances de critiques de travaux. Un thésard propose ainsi aux autres de lire un article à paraître ou un chapitre de sa thèse. Ses collègues en font la critique ; aussi bien sur la forme que sur le fond. Ces rencontres sont très appréciées des doctorants puisqu'elles permettent d'avoir un retour constructif et rapide sur leur travail.

## 1.2 Ma place au sein du laboratoire

Les tous premiers jours ont surtout étaient utiles pour m'habituer aux lieux et aux personnes qui y travaillent. De plus ces deux moments d'adaptation se sont faits facilement notamment par la gentillesse et l'écoute dont ont fait preuve thésards et chercheurs. Les premiers temps ont aussi étaient dédiés à la découverte de ce qu'est la sociologie. Cette découverte s'est fait par un double mouvement : d'abord les moments de discussions m'ont permis de percevoir les intérêts et les beautés de la sociologie, tandis que l'apprentissage d'un nouveau langage (SAS) m'a permis d'appréhender plus facilement les bases d'un raisonnement sociologique (et aussi d'en voir les aboutissements en remarquant comment l'utilisation de ce langage s'insérait dans le travail de recherche) et de percevoir rapidement les problèmes que pouvaient poser certain raisonnement et aussi la force qu'avait d'autre en bouleversant les préjugés.

Les nombreuses discussions ont ainsi participé naturellement à ma découverte de cette science. La sociologie est en effet une science beaucoup plus accessible que les sciences dites "dures", les moments d'échanges sont donc des instants privilégiés durant lesquels stagiaires, thésards et chercheurs s'entraident et s'ouvrent aux autres pour enrichir leurs pensées.

Ce sont ces différents facteurs qui m'ont permis de comprendre en quoi la sociologie est véritablement une science et pas seulement un ersatz de philosophie sociale. En effet on a coutume de rapprocher la sociologie à la philosophie : une matière qui explique la société à grand coups de concepts plus ou moins abstraits et plus ou moins partagés. En réalité l'utilisation de base de données

représentatives et l'analyse de ces données par des outils statistiques donne une objectivité et permettent des critiques qui ne révèle pas uniquement une interprétation personnelle. De plus les hypothèses et conjectures sont justifiées par des observations dont les critères de représentativité sont suffisants. Ceci rend chaque raisonnement objectivement critiquable, ce qui est le propre d'une science.

Dans cette organisation bien huilée trouver la place qu'il convient d'avoir n'est pas évident. Montrer que notre point de vue est enrichissant et proposer des critiques pertinentes est un exercice difficile. Ainsi durant les premiers exposés j'ai surtout écouté les remarques faites par les chercheurs pour essayer de comprendre ce qui est critiquable et ce qu'il est intéressant d'observer dans un travail.

Les lectures que m'a ensuite données ma maître de stage m'ont permis de m'entraîner et de faire l'exercice difficile de critiquer un travail alors même que je ne possède aucune compétence en sociologie. Cet exercice permet néanmoins de nous montrer les limites dans lesquelles on peut corriger le travail de l'autre.

Dans mes relations avec Mme. Le Donné ma position de stagiaire me permettait aussi de me faire expliquer des aspects théoriques du travail à fournir. L'explication des ces aspects forçait mon encadrant à faire preuve de pédagogie et ainsi de réapprendre elle même des notions qu'elle n'utilisait plus depuis quelques années. Ce qui m'a permis de bien comprendre les raisons des utilisations des différentes variables même si les détails des modèles utilisés sont restés parfois assez obscur.

Au cours de ce stage j'ai également pu proposer mon aide à un thésard pour vérifier des résultats de mathématique.

Tout au long du mois j'ai aussi pu constater que les membres du laboratoire avaient une vraie curiosité pour les mathématiques (que chacun maîtrisait avec plus ou moins de facilité), pour ce que pouvait être la recherche dans ce domaine et ce que pouvait être le résultat de ces recherches.

Durant la durée du stage j'ai également pu constater que l'ensemble des personnes du laboratoire évoluaient sur un pied d'égalité : chercheurs, doctorants et assistants de direction se tutoient et parlent à bâton rompu entre eux. Je dois avouer que j'avais quelques appréhensions lors de la première rencontre avec ma maître de stage, je ne savais pas si je devais utiliser le tutoiement ou non alors que je devais travailler durant un mois avec elle et que nos âges n'étaient pas si éloignés. Cela n'a pas duré longtemps puisque sa première demande a été d'utiliser le tutoiement. Le tutoiement des autres membres a alors naturellement suivi en voyant la chaleur de leur accueil et que pour leur part ils me tutoyaient. Ceci a favorisé mes questions et mes interventions à propos de mon travail et de celui des autres.

J'ai pu ressentir que j'étais aussi mis sur le même pied d'égalité que l'ensemble des membres du laboratoire, sensation qui se fortifie lorsque j'ai pu passer le repas à discuter avec M.Gollac, (habituellement appelé Michel par tous les membres) directeur du LSQ, de sujets très divers et surtout avec une grande simplicité. En réalité, s'il fallait décrire l'ensemble des rapports entre les membres du laboratoire, il serait plus exact de dire que ces relations étaient nettoyées des codes de hiérarchie (sauf naturellement à certains moments précis de la vie du

laboratoire, comme les réunions de décisions).

## 2 Une contribution de thèse

Noémie Le Donné travaille sur la réussite scolaire des élèves de 15 ans suivant leurs origines (sociales, géographiques,...). Pour son étude elle utilise la base de donnée des études PISA (2000, 2003, 2006, 2009). Le premier travail proposé est en rapport direct avec la thèse de Mme Le Donné.

Pour les années 2000 et 2006 une question sur les aspirations professionnelles des élèves de 15 ans a été posée, cette question porte sur le métier que les élèves espèrent avoir à l'âge de 30 ans. Noémie a déjà travaillé sur cette question mais dans le cadre d'une comparaison internationale, et montrant que les aspirations des adolescents dépendent beaucoup du niveau économique du pays (les étudiants des pays les plus défavorisés ayant les aspirations les plus "irréalistes" -i.e : le moins en accord avec le marché du travail- tandis que ceux des pays les plus développés sont plus "réalistes"). Cependant aucune étude approfondie n'a été réalisée sur le cas des aspirations des élèves français.

Ma première contribution a donc été d'étudier cette variable et d'essayer de l'expliquer en fonction d'autres variables (origines, niveaux, ...).

### 2.1 Préparation des données et premiers problèmes

La population étudiée est constituée d'environ 4 500 élèves âgés de 15 ans pendant les années 2000 et 2006. La question posée aux adolescents est : "Quel métier espèreriez vous pratiquer à 30 ans?". A cette question les élèves répondent par un intitulé clair du métier attendu, cet intitulé est ensuite traduit dans la nomenclature CITP 1988 (ou ISCO en anglais) qui est une nomenclature internationale subdivisant l'ensemble des métiers en neuf grandes catégories socioprofessionnelles elles même subdivisées en plusieurs sous catégories (par exemple le nombre 2113 représente les élèves voulant devenir chimiste [le chiffre 2 représente la grande catégorie, le 1 une sous-catégorie, le 1 une nouvelle sous-catégorie et le dernier chiffre un métier précis dans ces sous-catégories]). A partir de ces données il nous est donc possible de connaître exactement les aspirations des élèves sondés.

Plusieurs contraintes apparaissent cependant :

- Il n'est pas très judicieux d'étudier en détail les aspirations des étudiants, en effet le nombre de métiers est trop important. De plus un grand nombre de métiers peuvent avoir des statuts équivalents (deux chercheurs travaillant dans des domaines différents par exemple), il est alors intéressant de les regrouper dans une case "métiers de la recherche" ; cette opération nous permet toujours d'étudier le "niveau d'aspiration" des élèves et simplifie le travail,
- Cette nomenclature ne permet pas de définir ce que l'on pourrait appeler un "bon" ou "mauvais" métier. Elle ne peut pas être utile non plus pour construire un indicateur jugeant de la qualité du métier. L'étude du "niveau d'aspiration" des adolescents est alors très complexe. On aurait pu penser que la nomenclature CITP rapprochait les grands corps de métiers (comme les chercheurs, les médecins, les cadres,...) et baser notre étude dessus, cependant elle rapproche surtout les "employés du privé", "travailleurs qualifiés" qui couvre un échantillon

très divers de professions. Ainsi il n'est pas possible de juger de la qualité des aspirations,

- Cette nomenclature a subi des réajustements dans différents pays (réajustements propres à chaque pays) ce qui a eu pour effet la création de nouvelles classes qui ne peuvent être étudiées puisqu'à aucun moment il n'est renseigné ce qu'elles contiennent.

Les premiers problèmes soulevés demandent une définition précise et claire de la variable, qui sera la variable d'aspiration, sur laquelle repose l'ensemble de l'étude. En réalité la variable d'aspiration existe déjà, on l'appelle SEI. Le SEI est un chiffre compris entre 16 et 90 associé à chaque métier et qui prend en compte le nombre d'années d'études et le salaire moyen pour pratiquer tel ou tel métier. Plus cette quantité est proche de 90 plus la profession associée nécessite d'années d'étude et est bien rémunérée. C'est donc cette variable qui va nous servir à quantifier les aspirations des élèves. Par ce choix on se rend compte que dire qu'un élève a des aspirations élevées revient à dire qu'il espère avoir un métier socialement valorisant, ainsi la définition que l'on donne aux "bons" ou "mauvais" métiers est avant tout due à la subjectivité de notre société. On peut alors commencer à résoudre nos problèmes :

- Pour le premier je me suis contenté de regarder les deux premiers chiffres de la nomenclature CITEP pour classer les différents métiers (j'ai donc fait l'hypothèse que les professions situées dans la même classe des deux premiers chiffres étaient de niveau équivalent) pour étudier les choix des élèves et voir s'ils avaient changé au cours des six années,

- Pour le second j'ai naturellement utilisé la variable que j'ai introduit juste au dessus, tout en étant conscient des aspects arbitraires et subjectifs de cette définition (il faut faire remarquer cependant que le choix du SEI n'est pas tout à fait absurde puisque les futurs citoyens s'insèrent dans la société notamment par leur métier, il est donc normale de prendre le point de vue de la société pour définir ce qu'est un "bon" métier). L'autre avantage d'utiliser cette variable est qu'elle est renseignée dans les données des études PISA,

- Le dernier problème se résout naturellement après le choix de la variable d'aspiration, en effet le SEI est renseigné dans les données ce qui permet de ne pas se préoccuper des nouvelles classes.

Après avoir trouvé les variables principales que j'allais étudier (à savoir le SEI et les deux premiers chiffres de la nomenclature CITEP), j'ai dû regarder les variables qui vont avoir un impact potentiel sur ces deux variables.

Dans un premier temps l'origine géographique (française ou non) de l'élève ; j'ai considéré qu'un étudiant est de première génération si ses parents sont nés hors de France et que lui même n'y est pas né, il est de seconde génération lorsque ses parents ne sont pas nés en France, enfin il est français lorsque l'un au moins de ces parents est né en France.

Ensuite j'ai regardé la catégorie socioprofessionnelle des parents (le métier de chacun des parents est normalement renseigné avec le SEI associé, de plus l'enquête PISA donne le HISEI qui correspond au maximum des SEI des deux parents et c'est cette quantité qui va me servir pour définir la catégorie socioprofessionnelle des parents).

Puis je me suis occupé du plus haut niveau d'éducation des parents (qui représente le niveau d'étude maximale atteint par les parents : collège, lycée, étude

supérieure, ...).

Ensuite le niveau d'étude dans lequel se situe l'élève (au lycée [en classe de 2<sup>nd</sup> pour un élève de 15 ans "à l'heure"] professionnel ou général, au collège,...), on remarquera tout de même que les élèves qui sont sortis du système scolaire ne peuvent apparaître dans cette étude.

Le niveau scolaire des élèves est également une variable explicative (ce niveau scolaire est aussi étudié par l'étude PISA qui fait passer différents tests, ce sont ces tests que j'ai utilisé pour évaluer le niveau de chaque élèves).

## 2.2 Premières études et premiers résultats

La première partie de mon travail a donc été l'étude approfondie des choix des métiers des adolescents. J'ai donc séparé les élèves dans chacune des classes (à deux chiffres) que j'avais construites. J'ai pu constater la popularité de certains types de métiers chez les jeunes étudiant(e)s (aussi la désertion d'autres corps de métiers) : l'enseignement, la médecine et professions du privé sont les plus espérés. On constate aussi que ces grands domaines sont en générale ceux avec le SEI le plus important, ce qui permet de voir une première corrélation entre les attentes des élèves et ceux de la société (et de leurs parents) à leur égard.

Il a fallu expliquer certaines différences entre les années 2000 et 2006. Pour certains grands domaines les fréquences de choix avaient complètement évolué. Il s'est avéré que ces évolutions étaient en fait à mettre sur le compte de un ou deux métiers qui avaient été particulièrement choisis lors de l'une ou l'autre des années. Ces changements pouvaient s'expliquer par plusieurs choses : la première est le choix des lycées pour mener les tests (en effet il se peut que le test soit fait dans un lycée agricole et dans ce cas il est normal que l'ensemble des élèves choisissent le même domaine de profession : le métier sera alors sur-représenté), la seconde est que différentes campagnes de publicité ou politiques éducatives est mis en exergue une profession qui attire les nouveaux élèves. Savoir laquelle de ces influences est la plus prédominante est complexe, il faudrait en effet faire un travail sur l'ensemble des établissements et espérer que les filières qu'ils proposent ne soient pas trop vastes pour savoir quel type de classe a été interrogé. Ce travail n'a pas été fait, à la fois par manque de temps et pour favoriser d'autres études.

J'ai pu rapidement comparer les aspirations des filles et des garçons, on constate (et on pouvait s'y attendre) qu'il existe un certain nombre de métiers qui sont genrés (c'est à dire majoritairement choisis par des filles ou des garçons) : l'enseignement, le paramédical, la vente sont autant de métiers choisis majoritairement par les filles tandis que ouvriers, chercheurs, artisans sont surtout pris par les garçons.

On pourrait continuer à étudier les catégories socioprofessionnelles les plus choisies par les élèves, en regardant celles choisies dans les lycées professionnels, généraux et dans les collèges ; cependant ce sont toujours les aspirations qui nous intéressent et que l'on veut étudier avec détails.

La première "vérification" que j'ai faite et de regarder le SEI moyen de chacune des neuf familles des catégories socioprofessionnelles, pour voir si un SEI élevé correspondait à certaine classe de professions. J'ai ainsi pu voir que les premières catégories sont bien celles dont le SEI est le plus haut, on y retrouve



les hauts fonctionnaires, les professions intellectuelles, les employés du privé.

A partir de là on commence donc à étudier les corrélations entre notre variable d'aspiration et l'ensemble des autres variables explicatives.

Au niveau des sexes on constate une légère différence dans la valeur moyenne des aspirations : les filles ont en effet toujours des aspirations légèrement plus importantes que les garçons.

L'origine géographique donne des résultats assez intéressants car contre intuitifs. Les élèves avec les aspirations les plus hautes sont ainsi les immigrés de première génération (alors que statistiquement ce sont ceux qui réussissent le moins, à la fois au niveau des notes et plus tard professionnellement), ensuite les immigrés de seconde génération ont les aspirations les plus élevés et enfin ce sont les "français" qui ont les espoirs les moins haut. Ce type de résultats posent des problèmes d'interprétations évident : comment justifier logiquement une telle observation ? L'explication la plus simple est que les parents des enfants immigrés les poussent à voir l'école comme un ascenseur social et fonde ainsi plus d'espoir que les jeunes français sur leur capacité à s'en sortir même si leur niveau est faible. On voit toute la difficulté de telles interprétations pour un esprit qui n'a pas une certaine culture sociologique et qui n'est pas habitué à ce type de réflexions. De plus on voit ce que cette justification à d'arbitraire et de critiquable (surtout en l'absence de données pour accréditer notre thèse). Et c'est aussi tout l'enjeu de cette science : être capable de s'appuyer sur des travaux antérieurs ou plus théoriques, de les inscrire dans sa propre démarche sociologique et ainsi de justifier certaines conclusions ou en invalider d'autres.

La catégorie socioprofessionnelles des parents n'a pas une influence négative sur les aspirations des élèves, au contraire, comme on pouvait s'en douter. Pour constater une forte corrélation positive on peut ranger le HISEI des parents par quartile et voir la moyenne des SEI des enfants dans chaque quartile. J'ai pu alors constater que la corrélation est presque linéaire.

Cette dernière opération a également été profitable par le fait qu'il me faisait vérifier certain résultats que Mme. Le Donné avaient déjà. Or justement ces résultats ne concordaient pas : les quartiles semblaient avoir été mal définis.

J'ai pu constater alors tout l'intérêt d'avoir un deuxième point de vue sur ses propres travaux : celui de corriger les erreurs que l'on a pu laisser passer et qui sont presque invisibles (à moins de réécrire l'ensemble de ses programmes ce qui serait absurde). Cette observation m'est apparue à l'occasion d'une autre divergence de résultats que je présente ci dessous.

L'étude de la corrélation entre le plus haut niveau d'étude des parents et les espoirs de leurs enfants est quant à elle assez complexe. On constate que pour les parents ayant eux des formations supérieures et ayant finis le lycée ont des enfants avec de fortes aspirations, en revanche l'interprétation est beaucoup plus difficile pour les enfants des autres parents. On ne peut pas tirer de conclusions générale à partir des résultats qui sont très différents entre les années 2000 et 2006 et qui ne témoignent d'aucun régularité au sein de chaque année. On aboutit ici à un autre type de difficulté : celle de ne pas pouvoir interpréter les données. Il faut alors réussir à comparer ces résultats à d'autres, les mettre en parallèle et tenter d'aboutir à une explication.

C'est au cours de l'étude de ces résultats que j'ai pu constater des divergences entre mes données et celle de mon encadrant. En effet nous avons pas le même nombre de parents dans tous les niveaux d'études. En fait nos algorithmes étaient différents : alors que je construisais la donnée du niveau d'étude des parents moi même à partir de la question, Mme. Le Donné avait choisi d'utiliser une donnée remplie directement par PISA sur le niveau d'étude maximale. Cependant les données remplies par PISA l'avaient été selon des règles très discutables (avec lesquelles nous n'avons pas été d'accord par ailleurs). Le préjugé sur les effectifs n'était pas très important (une soixantaine d'individus) cependant la différence des résultats était particulièrement visible puisque l'on pouvait passer dans une catégorie de moins de 40 pour cent des effectifs à plus de 40 pour cent.

Les deux dernières variables explicatives sont le niveau de l'élève et sa classe. De manière tout à fait prévisible son niveau et ses aspirations sont corrélés de manière positive.

Quant à la classe de l'élève (ou plus précisément le type d'établissement fréquentés) on constate que les aspirations les plus hautes sont pour les lycées généraux mais qu'en revanche les élèves encore au collège ont des espoirs plus importants que ceux qui sont en lycée professionnel. Ceci s'explique facilement par le fait que les collégiens restent dans une voie généraliste.

On peut voir à travers tout cela la complexité de l'étude de ce type de données qui ne sont parfois pas suffisantes pour conclure quant à l'importance de l'influence de telle variable sur une autre. Dans ces cas le sociologue revient à une attitude plus hypothétique et tente tout de même de formuler une justification même si elle est souvent complexe et discutable.

### **3 Contribution à un projet d'étude sociologique d'un site internet humoristique**

Mme. Le Donné a aussi proposé à des élèves-ingénieurs (de l'ENSAE), dans le cadre d'un projet, l'étude du site internet [www.viedemerde.fr](http://www.viedemerde.fr). Un site internet créé en 2008, à but participatif, qui publie les anecdotes que chacun a vécu (nous rentrerons plus loin dans les détails de publications de ces anecdotes). Mis à part la partie sympathique de ce travail, cette étude peut également permettre une analyse des comportements sociaux (à travers le cadre des anecdotes racontées), des habitudes verbales des auteur(e)s ou encore de la vision que ces mêmes auteur(e)s peuvent avoir de leur situation ou de celle des autres. Ce site internet peut ainsi, potentiellement, permettre de répondre à plusieurs questions : De quoi rit-on ? Femmes et Hommes ont ils le même humour ?

#### **3.1 Construire une nomenclature et choisir les VDM**

Suite à la présentation du travail effectué par l'élève et Mme. Le Donné devant les membres du laboratoire (durant le séminaire interne dont j'ai parlé plus haut), la question de savoir s'il était possible de coder le comique de ces anecdotes (et ainsi de pouvoir énoncé des résultats sur cette variable très difficile à étudier) est venue rapidement et semblait très pertinente. Ma maître

de stage ne pouvait pas prendre le temps de faire se travail dans l'immédiat, je pouvais donc faire un travail complémentaire et innovant sur le projet précédent.

Dans un premier temps il a fallu trouver une nomenclature pour coder le comique représenté dans les anecdotes. Pour cela il a fallu lire plusieurs articles de sociologie. Il existe assez peu de travaux sur le comique en sociologie et aucun (à ma connaissance et à celle des membres du laboratoire) en sociologie quantitative. J'avais naturellement dans l'idée d'utiliser les comiques du théâtre comme base de réflexion. Il était cependant nécessaire d'avoir un argument d'autorité pour justifier le choix de notre nomenclature. L'article (enfin la série d'articles) de Bergson intitulé LE RIRE, m'a fourni cet argument d'autorité. Dans cet ouvrage l'auteur décrit, selon lui, ce que sont les différents types de comiques et les raisons pour lesquelles ils nous font rire. Les comiques énoncés dans le livre de Bergson sont en fait les comiques du théâtre (avec un ou deux comiques en plus).

J'ai pu ainsi avoir une nomenclature précise des comiques que je voulais repérer dans les VDM. Cependant dès la lecture des articles de Bergson j'ai pu voir la difficulté de classer certaine scène dans un seul comique. La lecture des premières VDM m'a conforté dans le fait que le comique reposait rarement sur un seul ressort, mais qu'il en utilisait de nombreux. Je décidais donc de pouvoir coder plusieurs comiques dans une seule anecdote. Cependant les quelques lectures de VDM me montraient que ce codage allait me plonger dans des questions existentielles pour savoir si le comique de la scène venait de la parole, de la posture ou d'un subtil mélange des deux.

Ma maître de stage et moi même avons ensuite décider de tirer 400 anecdotes au hasard (selon une loi uniforme). J'étais chargé de les lire et d'en trouver les différentes formes de comique. Je me suis rapidement rendu compte des difficultés (comme cité au dessus). Je voyais souvent très bien les arguments contre le classement de l'anecdote dans tel comique, mais lorsque je lisais une seconde fois l'anecdote je voyais aussi très bien les arguments qui pouvaient me permettre de la ranger dans ce même comique.

Finalement j'ai pris le partie de ne lire les anecdotes qu'une fois en me disant que leur grand nombre permettrait d'avoir des phénomènes de moyenne et ainsi de corriger mes choix parfois arbitraires.

## 3.2 Analyses et traitements des données

Avant de présenter mon travail il faut revenir sur les principaux résultats du travail qui ont motivé nos recherches. Le site internet viedemerde (VDM) est basé sur le principe de la participation des internautes, qu'ils soient ou non référencés sur le site. Pour proposer une anecdote il suffit de se rendre sur le site et de l'écrire. Quelques informations sont alors demandées : l'âge, le sexe, un pseudonyme, ... cependant ces informations ne sont pas obligatoirement précises (elles ne le sont pas la plus part du temps) ce qui rend les travaux d'analyse d'autant plus compliqués (il est néanmoins nécessaire ranger sa VDM dans une des catégories). L'anecdote ainsi proposée n'est pas rendue immédiatement publique, en effet les modérateurs du site votent pour les meilleures anecdotes, ensuite ce sont les administrateurs du site qui décident de rendre public les

anecdotes proposées par les internautes (plus d'un million d'anecdotes ont été soumises depuis la création du site en 2008 et 16 000 ont été diffusées).

Les premiers résultats de l'étude sont : la quantité d'homme et de femme (dans un quart des anecdotes le sexe est renseigné et les femmes sont deux fois plus nombreuses que les hommes à poster des anecdotes), la proportion d'anecdotes dans chacune des catégories, la proportion d'anecdotes postées par les femmes (respectivement les hommes) dans chacune des catégories, l'évolution du nombre d'anecdotes postées à chaque époque de l'année (depuis la création du site en 2008), la fréquentation du site par les internautes. Des analyses plus fines ont également été réalisées sur les liens entre le sexe et les catégories dans lesquelles les anecdotes ont été postées, ou encore sur le vocabulaire utilisé par les hommes et les femmes pour parler de leur couple, leurs amis, .... Ces résultats (intéressants) n'ont pas besoin d'être présentés en détail, ils m'ont néanmoins servi de socle pour corroborer mon propre travail (notamment pour montrer que mes effectifs étaient bien représentatifs)

Une fois l'ensemble des anecdotes codées le travail d'analyse a pu commencer. Le premier point était de vérifier que notre échantillon était bien représentatif de l'ensemble de notre base de données (constituées de 16 000 VDM). J'ai ainsi vérifié les proportions d'un certain nombre de paramètres. La fréquence des anecdotes anonymes, de sexe non renseigné, posté dans telle région du monde ou de France, de femmes et d'hommes, de la date du post sur internet.

Tous ces paramètres étaient en proportions semblables dans notre échantillon et dans notre base. Ce qui m'a naturellement rassurée sur l'étude de ces anecdotes.

Les anecdotes postées doivent avoir un thème (animaux, amour, enfant, argent, santé, sexe, inclassable ou travail), j'ai comparé l'utilisation de chaque comique dans chacun des thèmes par rapport à l'utilisation de ce comique en général. J'ai ainsi obtenu un tableau qui me donnait les sous ou sur-utilisation de tel comique dans tel thème.

De même pour les différents comiques utilisés par les hommes et les femmes, j'ai pu avoir un tableau du même aspect que le précédent.

Cependant malgré cela les résultats ne semblaient pas exploitables. En effet différents tests de significativité (surtout les tests de Chi 2) ont montré que la très grande majorité de résultats n'étaient pas exploitables (les données étaient trop volatiles pour que l'on puisse vraiment conclure à une sous-utilisation de tel procédé pour tel sexe). Cela n'empêche pas d'essayer de comprendre pourquoi cette significativité n'était pas suffisante. Mme. Le Donné pensait qu'il fallait simplement prendre un plus grand nombre d'anecdotes (pour l'étude de l'humour suivant les sexes seule une centaine d'anecdotes étaient utilisables). Je dois avouer que je me demandais fortement si l'hypothèse, que j'avais faites d'avoir un effet de moyenne en choisissant "rapidement" les comiques qui intervenaient dans chaque VDM, était vraiment réaliste.

Cela ne m'a pas empêché de continuer à étudier la ferveur que provoquait l'utilisation de tel ou tel comique. Pour cela j'ai regardé la quantité des participations des internautes à la VDM et j'ai regardé s'il y avait un lien entre ces participations et la présence d'un comique particulier.

Ces résultats ce sont aussi révélés assez peu intéressants, outre qu'ils n'étaient que très peu significatifs, car ils montraient surtout que les personnes réagissent de manière assez semblable que les comiques soient ou non présents.

Ainsi l'étude du comique des anecdotes s'est révélée assez complexe et ne donnant pas des résultats très marqués. Ceci peut néanmoins s'interpréter : on peut penser que aucune forme particulière du comique n'est avantagée par rapport aux autres (c'est l'événement en lui même qui fait rire et pas le fait d'utiliser tel ou tel forme de comique), ou croire que les petites différences que l'on a remarquées vont s'affirmer au fur et à mesure que l'on choisira un grand nombre de VDM. Naturellement le temps manquait pour essayer de vérifier l'une ou l'autre des hypothèses. Néanmoins ce travail aura permis de cerner les différents problèmes qui interviennent lorsque l'on se lance dans l'étude du comique. Pour les prochaines recherches l'angle d'approche pourra ainsi changer en fonction des objectifs désirés.

## 4 Conclusion

Au cours de ce stage j'ai donc pu me familiariser avec une science tout à fait nouvelle pour moi et pour laquelle je savais que j'avais de l'intérêt. Cependant je n'en ai pas fait un apprentissage classique (souvent théorique et appris dans des livres) mais plus empirique. J'ai pu en percevoir la force et aussi les limites. Les discussions et les questionnements m'ont montré cette discipline comme vivante, intéressante, pragmatique et ont relégué les aspects théoriques à une certaine forme d'ancienneté. Cependant j'ai pu voir en quoi cette ancienneté était vitale dans les cas complexes où le chercheurs s'aventure dans des domaines inexplorés. Se rattachant à ce qu'il connaît, il avance doucement, espère dépasser les vieilles théories ou leur redonner de la jeunesse en assurant une assise solide à l'argumentation qui plaide en leur faveur.

Je retiendrai de ce stage une certaine manière de raisonner, une ouverture d'esprit mais aussi est surtout une capacité à accepter la critique (sur le travail fourni) et à la rendre constructive. Je suis très content d'avoir pu suivre ce stage et d'avoir pu proposer un travail important et approfondi sur deux aspects très différents de la sociologie.



## Deuxième partie

### TERs





## Chapitre 4

# Du théorème de GOLAB

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Quelques définitions et remarques topologiques</b>	<b>1</b>
2.1	Une histoire de convergence d'ensembles . . . . .	1
2.2	Quelques conséquences immédiates . . . . .	3
2.3	Quelques conséquences beaucoup moins évidentes . . . . .	4
<b>3</b>	<b>Un peu de théorie géométrique de la mesure</b>	<b>6</b>
3.1	Premières définitions et lemme de sortie du territoire . . . . .	7
3.2	Théorème de recouvrement de Vitali et théorème de HAUSDORFF-RADON . . . . .	10
3.3	Les connexes de mesures finies sont connexes par arc . . . . .	14
3.4	Théorème de GOLAB . . . . .	18
<b>4</b>	<b>Applications</b>	<b>21</b>
<b>5</b>	<b>Annexe et théorème d'analyse</b>	<b>21</b>
5.1	A propos de la convergence faible d'une suite de mesures (finies) bornées . . . . .	22
5.2	A propos de la différentiabilité des fonctions lipschitziennes . . . . .	23

## 1 Introduction

Ce texte mathématique traite de topologie et de théorie de la mesure. Il n'a pas pour objectif d'être exhaustif ou de traiter en profondeur des aspects théoriques de topologie générale ou de théorie de la mesure. Le but de ces quelques pages est la démonstration du théorème de GOLAB. La démonstration du théorème de GOLAB fait référence à de nombreux résultats intermédiaires de difficultés variables. Nous avons, autant que possible, essayé de réduire ces difficultés. Dans un esprit de clarté le théorème de GOLAB sera considéré comme le Graal de ce texte et apparaîtra donc en fin de la partie qui lui est consacrée, les résultats nécessaires ayant été énoncés et expliqués précédemment. Ce travail repose sur le livre [AT09] (et [Rud98]). sur l'analyse dans les espaces métriques.

## 2 Quelques définitions et remarques topologiques

Le but de cette partie est de munir l'ensemble des parties d'un espace d'une distance, et ainsi de pouvoir parler de convergence pour les ensembles, voire parler de convergence sans avoir pour autant une distance (comme il est possible de le faire dans les espaces topologiques généraux).

### 2.1 Une histoire de convergence d'ensembles

**Définition.** Soit  $(E, d_E)$  un espace métrique et  $(C_n)_{n \in \mathbb{N}}$  une suite de fermés de  $E$

On dit que  $(C_n)_{n \in \mathbb{N}}$  converge vers  $C$  (supposé fermé) au sens de KURATOWSKI si et seulement si :

\* Pour toute suite  $(x_n)_{n \in \mathbb{N}}$  de  $E$  telle que pour tout  $n$  on a  $x_n \in C_n$  alors l'adhérence de  $(x_n)_{n \in \mathbb{N}}$  est incluse dans  $C$ .

\* Tout élément de  $C$  est limite d'une suite telle que décrite au dessus.

**Définition.** Dans un espace métrique  $(E, d_E)$ , soient  $A$  et  $B$  deux sous ensembles de  $E$  compacts. On se propose de poser la quantité :

$$d_H(A, B) = \sup\{\sup_{a \in A}(d_E(a, B)), \sup_{b \in B}(d_E(b, A))\}$$

On l'appelle distance de HAUSDORFF et on vérifie immédiatement qu'il s'agit bien d'une distance : \* Cette quantité est bien sur positive puisque toutes les quantités entrant dans la définition le sont.

\* Elle est symétrique car prendre le maximum de deux quantités est une opération symétrique.

\* Elle est définie :

$$\begin{aligned} d_H(A, B) = 0 &\iff \sup_{a \in A}(d_E(a, B)) = 0, \sup_{b \in B}(d_E(b, A)) = 0 \\ &\iff \forall (a, b) \in A \times B, d_E(a, B) = 0 = d_E(b, A) \\ &\iff \forall (a, b) \in A \times B, a \in B, b \in A \quad (\text{car } A \text{ et } B \text{ sont fermés}) \\ &\iff A = B \end{aligned}$$

\* Elle vérifie l'inégalité triangulaire :

Soient  $A, B, C$  des compacts de  $E$ . Par continuité de la fonction distance et compacité des ensembles on sait que :

$$\exists a_0 \in A, \exists b_0 \in B, d_H(A, B) = d_E(a_0, b_0)$$

Ainsi :  $\forall c \in C, d_E(a_0, b_0) \leq d_E(a_0, c) + d_E(c, b_0)$ .

la partie de droite n'a aucune dépendance en  $c$  on peut donc passer, sans rien changer, à l'infimum et on obtient :

$$d_E(a_0, b_0) \leq d_E(a_0, C) + d_E(C, b_0)$$

De plus par définition de la distance de HAUSDORFF on a :  $d_E(a_0, C) \leq d_H(A, C)$ . Et donc par majoration on arrive à :  $d_H(A, B) = d_E(a_0, b_0) \leq d_H(A, C) + d_H(C, B)$ .

**Proposition.** La quantité  $\|d_E(., A) - d_E(., B)\|_\infty$  est égale à la distance de HAUSDORFF de  $A$  à  $B$ .

*Démonstration.* On procède par double inégalité :

\*  $\forall x \in E, |d_E(x, A) - d_E(x, B)| \leq d_H(A, B)$  :

On se donne  $x$  dans  $E$ . On choisit  $b$  dans  $B$  tel que  $d_E(x, b) = d_E(x, B)$ .

Soit  $a$  appartenant à  $A$  tel que  $d_E(b, a) = d_E(b, A)$ . Alors on a :

$$d_E(x, A) \leq d_E(x, a) \leq d_E(x, b) + d_E(b, a) = d_E(x, B) + d_H(A, B)$$

Et finalement  $d_E(x, A) - d_E(x, B) \leq d_H(A, B)$ .

Par symétrie de la distance (en échangeant les rôles de  $A$  et  $B$ ) on obtient :

$$|d_E(x, A) - d_E(x, B)| \leq d_H(A, B)$$

On vient de montrer l'inégalité pour un  $x$  quelconque de  $E$ , on peut donc passer à la borne supérieure et on a une première inégalité.

$$* d_H(A, B) \leq \sup_{x \in E} (|d_E(x, A) - d_E(x, B)|) :$$

$$\sup_{a \in A} (d_E(a, B)) = \sup_{a \in A} (d_E(a, B) - d_E(a, A)) = \sup_{a \in A} (|d_E(a, B) - d_E(a, A)|) \leq \sup_{x \in E} (|d_E(x, B) - d_E(x, A)|),$$

□

On énonce ensuite le petit lemme suivant qui sera (sans être énoncé) utilisé à de nombreuses reprises.

**Lemme.** Soit  $I$  un ensemble et  $(f_i)_{i \in I}$  une famille de fonctions  $k$ -lipschitzienne alors  $f = \inf_{i \in I} (f_i)$  est également  $k$ -lipschitzienne.

*Démonstration.*  $\forall i \in I, |f_i(x) - f_i(y)| \leq k|x - y|$ . On a donc  $f(x) \leq f_i(x) \leq f_i(y) + k|x - y|$ . On peut donc extraire l'infimum :  $f(x) \leq f(y) + k|x - y|$  et finalement  $f(x) - f(y) \leq k|x - y|$  et par symétrie de la valeur absolue on obtient ce qu'on souhaite :  $|f(x) - f(y)| \leq k|x - y|$ . □

Dans la suite on assimilera la convergence par rapport à la distance de HAUSDORFF à la convergence uniforme (par analogie grâce à la proposition précédente), tandis que la convergence selon KURATOWSKI sera assimilée à la convergence simple (par analogie également).

## 2.2 Quelques conséquences immédiates

Les propriétés suivantes montrent la "force" des deux convergences définies plus haut.

**Proposition.** Si  $(C_n)_{n \in \mathbb{N}}$  converge uniformément alors cette suite converge simplement.

*Démonstration.* On a :  $\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N_\varepsilon, \sup_{x \in E} (|d_E(x, C_n) - d_E(x, C)|) \leq \varepsilon$ . On montre qu'alors les deux conditions de la convergence simple sont vérifiées :

\* Soit  $(x_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$  telle que pour tout  $n \in \mathbb{N}, x_n \in C_n$ .

Soit  $(x_{n_k})_{k \in \mathbb{N}}$  une sous suite convergente, on veut montrer que  $\lim_{k \rightarrow \infty} x_{n_k} \in C$ .

Il est aussi équivalent de dire que  $d_E(\lim_{k \rightarrow \infty} x_{n_k}, C) = 0$  (car  $C$  est fermé).

Et par continuité de la fonction distance on a

$$d_E(\lim_{k \rightarrow \infty} x_{n_k}, C) = \lim_{k \rightarrow \infty} d_E(x_{n_k}, C)$$

Et d'après la convergence de  $(C_n)_{n \in \mathbb{N}}$  vers  $C$  on a que pour un  $\varepsilon$  fixé il existe un  $n$  assez grand tel que

$$|d_E(x_n, C_n) - d_E(x_n, C)| \leq d_H(C_n, C) \leq \varepsilon$$

et donc pour  $k$  assez grand on a

$$|d_E(x_{n_k}, C_{n_k}) - d_E(x_{n_k}, C)| \leq d_H(C_{n_k}, C) \leq \varepsilon$$

et comme  $x_{n_k} \in C_{n_k}$  on a finalement que

$$|d_E(x_{n_k}, C)| \leq \varepsilon,$$

\* Soit  $x$  appartenant à  $C$ , montrons que  $x$  est limite d'une suite ayant les mêmes propriétés que plus haut. On a par définition de la convergence uniforme :

$$\forall k \in \mathbb{N}, \exists N_k \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N_k \Rightarrow d_H(C_n, C) \leq 1/k$$

On pose  $N_0 = 0$  et on obtient une suite  $(N_k)_{k \in \mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$  que l'on peut considérer comme strictement croissante. On a alors

$$\forall N_k \leq n \leq N_{k+1}, |d_E(x, C_n) - d_E(x, C)| \leq 1/k$$

On a de plus  $d_E(x, C) = 0$  et donc pour tout  $n$  compris entre  $N_k$  et  $N_{k+1}$  on peut trouver un  $x_n$  appartenant à  $C_n$  et tel qu'il soit à une distance inférieure à  $1/k$ . En généralisant ce procédé à tout intervalle  $N_k, N_{k+1}$  on obtient bien une suite convergent vers  $x$  et dont chacun des termes appartient à chacun des ensembles de la suite  $(C_n)_{n \in \mathbb{N}}$ . □

**Proposition.** *Si  $(E, d_E)$  est un espace compact, si  $(C_n)_{n \in \mathbb{N}}$  est une suite convergent simplement vers un ensemble  $C$  fermé (et donc compact) alors elle converge uniformément (ce théorème rappelle fortement le théorème de Dini).*

*Démonstration.* On veut montrer que :

$$\lim_{n \rightarrow \infty} d_H(C_n, C) = \lim_{n \rightarrow \infty} \sup_{z \in C_n} (d_E(z, C)), \sup_{x \in C} (d_E(x, C_n)) = 0$$

Par compacité on sait qu'il existe  $z_n \in C_n$  tel que  $\max_{z \in C_n} (d_E(z, C)) = d_E(z_n, C)$ . Et de même on a  $z \in C$  tel que  $\max_{x \in C_n} (d_E(x, C_n)) = d_E(z, C_n)$ . Or on sait déjà que  $\lim_{n \rightarrow \infty} d_E(z, C_n) = 0$  car par convergence simple on a que  $z$  est limite d'une suite d'élément des ensembles  $C_n$ . Si de plus on avait  $\lim_{n \rightarrow \infty} d_E(z_n, C) = 0$  alors on en déduirait que  $\lim_{n \rightarrow \infty} d_H(C_n, C) = 0$ . On sait, par convergence simple, que l'adhérence de  $(z_n)_{n \in \mathbb{N}}$  est incluse dans  $C$ . Si  $d_E(z_n, C)$  ne convergeait pas vers 0 alors on pourrait extraire une sous suite  $(z_{n_k})_{k \in \mathbb{N}}$  qui vérifie :

$$\exists \varepsilon > 0, \forall k \in \mathbb{N}, d_E(z_{n_k}, C) > \varepsilon$$

Et par compacité de  $E$  on peut extraire une sous suite de  $(z_{n_k})_{k \in \mathbb{N}}$  qui converge, et par convergence simple de la suite d'ensembles cette limite appartient à  $C$ , ce qui est contradictoire avec les propriétés de  $(z_{n_k})_{k \in \mathbb{N}}$ . On en déduit qu'une telle extraction n'existe pas et donc on a bien

$$\lim_{n \rightarrow \infty} d_E(z_n, C) = 0$$

Tout ceci nous permet d'avoir le résultat recherché. □

## 2.3 Quelques conséquences beaucoup moins évidentes

Les conséquences liées à la définition de la distance de HAUSDORFF nous donne des propriétés de "stabilité" importantes.

**Proposition.** *Si  $(E, d_E)$  est un espace compact, on note  $K_E$  l'ensemble des parties compacts non vides de  $E$ . Alors  $(K_E, d_H)$  est un espace compact.*

*Démonstration.* On a déjà que  $(K_E, d_H)$  est un espace métrique. On se donne  $(A_n)_{n \in \mathbb{N}} \in K_E^{\mathbb{N}}$  et on pose les fonctions associées :

$$f_n(x) = d_E(x, A_n)$$

La famille  $(f_n)_{n \in \mathbb{N}}$  est une famille de fonctions 1-lipschitziennes, on a donc une famille de fonctions équicontinue. On appelle  $D = \sup_{(x,y) \in E^2} (d_E(x, y))$  qui est fini car  $E$  est compact, on a alors :  $\forall n \in \mathbb{N}, \sup_{x \in E} (f_n(x)) \leq D$ .

On en conclut que  $\Phi(x) = \{f_n(x), n \in \mathbb{N}\}$  est borné, ainsi il est relativement compact pour tout  $x$  appartenant à  $E$ . On peut alors appliquer le théorème de Ascoli-Arzelà à la famille  $(f_n)_{n \in \mathbb{N}}$  de fonctions. On a ainsi l'existence d'une suite extraite qui converge uniformément vers une fonction  $f : E \rightarrow \mathbb{R}_+$ . L'objectif est de montrer qu'une suite de compact converge uniformément vers un compact. Ainsi si la fonction  $f$  était de la forme  $f(x) = d_E(x, A)$ , avec  $A$  un compact, alors on aurait  $d_H(A_n, A) = \sup_{x \in E} (|f_n(x) - f(x)|)$  et donc la suite de compact de  $E$  convergerait uniformément vers  $A$ .

On va donc montrer que  $f$  est bien de la forme espérée. On procède donc par analyse-synthèse : si  $f$  était de la forme souhaitée alors on aurait  $A = f^{-1}(\{0\})$ . On voit que  $A$  est fermé comme image réciproque d'un fermé par une application continue, il est même compact car  $E$  est compact. On montre ensuite par double inégalité que  $f$  est bien la distance cherchée :

\*  $f(x) \leq d_E(x, A)$  :  
 $f$  est 1-lipschitzienne comme limite de fonctions 1-lipschitziennes :

$$\forall a \in A, |f(x) - f(a)| \leq d_E(x, a)$$

Et par définition de l'ensemble  $A$  on a  $f(a) = 0$ . On obtient donc l'inégalité :  $|f(x)| \leq d_E(x, a)$ . Puis par passage à la borne inférieure

$$|f(x)| \leq d_E(x, A)$$

\*  $f(x) \geq d_E(x, A)$  :

$$\forall n \in \mathbb{N}, \exists a_n \in A_n, f_n(x) = d_E(x, a_n)$$

et donc  $f_n(a_n) = 0$ .  $(a_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$  donc quitte à extraire une sous suite on peut supposer que cette suite converge vers  $a$ . On montre ensuite que la suite  $(f_n(a_n))_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$  converge vers  $f(a)$ . En effet : soit  $\varepsilon$  strictement positif, il existe un rang à partir duquel  $\sup_{x \in E} (|f_n(x) - f(x)|)$  est plus petit que  $\varepsilon$ . De plus comme  $f$  est continue alors il existe un autre rang à partir duquel  $|f(a_n) - f(a)|$  est également plus petit que  $\varepsilon$ . On en déduit immédiatement que pour un rang assez grand on a :

$$|f_n(a_n) - f(a)| \leq |f_n(a_n) - f(a_n)| + |f(a_n) - f(a)| \leq 2\varepsilon$$

Finalement

$$\lim_{n \rightarrow \infty} f_n(a_n) = f(a) = 0, \text{ ainsi } a \in A$$

Par conséquent

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} d_E(x, a_n) = d_E(x, a) \geq d_E(x, A)$$

On a donc bien  $f(x) = d_E(x, A)$  où  $A = f^{-1}(\{0\})$ .

On montre très rapidement que  $A$  n'est pas vide par l'existence de la limite de la suite  $(a_n)_{n \in \mathbb{N}}$  précédente qui appartient à  $A$ . □

**Théorème.** Si  $(E, d_E)$  est un espace compact et  $(C_n)_{n \in \mathbb{N}} \in K_E^{\mathbb{N}}$  une suite de connexes fermés qui converge uniformément vers  $C$ , alors  $C$  est connexe.

*Démonstration.* On raisonne par l'absurde : Soient  $A$  et  $B$  deux ouverts de  $E$  tels que

$$C \subset A \cup B \text{ et } (C \cap A \neq \emptyset \text{ et } C \cap B \neq \emptyset)$$

On va montrer alors qu'il existe  $n_0 \in \mathbb{N}$  tel que  $C_{n_0}$  n'est pas connexe. On va simplement montrer que  $C_{n_0} \subset A \cup B$  et  $(C_{n_0} \cap A \neq \emptyset \text{ et } C_{n_0} \cap B \neq \emptyset)$ .

\* On commence par montrer qu'à partir d'un certain rang on a nécessairement  $C_n \cap A \neq \emptyset$ . En effet si  $a \in A \cap C$  alors

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N}, \forall n \geq N_\varepsilon \Rightarrow d_H(C_n, C) \leq \varepsilon$$

Et alors :  $d_E(a, C_n) \leq d_H(C, C_n) \leq \varepsilon$  pour un  $n$  assez grand. De plus  $A$  est un ouvert donc il existe  $\rho$  positif tel que  $B(a, \rho) \subset A$ . Donc il existe un rang  $N_{\rho/2}$  à partir duquel on a

$$d_E(a, C_n) \leq \rho/2$$

Par conséquent il existe  $a_n \in C_n$  tel que  $d_E(a, a_n) < \rho$ , on en déduit alors que  $a_n \in B(a, \rho) \subset A$  et finalement  $C_n \cap A \neq \emptyset$  à partir d'un rang. De manière analogue on obtient qu'à partir d'un certain rang on a

$$C_n \cap B \neq \emptyset$$

\* On montre ensuite qu'à partir d'un certain rang on a :  $C_n \subset A \cup B$ . On pose  $K = \overline{C}_E \cap A \cup B$ , et on veut qu'à partir d'un rang on ait

$$\forall x \in C_n, d_E(x, K) > 0$$

On aurait alors  $C_n \cap K = \emptyset$  (car  $K$  est un fermé de  $E$ ).

On a déjà que  $d = \inf_{c \in C} (d_E(c, K)) > 0$  car  $C \cap K = \emptyset$ , que  $K$  est fermé et  $C$  est compact. Soit  $a_n \in C_n$  et  $k_{a_n} \in K$  tels que  $d_E(a_n, k_{a_n}) = d_E(a_n, K)$ . On a alors :

$$\forall c \in C, d \leq d_E(c, k_{a_n}) \leq d_E(c, a_n) + d_E(a_n, k_{a_n})$$

On peut passer à l'infimum (car  $d$  ne dépend pas du choix de l'élément de  $C_n$ ) et on obtient :

$$d \leq d_E(c, C_n) + d_E(a_n, K) \leq d_H(C, C_n) + d_E(a_n, K)$$

Finalement  $d_E(a_n, K) \geq d - d_H(C, C_n)$ . On choisit donc  $\varepsilon = d/2$  et il existe  $N_\varepsilon$  tel que pour tout entier naturel plus grand on a :  $d_H(C_n, C) \leq d/2$  et donc  $d_E(a_n, K) \geq d/2 > 0$ . On a donc bien ce que l'on souhaite : tout élément de  $C_n$  est à une distance strictement positive de  $K$  ainsi il appartient au complémentaire de  $K$  ainsi on a que  $C_n$  est bien inclus dans l'union de deux ouverts disjoints non vides.

□

### 3 Un peu de théorie géométrique de la mesure

Dans un premier temps on va introduire la mesure de HAUSDORFF puis quelques quantités en lien avec les mesures. Nous allons ensuite donner une suite de lemmes, propriétés, théorèmes plus ou moins généraux sur les mesures. On passera ensuite à deux théorèmes "géométriques" sur les liens entre les connexes de mesure de HAUSDORFF finie et les connexes par arc. Enfin le théorème de GOLAB qui sera chargé de mettre tout cela de concert pour donner le résultat attendu.

### 3.1 Premières définitions et lemme de sortie du territoire

**Définition.** On pose  $\omega_k = \frac{\pi^{k/2}}{\Gamma(1+k/2)}$  où  $\Gamma$  est la fonction factorielle d'Euler. On pose ensuite :

$$\mathcal{H}_\delta^k(A) = \frac{\omega_k}{2^k} \inf \left\{ \sum_{i \in I} \text{diam}(A_i)^k, \text{diam}(A_i) \leq \delta, A \subseteq \bigcup_{i \in I} A_i \right\}$$

Et finalement on pose  $\mathcal{H}^k(A) = \sup_{\delta > 0} \mathcal{H}_\delta^k(A)$  que l'on appelle mesure de HAUSDORFF de  $A$ .

Il faut ensuite montrer que l'on a bien une mesure :

\*  $\mathcal{H}^k(\emptyset) = 0$  car  $\forall \delta > 0, \mathcal{H}_\delta^k(\emptyset) = 0$  et donc en passant à la borne supérieure on a ce qu'on souhaite,

\* Si on se donne une famille dénombrable d'ensembles  $(A_i)_{i \in \mathbb{N}}$  alors on a que pour un recouvrement de  $\bigcup_{i \geq 0} A_i$  à  $\delta$ -près on a un recouvrement de chacun des éléments  $A_i$ , ainsi

$$\mathcal{H}_\delta^k \left( \bigcup_{i \geq 0} A_i \right) \leq \sum_{i \geq 0} \mathcal{H}_\delta^k(A_i)$$

puis par passage à la borne supérieure on a

$$\mathcal{H}^k \left( \bigcup_{i \geq 0} A_i \right) \leq \sum_{i \geq 0} \mathcal{H}^k(A_i)$$

On a même un peut mieux si l'on considère la mesure de HAUSDORFF sur l'espace  $(K_E, d_H)$ , en effet on a la formule d'égalité :  $\mathcal{H}^k(A \cup B) = \mathcal{H}^k(A) + \mathcal{H}^k(B)$  lorsque  $A$  et  $B$  sont disjoints. Il suffit en effet de prendre un recouvrement de  $A$  et de  $B$  à  $\inf_{a \in A, b \in B} (d_E(a, b))/2$ -près et alors on a que :  $\mathcal{H}^k(A) + \mathcal{H}^k(B) \geq \mathcal{H}^k(A \cup B)$  ce qui nous donne finalement l'égalité recherchée. Dans la suite de ce projet nous n'avons pas réellement besoin que la mesure de HAUSDORFF soit autre chose qu'une mesure extérieure, on peut donc s'arrêter ici sur les propriétés élémentaires de cette mesure.

**Définition.** On définit les quantités suivantes :

$$\overline{\Theta}_k(\mu, x) = \overline{\lim}_{r \rightarrow 0^+} \frac{\mu(B(x, r))}{\omega_k r^k}$$

$$\underline{\Theta}_k(\mu, x) = \underline{\lim}_{r \rightarrow 0^+} \frac{\mu(B(x, r))}{\omega_k r^k}$$

Que l'on appelle respectivement la densité  $k$ -dimensionnelle supérieure ou inférieure de la mesure  $\mu$ . On pose de plus ce qu'on appelle les densités sphériques supérieure et inférieure d'un ensemble  $B$  qui valent :

$$\overline{\Theta}_k(B, x) = \overline{\Theta}_k(\mathcal{H}_B^k, x) = \overline{\lim}_{r \rightarrow 0^+} \frac{\mathcal{H}^k(B \cap B(x, r))}{\omega_k r^k}$$

$$\underline{\Theta}_k(B, x) = \underline{\Theta}_k(\mathcal{H}_B^k, x) = \underline{\lim}_{r \rightarrow 0^+} \frac{\mathcal{H}^k(B \cap B(x, r))}{\omega_k r^k}$$



Il est à noter que l'indice inférieur lié à la mesure de HAUSDORFF renvoie à la définition de mesure restreinte et n'est pas du même ordre que l'indice  $\delta$  utilisé pour définir cette mesure.

**Remarque :** Dans le cas particulier où  $k$  vaut 1, 2 ou 3 on constate que la mesure de HAUSDORFF correspond respectivement à la longueur, l'aire ou le volume de l'ensemble étudié (On peut notamment remarquer que dans le cas particulier de  $\mathbb{R}$  les mesures de HAUSDORFF et de Lebesgue correspondent). L'intérêt de cette mesure est par exemple de pouvoir calculer l'aire d'une sous variété (par exemple un cercle) dans un espace de dimension supérieure (dans  $\mathbb{R}^3$ ), ou encore de mesurer les tailles d'ensembles en dehors des espaces communs  $\mathbb{R}^n$ .

**Lemme.** Soit  $(E, d_E)$  un espace métrique. Si on se donne une suite de mesures finies  $(\mu_n)_{n \in \mathbb{N}}$  qui converge faiblement vers une mesure  $\mu$  alors pour tout fermé  $\mathcal{F}$  on a

$$\mu(\mathcal{F}) \geq \overline{\lim}_{n \rightarrow \infty} \mu_n(\mathcal{F}) \text{ et } \mu(\mathcal{O}) \leq \underline{\lim}_{n \rightarrow \infty} \mu_n(\mathcal{O})$$

*Démonstration.* On commence par montrer le lemme pour  $\mathcal{F}$  fermé puis par passage au complémentaire on pourra le montrer pour  $\mathcal{O}$  ouvert.

On pose la fonction  $\phi_\varepsilon(x) = \max\{0, 1 - \frac{1}{\varepsilon}d_E(x, F)\}$ .

On a donc

$$\forall \varepsilon > 0, \forall x \in E, \mathbb{1}_{\mathcal{F}}(x) \leq \phi_\varepsilon(x)$$

Ainsi

$$\int_E \mathbb{1}_{\mathcal{F}} d\mu_n = \mu_n(\mathcal{F}) \leq \int_E \phi_\varepsilon d\mu_n$$

et

$$\phi_\varepsilon \in \mathcal{C}_b(E, \mathbb{R})$$

donc

$$\overline{\lim}_{n \rightarrow \infty} \mu_n(\mathcal{F}) \leq \overline{\lim}_{n \rightarrow \infty} \int_E \phi_\varepsilon d\mu_n = \int_E \phi_\varepsilon d\mu$$

On a de plus

$$\lim_{\varepsilon \rightarrow 0} \phi_\varepsilon(x) = \mathbb{1}_{\mathcal{F}}(x)$$

et donc par le théorème de convergence dominée (que l'on peut utiliser car la mesure  $\mu$  est finie et que la famille  $(\phi_\varepsilon)_{\varepsilon > 0}$  est bornée par 1 pour tout  $\varepsilon$ ) on a

$$\lim_{\varepsilon \rightarrow 0} \int_E \phi_\varepsilon d\mu = \int_E \mathbb{1}_{\mathcal{F}} d\mu = \mu(\mathcal{F})$$

Finalement on a bien l'inégalité recherchée.

On se donne  $\mathcal{O}$  un ouvert. On a alors

$$\mu(\mathbb{C}_E \mathcal{O}) \geq \overline{\lim}_{n \rightarrow \infty} \mu_n(\mathbb{C}_E \mathcal{O})$$

Donc :

$$\mu(E) - \mu(\mathcal{O}) \geq \overline{\lim}_{n \rightarrow \infty} \mu_n(E) - \mu_n(\mathcal{O})$$

ainsi

$$\mu(\mathcal{O}) \leq - \overline{\lim}_{n \rightarrow \infty} -\mu_n(\mathcal{O})$$

et

$$- \overline{\lim}_{n \rightarrow \infty} -\mu_n(\mathcal{O}) = \underline{\lim}_{n \rightarrow \infty} \mu_n(\mathcal{O})$$

finalement on a bien ce qu'on souhaite. □

Nommer le lemme suivant est surtout utile parce qu'on y fera abondamment référence et qu'il est plus agréable d'avoir un nom que de dire : "d'après le lemme, vous vous rappelez le truc bizarre qui énonce quelque chose de quelconque!!".

**Lemme** (de sortie du territoire). *Soit  $(E, d_E)$  un espace métrique et  $C$  un connexe de  $E$ . De plus si  $0 < r < \text{diam}(C)/2$  alors*

$$\forall x \in E, \mathcal{H}^1(C \cap B(x, r)) \geq r$$

*Démonstration.* On fait, dans un premier temps, remarquer que si  $\phi$  est une fonction  $k$ -lipschitzienne alors

$$\mathcal{H}^1(\phi(C \cap B(x, r))) \leq k\mathcal{H}^1(C \cap B(x, r))$$

En effet on a  $\forall (x, y) \in E^2, d_E(\phi(x), \phi(y)) \leq kd_E(x, y)$  ainsi si on se donne  $(A_i)_{i \in I}$  un recouvrement de  $C \cap B(x, r)$  alors on peut construire un recouvrement de  $\phi(C \cap B(x, r))$  en prenant  $(\phi(A_i))_{i \in I}$ , car si  $z \in C \cap B(x, r)$  alors il existe  $i_0$  tel que  $z \in A_{i_0}$  et alors  $\phi(z) \in \phi(A_{i_0})$  et  $\phi(z) \in \phi(C \cap B(x, r))$  on a donc bien un recouvrement. On fait remarquer que  $\omega_1/2 = 1$  et on obtient la suite d'inégalité suivante :

$$\begin{aligned} \mathcal{H}_\delta^1(A) &= \inf \left\{ \sum_{i \in I} \text{diam}(A_i), \text{diam}(A_i) \leq \delta, A \subseteq \bigcup_{i \in I} A_i \right\} \\ &\geq \inf \left\{ \sum_{i \in I} \frac{\text{diam}(\phi(A_i))}{k}, \text{diam}(A_i) \leq \delta, A \subseteq \bigcup_{i \in I} A_i \right\} \\ &\geq \frac{1}{k} \inf \left\{ \sum_{i \in I} \text{diam}(\phi(A_i)), \text{diam}(A_i) \leq \delta, A \subseteq \bigcup_{i \in I} A_i \right\} \\ &\geq \frac{1}{k} \inf \left\{ \sum_{i \in I} \text{diam}(\phi(A_i)), \text{diam}(\phi(A_i)) \leq k\delta, \phi(A) \subseteq \bigcup_{i \in I} \phi(A_i) \right\} \\ &\geq \frac{1}{k} \inf \left\{ \sum_{i \in I} \text{diam}(\phi(A_i)), \text{diam}(B_i) \leq k\delta, \phi(A) \subseteq \bigcup_{i \in I} B_i \right\} \\ &\geq \frac{1}{k} \mathcal{H}_{k\delta}^1(\phi(A)) \end{aligned}$$

Puis par passage à la limite on a :

$$k\mathcal{H}^1(C \cap B(x, r)) = k \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^1(C \cap B(x, r)) \geq \lim_{\delta \rightarrow 0} \mathcal{H}_{k\delta}^1(\phi(C \cap B(x, r))) = \mathcal{H}^1(\phi(C \cap B(x, r)))$$

On pose alors

$$\forall y \in C, \forall x \in C, \phi_x(y) = d_E(x, y)$$

qui est une application 1-lipschitzienne donc

$$\mathcal{H}^1(\phi_x(C \cap B(x, r))) \leq \mathcal{H}^1(C \cap B(x, r))$$

On a que  $\phi_x(C \cap B(x, r)) = \phi_x(C) \cap [0, r[$  (résultat que l'on prouve par double inclusion) est un connexe de  $\mathbb{R}$ , c'est donc un intervalle de la forme  $[0, a[$  où  $a < +\infty$  (l'intervalle est ouvert ou fermé à droite, en revanche il est borné car  $C \cap B(x, r)$  est borné). Montrons en fait que  $a = r$ , en effet on a que  $r < \text{diam}(C)/2$  et donc il existe  $y$  tel que  $d_E(x, y) = \phi_x(y) > r$  (sinon on aurait

que  $C$  est compris dans une boule de diamètre strictement inférieur au diamètre de  $C$ ) et ainsi  $[0, r] \subset \phi_x(C)$ . On en déduit donc que  $a = r$  et on a ainsi ce qu'on cherchait car

$$\mathcal{H}^1(\phi(C \cap B(x, r))) = a = r \leq \mathcal{H}^1(C \cap B(x, r))$$

□

**Remarque :** Le lemme précédent est en fait une mise en forme formelle d'une idée assez simple à se représenter dans le plan. Si on se donne une courbe du plan et une boule centrée en l'un des points de la courbe et dont le diamètre est assez petit, alors la courbe sortira de la boule et il lui faudra au moins parcourir la distance du centre à la frontière de la boule. Les arguments utilisés dans la preuve sont eux mêmes ceux que l'on serait tenté d'utiliser pour le cas réel : on regarde quelle est la valeur de la distance du centre de la boule au point de la courbe le plus éloigné appartenant à la boule et on montre que cette valeur vaut en fait le rayon de la boule par la connexité de la courbe. Et comme dans le plan le plus court chemin est la droite alors la distance parcourue par la courbe dans la boule est forcément plus grande que le rayon de la boule. Une autre image : si vous voulez sortir du territoire français par la route vous devrez, obligatoirement, parcourir au moins la distance qui vous sépare de la frontière la plus proche.

### 3.2 Théorème de recouvrement de Vitali et théorème de HAUSDORFF-RADON

**Définition.** On dit que  $\mathfrak{F}$  est un recouvrement fin de  $B$  si et seulement si  $\mathfrak{F}$  est une famille de fermés et  $\forall \varepsilon > 0, \forall x \in B, \exists V \in \mathfrak{F}, \text{diam}(V) < \varepsilon$  et  $x \in V$ .

**Théorème (VITALI).** : Si  $(E, d_E)$  est un espace métrique et  $B$  une partie de  $E$ . Soient  $\alpha \geq 0$  et  $\mathfrak{F}$  un recouvrement fin de  $B$ . Alors il existe  $(V_i)_{i \in \mathbb{N}}$  une famille d'éléments de  $\mathfrak{F}$  finie ou dénombrable et disjointe telle que l'une, au moins, des conditions suivantes est vérifiée :

$$\sum_{i=0}^{\infty} \text{diam}(V_i)^\alpha = \infty$$

$$H^\alpha(B \setminus \bigcup_{i \in \mathbb{N}} V_i) = 0$$

*Démonstration.* La démonstration du théorème de Vitali est constructive : elle va clairement exhiber une suite  $(V_i)_{i \in \mathbb{N}}$  qui vérifiera au moins l'une des deux conditions. Pour construire cette suite on se contente de choisir successivement des éléments de  $\mathfrak{F}$  tels qu'ils recouvrent suffisamment  $B$  et qu'ils soient disjoints des précédents. On montre ensuite que cette manière suffit à trouver une telle suite.

On commence par faire remarquer qu'il est possible de borner l'ensemble

$$\mathfrak{E} = \{\text{diam}(V), V \in \mathfrak{F}\}$$

en effet si  $\mathfrak{F}$  est un recouvrement fin alors

$$\tilde{\mathfrak{F}} = \{V, V \in \mathfrak{F}, \text{diam}(V) \leq \beta\}$$

est également un recouvrement fin (pour tout  $\beta$  positif). On en déduit que quitte à tronquer  $\mathfrak{F}$  on peut supposer que  $\mathfrak{E}$  est borné. On choisit alors un élément  $V_1$  tel que

$$\text{diam}(V_1) \geq \frac{1}{2} \sup \mathfrak{E}$$

On pose ensuite

$$\mathfrak{F}_1 = \{V \in \mathfrak{F}, V \cap V_1 = \emptyset\}, \mathfrak{E}_1 = \{\text{diam}(V), V \in \mathfrak{F}_1\} \text{ et } d_1 = \sup \mathfrak{E}_1$$

On choisit ensuite un élément  $V_2 \in \mathfrak{F}_1$  tel que

$$\text{diam}(V_2) \geq \frac{1}{2}d_1$$

On continue ainsi de suite en posant pour tout entier naturel  $k$ ,

$$\mathfrak{F}_k = \{V \in \mathfrak{F}, V \cap \bigcup_{i=1}^k V_i = \emptyset\}, \mathfrak{E}_k = \{\text{diam}(V), V \in \mathfrak{F}_k\}, d_k = \sup \mathfrak{E}_k \text{ et } V_{k+1} \in \mathfrak{F}_k$$

tel que  $\text{diam}(V_{k+1}) \geq \frac{1}{2}d_k$ . On vient de construire une suite qui doit recouvrir  $B$ , on va voir que celle ci recouvre en effet presque entièrement  $B$  (c'est à dire à un ensemble de mesure nulle près). Plusieurs cas peuvent se produire :

\* Si le processus s'arrête (i.e : s'il n'est plus possible de choisir des éléments  $V_{k+1}$  qui vérifient ce que l'on souhaite) alors on a un recouvrement de  $B$ . Si tel n'était pas le cas alors on aurait  $(V_i)_{1 \leq i \leq n}$  notre suite ainsi construite,

$$x \in B \setminus \bigcup_{i=1}^n V_i \text{ et } d_x = \min(d_E(x, V_i), 1 \leq i \leq n)$$

Or chacun des  $V_i$  est fermé donc on a  $d_x > 0$ , on peut donc choisir un élément  $V$  de  $\mathfrak{F}$  tel que  $\text{diam}(V) < d_x$  et  $x \in V$ . Dans ce cas on a que

$$V \cap \bigcup_{i=1}^n V_i = \emptyset$$

ainsi  $\mathfrak{F}_n$  n'est pas vide, donc  $\mathfrak{E}_n$  non plus, et on peut donc extraire un élément  $V_{n+1}$  ce qui est contradictoire avec l'hypothèse selon laquelle la construction ne pouvait aller plus loin. Ainsi on a un recouvrement de  $B$  et donc

$$H^\alpha(B \setminus \bigcup_{i=1}^n V_i) = 0$$

\* Si la construction ne s'arrête pas et de plus que  $\sum_{i=1}^{\infty} \text{diam}(V_i)^\alpha < \infty$ , montrons alors que  $(V_i)_{i \in \mathbb{N}}$  recouvre presque entièrement  $B$ . On commence par montrer que tout élément  $x$  de  $B$  se laisse approcher autant qu'on le veut par la suite  $(V_i)_{i \in \mathbb{N}}$ . On se donne  $x \in B \setminus \bigcup_{i=1}^k V_i$  pour un entier  $k$  fixé arbitraire. Comme  $\mathfrak{F}$  est un recouvrement fin on a l'existence d'un élément  $V \in \mathfrak{F}_k$  tel que

$$\text{diam}(V) \leq d_k < 2\text{diam}(V_{k+1})$$

et qui contient  $x$  (en effet il suffit de faire remarquer que  $d_x = \min_{1 \leq i \leq k} (d_E(x, V_i)) > 0$  car les  $V_i$  sont des fermés et de choisir  $V \in \mathfrak{F}$  tel que  $\text{diam}(V) < d_x$ ).

Supposons alors que pour tout entier  $n > k$  on ait  $V \cap \bigcup_{i=1}^n V_i = \emptyset$ . On aurait alors que  $V \in \mathfrak{F}_n$  pour tout  $n > k$  et donc que

$$\text{diam}(V) \leq d_n < 2\text{diam}(V_{n+1})$$

Or par convergence de  $\sum_{i=1}^{\infty} \text{diam}(V_i)^\alpha$  on a que  $\lim_{i \rightarrow \infty} \text{diam}(V_i) = 0$  et donc à partir d'un certain rang on a nécessairement  $2\text{diam}(V_n) \leq \text{diam}(V)$  ce qui est contradictoire. On en déduit qu'il existe un entier  $n > k$  tel que  $V \cap \bigcup_{i=1}^n V_i \neq \emptyset$  et cela quelque soit le diamètre de  $V$ . Si  $y \in V \cap V_n$  on a alors :

$$d_E(x, V_n) \leq d_E(x, x_n) \leq d_E(x, y) + d_E(y, x_n) \leq \text{diam}(V) + \text{diam}(V_n) \leq 3\text{diam}(V_n)$$

On en déduit que pour  $k$  bien choisi on peut approcher  $x$  aussi près que l'on veut par la suite  $(V_i)_{i \in \mathbb{N}}$ . On a que quelque soit  $(x_i)_{i \in \mathbb{N}} \in E^{\mathbb{N}}$  telle que pour tout  $i$  on ait  $x_i \in V_i$  alors  $(B(x_i, 3\text{diam}(V_i)))_{i \geq k+1}$  recouvre  $B \setminus \bigcup_{i=1}^k V_i$ . Donc

$$\forall \delta > 0, H_\delta^\alpha(B \setminus \bigcup_{i=1}^k V_i) \leq H_\delta^\alpha\left(\bigcup_{i \geq k+1} B(x_i, 3\text{diam}(V_i))\right) \leq \sum_{i \geq k+1} H_\delta^\alpha(B(x_i, 3\text{diam}(V_i)))$$

On peut choisir  $k$  tel que  $3\text{diam}(V_i) < \delta, \forall i > k$ . En remarquant de plus que :

$$B \setminus \bigcup_{i \geq 1} V_i \subset B \setminus \bigcup_{i=1}^k V_i$$

$$H_\delta^\alpha(B \setminus \bigcup_{i \geq 1} V_i) \leq H_\delta^\alpha(B \setminus \bigcup_{i=1}^k V_i) \leq \sum_{i \geq k+1} H_\delta^\alpha(B(x_i, 3\text{diam}(V_i))) \leq \sum_{i=k+1}^{\infty} (3\text{diam}(V_i))^\alpha < \infty$$

Puis en prenant la limite lorsque  $k$  tend vers l'infini on a :

$$H_\delta^\alpha(B \setminus \bigcup_{i \geq 1} V_i) \leq \lim_{k \rightarrow \infty} \sum_{i=k+1}^{\infty} (3\text{diam}(V_i))^\alpha = 0$$

On peut ensuite faire tendre  $\delta$  vers 0 et on obtient le résultat attendu. □

**Remarque :** Le théorème de Vitali ne fait qu'énoncer le fait qu'il est toujours possible de recouvrir presque entièrement tout ensemble possédant un recouvrement fin. Ainsi on peut par un procédé explicite et "techniquement faisable" trouver une famille d'ensemble disjoints qui recouvre aussi précisément qu'on le souhaite notre ensemble de départ.

**Théorème (HAUSDORFF-RADON).** *Si  $\mu$  est une mesure finie sur  $E$  (ou  $E$  est bien sûr un espace métrique, condition nécessaire puisqu'on parle de la mesure de HAUSDORFF) alors pour tout  $t \in [0, +\infty[$  et  $B \in \mathcal{B}(E)$  on a :*

$$\overline{\Theta}_k(\mu, x) \geq t, \forall x \in B \Rightarrow \mu(B) \geq t\mathcal{H}^k(B)$$

*Démonstration.* On commence par supposer que  $t = 1$  et que  $B$  est borné. On se donne  $\delta \in [0, 1]$  et  $A$  un ouvert borné contenant  $B$ , on pose :

$$\mathfrak{C}_\delta = \{B(x, r), B(x, r) \subset A, x \in B, r < \delta, \mu(B(x, r)) \geq (1 - \delta)\omega_k(\text{diam}(B(x, r)))^k / 2^k\}$$

et on vérifie alors deux choses : la première que  $\mathfrak{C}_\delta$  est non vide et la seconde que cet ensemble est un recouvrement fin de  $B$ . On a

$$\overline{\Theta}_k(\mu, x) = \overline{\lim}_{r \rightarrow 0^+} \frac{\mu(B(x, r))}{\omega_k r^k} \geq 1$$

ainsi on a :

$$\exists (r_i)_{i \in \mathbb{N}} \in \mathbb{R}_+^*, r_i \rightarrow 0, \forall i \in \mathbb{N}, \frac{\mu(B(x, r_i))}{\omega_k r_i^k} \geq (1 - \delta)$$

et donc

$$\mu(B(x, r_i)) \geq (1 - \delta) \omega_k r_i^k \text{ et } r_i = \text{diam}(B(x, r_i))/2$$

On en déduit :

$$\forall i \in \mathbb{N}, \mu(B(x, r_i)) \geq (1 - \delta) \omega_k (\text{diam}(B(x, r_i)))^k / 2^k$$

Ainsi on a bien l'existence d'un élément dans l'ensemble  $\mathfrak{C}_\delta$  qui est donc non vide, on a même mieux : c'est qu'un ensemble de boules de suite de rayons convergent vers 0 appartient à  $\mathfrak{C}_\delta$  et ainsi on a un recouvrement fin de  $B$ . On construit  $(C_i)_{i \in \mathbb{N}}$  comme on l'a fait dans la démonstration du théorème de Vitali, on peut alors utiliser la théorème de recouvrement de Vitali et on a :

$$\frac{\omega_k}{2^k} \sum_{i \in \mathbb{N}} \text{diam}(C_i)^k \leq \frac{1}{1 - \delta} \sum_{i \in \mathbb{N}} \mu(C_i) \leq \frac{1}{1 - \delta} \mu(A) < +\infty$$

et donc nous sommes dans le second cas du théorème de Vitali :

$$\mathcal{H}^k(B) = \mathcal{H}^k\left(\bigcup_{i \in \mathbb{N}} C_i\right) \leq \frac{\omega_k}{2^k} \sum_{i \in \mathbb{N}} \text{diam}(C_i)^k \leq \frac{1}{1 - \delta} \mu(A)$$

Le résultat qui vient d'être montré est vrai pour tout  $\delta$  et pour tout  $A$  et donc  $\mathcal{H}^k(B) \leq \mu(B)$ . Pour obtenir le résultat plus général on peut remarquer que les raisonnements précédents restent vrais pour des ensembles  $B$  non bornés. Enfin pour montrer l'inégalité pour tous  $t$  et non plus seulement pour 1, on pose toujours le même ensemble  $\mathfrak{C}_\delta$  en demandant cette fois que

$$\mu(B(x, r)) \geq (1 - \delta) t \omega_k (\text{diam}(B(x, r)))^k / 2^k$$

et par les mêmes considérations on a :

$$\exists (r_i)_{i \in \mathbb{N}}, \forall i \in \mathbb{N}, \frac{\mu(B(x, r_i))}{\omega_k r_i^k} \geq t(1 - \delta)$$

et donc par la même suite d'arguments on obtient pour tout  $\delta$  et pour tout  $A$  choisit comme précédemment

$$\mathcal{H}^k(B) = \mathcal{H}^k\left(\bigcup_{i \in \mathbb{N}} C_i\right) \leq \frac{\omega_k}{2^k} \sum_{i \in \mathbb{N}} \text{diam}(C_i)^k \leq \frac{1}{t(1 - \delta)} \sum_{i \in \mathbb{N}} \mu(C_i) \leq \frac{1}{t(1 - \delta)} \mu(A) < +\infty$$

et finalement on a bien le résultat souhaité :  $\mu(B) \geq t \mathcal{H}^k(B)$ .

□

### 3.3 Les connexes de mesures finies sont connexes par arc

En réalité ce qui nous intéresse est de montrer que tout connexe de mesure de HAUSDORFF finie est presque entièrement recouvrable par une quantité dénombrable de courbes lipschitziennes. Ce résultat est en fait une conséquence du fait que le connexe considéré est connexe par arc.

**Théorème.** *Si  $(E, d_E)$  est un espace métrique séparable alors il existe une isométrie de  $E$  dans  $l^\infty$ ,  $\mathcal{J} : E \rightarrow l^\infty$*

*Démonstration.* Si on nomme  $(x_n)_{n \in \mathbb{N}}$  la famille dénombrable dense de  $E$  et si on pose

$$\forall x \in E, \mathcal{J}(x) = (d_E(x, x_n) - d_E(x_1, x_n))_{n \in \mathbb{N}}$$

On a déjà que  $\mathcal{J}(x) \in l^\infty$ , en effet par l'inégalité triangulaire on a :

$$\forall n \in \mathbb{N}, \forall x \in E, d_E(x, x_n) - d_E(x_1, x_n) \leq d_E(x, x_1)$$

La symétrie nous donne :  $|d_E(x, x_n) - d_E(x_1, x_n)| \leq d_E(x, x_1) < +\infty$ . Et donc :  $\forall x \in E, \mathcal{J}(x) \in l^\infty$ .

Pour montrer l'isométrie on procède ensuite par double inégalité :

$$* \|\mathcal{J}(x) - \mathcal{J}(y)\|_{l^\infty} \leq d_E(x, y) :$$

Soient  $x$  et  $y$  appartenant à  $E$  alors

$$\mathcal{J}(x) - \mathcal{J}(y) = (d_E(x, x_n) - d_E(y, x_n))_{n \in \mathbb{N}}$$

et par l'inégalité triangulaire on a

$$\forall n \in \mathbb{N}, |d_E(x, x_n) - d_E(y, x_n)| \leq d_E(x, y)$$

$$* \|\mathcal{J}(x) - \mathcal{J}(y)\|_{l^\infty} \geq d_E(x, y) :$$

$$\|\mathcal{J}(x) - \mathcal{J}(y)\|_{l^\infty} \geq |d_E(x, x_n) - d_E(x_1, x_n) - d_E(y, x_n) + d_E(x_1, x_n)| = |d_E(x, x_n) - d_E(y, x_n)|$$

et par suite d'inégalité triangulaire on a

$$\begin{aligned} |d_E(x, x_n) - d_E(y, x_n)| &= |d_E(x, x_n) - d_E(x, y) + d_E(x, y) - d_E(y, x_n)| \\ &\geq |d_E(y, x_n) - d_E(x, y) + d_E(x, y)| - |d_E(x, x_n)| \\ &\geq |d_E(x, y)| - |d_E(y, x_n) - d_E(x, y)| - |d_E(y, x_n)| \\ &\geq d_E(x, y) - 2\varepsilon \end{aligned}$$

□

**Théorème.** *Si  $(E, d_E)$  est un espace métrique complet et  $C$  une partie de  $E$  qui est fermée connexe telle que  $\mathcal{H}^1(C) < +\infty$  alors  $C$  est compact et est connexe par arc.*

*Démonstration.* On commence par montrer que  $C$  est compact. Pour cela on va montrer que  $C$  est toujours recouvert par un nombre fini de boules de rayons aussi petit que souhaité (propriété de Borel-Lebesgue) et comme  $C$  est fermé dans un espace complet alors  $C$  est lui même compact ce qui nous donne bien une caractérisation de la compacité. On se donne donc  $x_0 \in C$  et  $\varepsilon < \text{diam}(C)$  (si  $\text{diam}(C) = 0$  alors  $C = \{x_0\}$  et  $C$  est donc bien compact) et on construit

deux suites  $(x_n)_{n \in \mathbb{N}}$  et  $(B(x_n, \varepsilon))_{n \in \mathbb{N}}$  où  $x_n$  est choisit de manière à ce que  $x_n \in C \setminus \bigcup_{i=0}^{n-1} B(x_i, \varepsilon)$ .

On affirme de plus que par ce procédé on obtient un recouvrement fini de  $C$  par des boules de diamètre  $\varepsilon$  (autrement dit que la construction proposée s'arrête au bout d'un nombre fini d'étapes). On a en effet que  $(B(x_n, \varepsilon/2))_{n \in \mathbb{N}}$  est une famille de boules disjointes deux à deux (soient  $x_i$  et  $x_j$  deux éléments différents de la suite précédemment définie, on a que  $x_i \notin B(x_j, \varepsilon)$  donc  $d_E(x_i, x_j) \geq \varepsilon$  on en déduit que si  $y \in B(x_i, \varepsilon/2)$  alors  $d_E(y, x_j) \geq d_E(x_i, x_j) - d_E(y, x_i) \geq \varepsilon/2$  et donc  $y \notin B(x_j, \varepsilon/2)$ , de manière symétrique on en déduit qu'aucun élément de l'autre boule n'appartient à sa voisine et donc  $B(x_i, \varepsilon/2) \cap B(x_j, \varepsilon/2) = \emptyset$ ). Ainsi on a que  $(C \cap B(x_i, \varepsilon/2))_{i \in \mathbb{N}}$  est une famille d'ensembles disjointes deux à deux,

de plus  $\bigcup_{i=0}^{n-1} (C \cap B(x_i, \varepsilon/2)) \subset C$  donc on a :

$$\mathcal{H}^1(C) \geq \sum_{i=0}^{n-1} \mathcal{H}^1(C \cap B(x_i, \varepsilon/2)) \geq n\varepsilon/2$$

La dernière inégalité venant toujours du lemme de sortie du territoire. On a ainsi une minoration de la mesure de HAUSDORFF de notre connexe, et comme la mesure de notre ensemble est finie alors nécessairement la valeur de  $n$  est finie, ce qui nous donne le nombre d'étapes nécessaire pour construire notre recouvrement.

La fin de la démonstration va consister à montrer que  $C$  peut être paramétré par une courbe. Pour cela on va construire plus ou moins explicitement une courbe lipschitzienne qui va approcher à  $\varepsilon$  près notre connexe  $C$  puis par passage à la limite montrer qu'il existe une courbe  $\gamma$  qui répond à notre problème. On commence par faire remarquer que quitte à réduire la surface de  $E$  on peut estimer que  $E$  est compact et donc séparable, par conséquent quitte à composer par une isométrie on peut estimer que  $E$  a une structure d'espace vectoriel (celui de  $l^\infty$ ).

Pour  $\varepsilon > 0$  et  $x \in C$  on définit l'ensemble  $C'_x = \{y \in C, y \text{ et } x \text{ sont } \varepsilon\text{-connectés dans } C\}$ . Et on dit que deux points  $y$  et  $x$  sont  $\varepsilon$ -connectés dans  $C$  lorsqu'il existe un nombre fini d'éléments  $(x_i)_{0 \leq i \leq n}$  tels que :

\*  $x_0 = x$  et  $x_n = y$

\* Pour tout  $i$  dans on a  $x_i \in C$

\* Pour tout  $i$  dans on a  $d_E(x_i, x_{i+1}) \leq \varepsilon$ .

Il est clair que si  $C$  est connexe et compact alors  $C'_x$  est fermé. En effet si on se donne une suite convergente de  $C'_x$  alors elle converge dans  $C$  vers  $z$ , il suffit ensuite de montrer que  $z$  est  $\varepsilon$ -connecté avec  $x$ , ce qui est à peu près évident car à partir d'un certain rang les éléments de la suite sont proches à moins de  $\varepsilon$  donc ces éléments sont  $\varepsilon$ -connectés avec  $z$  et comme être  $\varepsilon$ -connecté est une opération qui est transitive on a ce qu'on souhaite.

On a de plus que  $C'_x$  est ouvert (de  $C$ ), en effet il suffit de prendre la boule  $B(x, \varepsilon/2)$  car tout élément de  $C \cap B(x, \varepsilon/2)$  appartient à  $C$  et est à moins de  $\varepsilon$  de  $x$  ainsi  $C \cap B(x, \varepsilon/2) \subset C'_x$ .

De ces deux résultats on en déduit que  $C'_x = C$  par connexité de  $C$ . Ainsi pour tout couple  $(x, y) \in C^2$  on peut se donner une suite finie d'éléments  $(x_i)_{0 \leq i \leq n}$  qui relie  $x$  et  $y$  par "saut" de taille  $\varepsilon$ . On peut aussi supposer que  $|i - j| > 1 \Rightarrow d_E(x_i, x_j) \geq \varepsilon$ . Cette dernière supposition est évidente, si jamais elle n'est pas vérifiée il suffit de retirer l'élément (ou les éléments) situé entre  $i$  et  $j$ . Et donc on a  $|i - j| > 1 \Rightarrow B(x_i, \varepsilon/2) \cap B(x_j, \varepsilon/2) = \emptyset$ . Et donc on a :

$$\mathcal{H}^1(C) \geq \mathcal{H}^1(C \cap \bigcup_{0 \leq i \leq n, i=\text{impair}} B(x_i, \varepsilon/2)) = \sum_{0 \leq i \leq n, i=\text{impair}} \mathcal{H}^1(C \cap B(x_i, \varepsilon/2))$$

Et par le lemme de sortie du territoire on a :  $\mathcal{H}^1(C) \geq \sum_{0 \leq i \leq n, i=\text{impair}} \frac{\varepsilon}{2}$ .

Et par les mêmes considérations paires on obtient finalement :  $2\mathcal{H}^1(C) \geq n\frac{\varepsilon}{2}$ . Tout ceci



nous permet d'avoir une majoration de la taille de la suite nécessaire pour relier  $x$  et  $y$  par des  $\varepsilon$ -chemins.

On se donne  $(x, y) \in C^2$  et  $(x_i)_{0 \leq i \leq n}$  un  $\varepsilon$ -chemin entre  $x$  et  $y$ . On sait déjà que  $n \leq 4\mathcal{H}^1(C)/\varepsilon$ , on pose  $L = 4\mathcal{H}^1(C)$  et on construit  $\gamma_\varepsilon : [0, L] \rightarrow E$  de la manière suivante :

- \* On se donne une subdivision  $0 = t_0 < t_1 < \dots < t_n \leq L$  telle que :
- $\forall i \leq n-1, t_{i+1} - t_i = d_E(x_i, x_{i+1}) < \varepsilon$ ;
- \* On pose alors sur chaque intervalle  $[t_i, t_{i+1}]$ ,  $\gamma_\varepsilon(s) = x_i + \frac{(s-t_i)}{(t_{i+1}-t_i)}(x_{i+1} - x_i)$ ;
- \* Et sur l'intervalle  $[t_n, L]$ ,  $\gamma_\varepsilon(s) = x_n = y$

On a alors que la courbe  $\gamma_\varepsilon$  vérifie bien évidemment les propriétés suivantes (puisqu'elle a été construite pour ça) :

- \*  $\gamma_\varepsilon$  est 1-lipschitzienne;
  - \*  $\forall s \in [0, L], \gamma_\varepsilon(s) \in C_\varepsilon = \{z \in E, d_E(z, C) \leq \varepsilon\}$ ;
- Pour la lipschitzianité il suffit d'écrire  $|\gamma_\varepsilon(s) - \gamma_\varepsilon(t)|$  de simplifier et de s'apercevoir que c'est bien inférieur à  $|s - t|$ . Quand au fait que la courbe est incluse dans une  $\varepsilon$ -enveloppe autour de  $C$ . On a pour tout  $t \in [t_i, t_{i+1}]$  que  $|x_i - \gamma_\varepsilon(t)| = \left| \frac{(t-t_i)}{(t_{i+1}-t_i)}(x_{i+1} - x_i) \right| \leq \frac{(t-t_i)}{(t_{i+1}-t_i)}\varepsilon \leq \varepsilon$ .

On en déduit que la famille  $(\gamma_\varepsilon)_{10^{-n} > \varepsilon > 0}$  est équicontinue et  $\Gamma_\varepsilon(t) = \{\gamma_\varepsilon(t), 10^{-n} > \varepsilon > 0\}$  est borné pour tout  $t \in [0, L]$  donc relativement-compact. On peut donc appliquer le théorème de Ascoli-Arzelà et on en déduit l'existence d'une courbe  $\gamma$  qui est 1-lipschitzienne et qui est limite uniforme de  $(\gamma_{\phi(n)})_{n \in \mathbb{N}}$ . Et pour tout  $t \in [0, L], \gamma(t) \in C_0 = C$ .  $\square$

**Théorème.** Si  $(E, d_E)$  est un espace métrique complet et  $C$  un fermé-connexe de  $E$  tel que  $\mathcal{H}^1(C) < +\infty$  alors il existe une famille dénombrable de courbes lipschitziennes  $(\gamma_i)_{i \in \mathbb{N}}$  de  $[0, 1]$  dans  $C$  telle que :

$$\mathcal{H}^1\left(C \setminus \bigcup_{i \geq 0} \gamma_i([0, 1])\right) = 0$$

*Démonstration.* Une fois de plus la démonstration va être constructive. Elle va consister à recouvrir successivement notre connexe par des courbes paramétrées, pour cela on va abondamment utiliser le théorème précédent qui va nous permettre de relier deux points de  $C$  par un arc. On va d'abord choisir les deux points les plus éloignés et considérer la courbe  $\Gamma_1$  qui les relie, ensuite on choisit un point qui n'appartient pas à cette courbe et qui se situe le plus loin (au sens de la distance  $d_E$ ) de la courbe, on relie ensuite ce point à  $\Gamma_1$  par une courbe  $\Gamma_2$  et on continue ainsi pour tout  $k \in \mathbb{N}$ . On formalise tout cela de la manière suivante :

On se donne  $(x, y) \in E^2$  tel que  $d_E(x, y) = \text{diam}(C)$  (car  $C$  est compact d'après le théorème précédent). Il existe  $\Gamma_0 : [0, 1] \rightarrow C$  tel que  $\Gamma_0(0) = x$  et  $\Gamma_0(1) = y$  et pour tout  $t \in [0, 1], \Gamma_0(t) \in C$ . On construit ainsi  $(\Gamma_i)_{0 \leq i \leq k}$  tel que :

- \*  $\Gamma_i([0, 1]) \subset C$  pour tout  $i$  compris entre 0 et  $k$ ;
- \*  $\Gamma_i([0, 1]) \cap \bigcup_{j < i} \Gamma_j([0, 1])$  est réduit à un point pour tout  $i$  compris entre 0 et  $k$ ;

On définit  $d_k = \sup_{x \in C} (d_E(x, \bigcup_{i=0}^k \Gamma_i([0, 1])))$ .

- \* Si  $d_k = 0$  alors on a  $\bigcup_{i=0}^k \Gamma_i([0, 1]) = C$  et le théorème est prouvé;

\* Si  $d_k > 0$  alors on pose  $x_{k+1} \in C$  et  $y_{k+1} \in \bigcup_{i=0}^k \Gamma_i([0, 1])$  tel que  $d_k = d_E(x_{k+1}, y_{k+1})$ , et on note  $\Gamma'_{k+1}$  la courbe lipschitzienne qui relie  $x_{k+1}$  et  $y_{k+1}$ . A partir de cette courbe on va construire

une nouvelle courbe  $\Gamma_{k+1}$  qui va relier  $x_{k+1}$  à  $\bigcup_{i=0}^k \Gamma_i([0, 1])$ , pour cela on pose :

$t_{k+1} = \inf\{t \in [0, 1], \Gamma'_{k+1}(t) \in \bigcup_{i=0}^k \Gamma_i([0, 1])\}$  et on appelle  $\Gamma_{k+1} : [0, t_{k+1}] \rightarrow C$  la courbe lipschitzienne recherchée qui pour tout  $t \in [0, t_{k+1}]$  vérifie  $\Gamma_{k+1}(t) = \Gamma'_{k+1}(t)$ . On peut ensuite reparamétriser cette nouvelle courbe pour qu'elle soit définie sur  $[0, 1]$ , et on a bien que  $\Gamma_{k+1}([0, 1]) \cap \bigcup_{j \leq k} \Gamma_j([0, 1])$  est réduit à un point (le point  $\Gamma_{k+1}(1)$ ).

Il nous reste à montrer que notre construction recouvre presque entièrement  $C$ . Pour cela on va commencer par montrer que  $C \subset \overline{\bigcup_{i=0}^{\infty} \Gamma_i([0, 1])}$  et enfin par le théorème de HAUSDORFF-RADON

on va pouvoir majorer la quantité  $\mathcal{H}^1(C \setminus \bigcup_{i=0}^{\infty} \Gamma_i)$  par quelque chose d'aussi petit que souhaité.

On a déjà :  $d_k \leq d_E(x_{k+1}, \Gamma_{k+1}(1))$  car  $d_k$  est la longueur de  $x_{k+1}$  à l'union des courbes précédemment construites et que  $\Gamma_{k+1}(1)$  appartient à cette même union. De plus on a :

$$d_E(x_{k+1}, \Gamma_{k+1}(1)) \leq \mathcal{H}^1(\Gamma_{k+1}([0, 1])),$$

en effet si on pose  $\phi_{x_{k+1}}(z) = d_E(x_{k+1}, z)$  alors :  $\phi_{x_{k+1}}(\Gamma_{k+1}([0, 1])) = [0, p]$  et  $p \geq d_E(x_{k+1}, \Gamma_{k+1}(1))$  et comme  $\phi_{x_{k+1}}$  est 1-lipschitzienne on a :

$$d_E(x_{k+1}, \Gamma_{k+1}(1)) \leq p = \mathcal{H}^1(\phi_{x_{k+1}}(\Gamma_{k+1}([0, 1]))) \leq \mathcal{H}^1(\Gamma_{k+1}([0, 1]))$$

On trouve ainsi :  $d_k \leq \mathcal{H}^1(\Gamma_{k+1}([0, 1]))$ . On a donc ensuite :

$$\sum_{i=0}^{k+1} d_i \leq \sum_{i=0}^{k+2} \mathcal{H}^1(\Gamma_i([0, 1])) \leq \mathcal{H}^1(C)$$

Ce qui nous donne que la série de terme générale  $(d_i)_{i \in \mathbb{N}}$  est convergente et donc que la suite tend vers 0 lorsque  $n$  tend vers l'infini. Si on pose :  $(B)_\varepsilon = \{x \in E, d_E(x, B) \leq \varepsilon\}$ , on a alors par construction :

$$C \subset \left(\bigcup_{j \leq k} \Gamma_j([0, 1])\right)_{2d_k} \subset \left(\bigcup_{j \geq 0} \Gamma_j([0, 1])\right)_{2d_k}$$

Ceci est vrai pour tout  $k \in \mathbb{N}$  donc on a :  $C \subset \overline{\bigcup_{j \geq 0} \Gamma_j([0, 1])}$

Posons  $\mu(B) = \mathcal{H}^1(B \cap \bigcup_{i=k+1}^{\infty} \Gamma_i([0, 1]))$ , si on avait  $\overline{\Theta}_1(\mu, x) \geq 1/2, \forall x \in C \setminus \bigcup_{i=0}^k \Gamma_i([0, 1])$  alors par application du théorème de HAUSDORFF-RADON on obtiendrait :

$$\mathcal{H}^1\left(\bigcup_{i=k+1}^{\infty} \Gamma_i([0, 1])\right) \geq \mu\left(C \setminus \bigcup_{i=0}^k \Gamma_i([0, 1])\right) \geq \frac{1}{2} \mathcal{H}^1\left(C \setminus \bigcup_{i=0}^k \Gamma_i([0, 1])\right)$$

Ainsi

$$\mathcal{H}^1\left(C \setminus \bigcup_{i=0}^{\infty} \Gamma_i([0, 1])\right) \leq \mathcal{H}^1\left(C \setminus \bigcup_{i=0}^k \Gamma_i([0, 1])\right) \leq 2\mathcal{H}^1\left(\bigcup_{i=k+1}^{\infty} \Gamma_i([0, 1])\right) \leq 2 \sum_{i=k+1}^{\infty} \mathcal{H}^1(\Gamma_i([0, 1]))$$

Or d'après ce qui précède on a que la série de terme général  $(\mathcal{H}^1(\Gamma_i([0, 1]))_{i \in \mathbb{N}}$  est convergente et donc pour  $k$  assez grand le dernier terme de l'inégalité est aussi petit qu'on le souhaite. Par passage à la limite on obtient que :  $\mathcal{H}^1(C \setminus \bigcup_{i=0}^{\infty} \Gamma_i([0, 1])) = 0$  ce qui nous donnerait le résultat attendu.

L'objectif est donc de minorer la densité 1-dimensionnelle de la mesure  $\mu$ . Pour cela on va montrer que l'on a (pour un  $r$  assez petit et  $x \in C \setminus \bigcup_{i=0}^k \Gamma_i([0, 1])$ ),  $\mathcal{H}^1(B(x, r) \cap \bigcup_{i=k+1}^{\infty} \Gamma_i([0, 1])) \geq r$  (en divisant par  $r$  et en passant à la limite supérieure on obtient immédiatement le résultat recherché sur la densité 1-dimensionnelle de la mesure  $\mu$ ). On commence par donner une vision intuitive de ce résultat : si on se donne  $x \in C \setminus \bigcup_{i=0}^k \Gamma_i([0, 1])$  alors  $x$  va appartenir à l'une des courbes  $(\Gamma_i([0, 1]))_{i > k}$  ( $\Gamma_p([0, 1])$  par exemple) et alors on aura  $r \leq d_E(x, \Gamma_p(1))$  (pour  $r$  assez petit), on sait alors que la courbe est obligée de sortir de la boule et donc par un raisonnement analogue à celui fait pour la démonstration du lemme de sortie du territoire on obtient :  $\mathcal{H}^1(B(x, r) \cap \bigcup_{i=k+1}^{\infty} \Gamma_i([0, 1])) \geq \mathcal{H}^1(B(x, r) \cap \Gamma_p([0, 1])) \geq r$ . Cependant  $x$  peut se trouver dans l'adhérence de l'une des courbes et non exactement appartenir à celle-ci.

On se donne  $x \in C \setminus \bigcup_{i=0}^k \Gamma_i([0, 1])$  et  $r$  tel que  $\overline{B(x, r)} \cap \bigcup_{i=0}^k \Gamma_i([0, 1]) = \emptyset$ . On sait qu'il existe une suite  $(x_h)_{h \in \mathbb{N}} \in C^{\mathbb{N}}$  qui converge vers  $x$  et telle que pour tout entier naturel  $h$  on ait  $x_h \in \Gamma_{\phi(h)}$ , où  $\phi(h) > k$  et  $\phi$  n'est pas décroissante. On pose ensuite  $C_h = \bigcup_{i=0}^{\phi(h)} \Gamma_i([0, 1])$  ( $C_h$  est évidemment connexe car il est connexe par arc). Pour tout  $r' < r$  et  $h$  assez grand on a que  $C_h$  sort de la boule  $B(x_h, r')$  (sinon on aurait que le compact  $C$  est contenu dans une boule de rayon  $r$  et donc on ne pourrait avoir  $\overline{B(x, r)} \cap \bigcup_{i=0}^k \Gamma_i([0, 1]) = \emptyset$ ). Ainsi on a :  $\mathcal{H}^1(C_h \cap B(x_h, r')) \geq r'$ . On sait aussi que pour  $h$  assez grand on a :  $B(x_h, r') \subset B(x, r)$  (car la suite converge vers  $x$ ). On a alors la suite d'inégalité suivante :

$$\mathcal{H}^1(B(x, r) \cap \bigcup_{i=k+1}^{\phi(h)} \Gamma_i([0, 1])) = \mathcal{H}^1(B(x, r) \cap C_h) \geq \mathcal{H}^1(B(x_h, r') \cap C_h) \geq r'$$

La première égalité vient du fait que la boule  $B(x, r)$  ne coupe pas les  $k$  premières courbes, les autres inégalités sont des conséquences immédiates de tout ce qui a été dit. Le résultat prouvé est vrai pour tout  $r' < r$  et donc par passage à la limite on a que :

$$\mathcal{H}^1(B(x, r) \cap \bigcup_{i=k+1}^{\phi(h)} \Gamma_i([0, 1])) \geq r,$$

ce qui nous donne le résultat tant espéré. □

### 3.4 Théorème de GOLAB

**Théorème (GOLAB).** *Soit  $(E, d_E)$  un espace métrique compact et  $(C_n)_{n \in \mathbb{N}}$  une famille de fermés connexes qui converge uniformément vers  $C$  alors on a que  $C$  est connexe(déjà vu) et :*

$$\mathcal{H}^1(C) \leq \liminf_{n \rightarrow \infty} \mathcal{H}^1(C_n)$$

*Démonstration.* On commence par faire cette suite d'hypothèses :

\* Quitte à extraire une sous suite on peut estimer que  $\lim_{n \rightarrow \infty} \mathcal{H}^1(C_n) = L$  existe et est finie.

\* Quitte à supprimer des termes de la suite on peut supposer que :  $\forall n \in \mathbb{N}, \mathcal{H}^1(C_n) < L + 1$ .

\* Si  $d_n = \text{diam}(C_n)$  alors quitte à extraire une sous suite on peut supposer que  $(d_n)_{n \in \mathbb{N}}$  est une suite qui converge vers  $d$ , on montre de plus que nécessairement  $d = \text{diam}(C)$  : en effet si  $(x, y) \in C^2$  tels que  $d_E(x, y) = \text{diam}(C)$  alors il existe  $(x_n)_{n \in \mathbb{N}}$  et  $(y_n)_{n \in \mathbb{N}}$  qui convergent respectivement vers  $x$  et  $y$  et telles que pour tout entier naturel  $n$   $(x_n, y_n) \in C_n^2$ , alors :  $d_E(x_n, y_n) \leq d_n$  et par passage à la limite  $d_E(x, y) \leq d$ . De manière analogue on a qu'il existe  $(x_n)_{n \in \mathbb{N}}$  et  $(y_n)_{n \in \mathbb{N}}$  telles que pour tout entier naturel  $n$   $(x_n, y_n) \in C_n^2$  et  $d_E(x_n, y_n) = d_n$  de plus quitte à extraire des sous suites on peut estimer que chacune des suites convergent vers  $x$  et  $y$  qui appartiennent à  $C$  et ainsi :  $d = d_E(x, y) \leq \text{diam}(C)$ . On obtient donc par double inégalités que  $d = \text{diam}(C)$  (On remarque que cette égalité est en fait du au fait que la suite de compact converge simplement vers  $C$ ).

\* Si  $\mu_n(B) = \mathcal{H}^1(B \cap C_n)$  alors quitte à extraire une sous suite on peut supposer que  $(\mu_n)_{n \in \mathbb{N}}$  est une suite de mesures finies qui converge faiblement vers  $\mu$  (pour avoir plus de précisions sur le fait que l'on puisse effectivement faire ceci regarder la première partie de l'annexe d'analyse). Le principe de la preuve est de majorer la valeur  $\mathcal{H}^1(C)$  par la valeur  $\mu(E)$ . La mesure  $\mu$  est plus agréable pour plusieurs raisons : la première étant que  $\mu$  est une limite de mesure et possède ainsi plusieurs propriétés d'inégalités (avec les ouverts et les fermés), la seconde est que les éléments de la suite qui définit  $\mu$  ont eux même des propriétés très intéressantes (notamment celles liées au lemme de sortie du territoire), la dernière est que le théorème de HAUSDORFF-RADON donne un lien direct entre les deux mesures sous des conditions qui vont être vérifiées grâce aux propriétés précédentes. La preuve du théorème de GOLAB va donc reposer principalement sur le théorème de HAUSDORFF-RADON et, comme on va le voir, la première partie va montrer qu'on peut facilement minorer la densité 1-dimensionnelle supérieure de la mesure  $\mu$  par un facteur 1/2, tandis que dans la seconde partie on va travailler pour minorer cette même quantité par un facteur 1. Comme indiqué on veut utiliser le théorème de HAUSDORFF-RADON, on va donc commencer par essayer de minorer la quantité  $\mu(B(x, r))$  pour  $x$  dans  $C$ . Soient  $x \in C$  et  $r' < r < \text{diam}(C)/2$  alors on a :

$$\mu(B(x, r)) \geq \mu(\overline{B(x, r')}) \geq \overline{\lim}_{n \rightarrow \infty} \mu_n(\overline{B(x, r')}) = \overline{\lim}_{n \rightarrow \infty} \mathcal{H}^1(C_n \cap \overline{B(x, r')}) \geq r'$$

La dernière inégalité vient du "lemme de sortie du territoire" (il est à remarquer que l'on a bien  $r' < \text{diam}(C_n)/2$  à partir d'un certain rang car d'après les hypothèses on a convergence des diamètres de la suite vers le diamètre de  $C$ ). On vient de montrer que  $\mu(B(x, r)) \geq r'$  pour tout  $r' < r$ , par passage à la limite on obtient finalement :  $\forall x \in C, \forall 0 < r < \text{diam}(C)/2, \mu(B(x, r)) \geq r$ . Et on a  $\omega_1 = 2$ . On a donc le résultat :

$$\forall x \in C, \frac{\mu(B(x, r))}{r} \geq 1 \Rightarrow \forall x \in C, \frac{\mu(B(x, r))}{\omega_1 r} \geq 1/2 \Rightarrow \forall x \in C, \overline{\Theta}_1(\mu, x) \geq 1/2$$

Et par le théorème déjà trop de fois nommé on a :

$$\frac{1}{2} \mathcal{H}^1(C) \leq \mu(C) \leq \mu(E) \leq \overline{\lim}_{n \rightarrow \infty} \mu_n(E) = \overline{\lim}_{n \rightarrow \infty} \mathcal{H}^1(C_n) = L$$

On a donc que la "taille" de  $C$  par la mesure de HAUSDORFF est bien finie et est majorée par deux fois la taille limite de la suite de compacts connexes.

Dans la seconde partie on va utiliser des propriétés de recouvrements de l'ensemble  $C$  par des courbes paramétrées (car la mesure de HAUSDORFF de  $C$  est finie). On sait que pour presque

tout  $x_0 \in C$  on a l'existence d'une courbe paramétrées  $k$ -lipschitzienne de  $[0, 1]$  dans  $C$  telle que  $x_0 = \gamma(t_0)$ . De plus quitte à reparamétriser on peut supposer que  $|\dot{\gamma}|(t_0) = 1$ . On a de surcroît une courbe qui est  $k$ -lipschitzienne et presque partout différentiable (la seconde partie de l'annexe d'analyse justifie cette affirmation). On peut ainsi supposer que la quantité suivante à bien un sens

$$\lim_{h>0, h \rightarrow 0} \frac{d_E(\gamma(t_0 + h), \gamma(t_0 - h))}{2|h|} = 1$$

On se donne  $x_0 \in C$  et  $\gamma$  qui vérifient les conditions précédentes ( $x_0$  ne fait donc pas partie de cet ensemble négligeable de points qui ne suit pas les propriétés qui ont été énoncées). On se donne de plus un élément  $\sigma \in ]0, 1[$ . Par définition de la limite on a l'existence d'une valeur  $h_1$  telle que :

$$\forall 0 < h < h_1, d_E(\gamma(t_0 + h), \gamma(t_0 - h)) \geq (2 - \sigma)h$$

De la même manière (et parce que  $|\dot{\gamma}|(t_0) = 1$ ) on a l'existence d'une valeur  $h_2$  telle que :

$$\forall 0 < h < h_2, (1 - \sigma)h \leq d_E(\gamma(t_0 + h), \gamma(t_0)) \leq (1 + \sigma)h$$

On peut ainsi trouver une valeur  $\eta < \sigma/(1 + \sigma)$  et qui vérifie les deux approximations précédentes (il suffit de prendre  $\eta < \min(\sigma/(1 + \sigma), h_1, h_2)$ ). On pose ensuite  $y = \gamma(t_0 - \eta)$  et  $z = \gamma(t_0 + \eta)$  et  $r = \max(d_E(y, x_0), d_E(z, x_0))$ . On obtient les résultats :

$$r = d_E(\gamma(t_0 + \eta), \gamma(t_0)) \leq (1 + \sigma)\eta < (1 + \sigma)\sigma/(1 + \sigma) < \sigma$$

$$d_E(\gamma(t_0 + \eta), \gamma(t_0 - \eta)) = d_E(y, z) \geq (2 - \sigma)\eta \geq \frac{(2 - \sigma)}{(1 + \sigma)}r$$

La fin du travail va consister à minorer la densité 1-dimensionnelle supérieure de la mesure  $\mu$  par la quantité  $\frac{(2 - \sigma)}{(1 + \sigma)\omega_1}$ . Il suffit pour cela de minorer  $\mu(B(x_0, r))$  par  $\frac{(2 - \sigma)}{(1 + \sigma)}r$  ou encore de minorer  $\mu(B(x_0, r))$  par  $d_E(y, z)$ . Pour cela on se donne  $r' = (1 + \sigma)r$  qui va nous donner une marge de "sécurité" autour de  $B(x_0, r)$ . On rappelle alors la première propriété de cet article (oui c'est déjà loin tout ça : une histoire bizarre de convergence pour des topologies curieuses) qui disait que la convergence de HAUSDORFF impliquait celle de KURATOWSKI, les points  $y$  et  $z$  sont des éléments de  $C$  (la limite hausdorffienne de notre suite de connexe) et donc il existe deux suites  $(y_n)_{n \in \mathbb{N}}$  et  $(z_n)_{n \in \mathbb{N}}$  qui convergent respectivement vers  $y$  et  $z$  et telles que :  $\forall n \in \mathbb{N}, y_n \in C_n, z_n \in C_n$ . Ainsi à partir d'un certain rang on a que  $y_n \in B(x_0, r')$  et  $z_n \in B(x_0, r')$ . On pose de plus  $\phi_z(x) = d_E(z, x)$  qui est une application 1-lipschitzienne ; on a alors :

$$\mu_n(\overline{B(x_0, r')}) = \mathcal{H}^1(C_n \cap \overline{B(x_0, r')}) \geq \mathcal{H}^1(\phi_z(C_n \cap \overline{B(x_0, r')})) = \sup_{w \in C_n \cap \overline{B(x_0, r')}} \phi_z(w) \geq d_E(z, y_n)$$

Pour un  $n$  assez grand. On peut ainsi prendre ensuite la limite supérieure sur cette inégalité et on obtient :

$$\mu(\overline{B(x_0, r')}) \geq \overline{\lim}_{n \rightarrow \infty} \mathcal{H}^1(C_n \cap \overline{B(x_0, r')}) \geq \overline{\lim}_{n \rightarrow \infty} d_E(z, y_n) = d_E(z, y) \geq \frac{(2 - \sigma)}{(1 + \sigma)}r \geq \frac{(2 - \sigma)}{(1 + \sigma)^2}r'$$

Ceci est vrai pour tout  $\sigma \in ]0, 1]$  on peut donc faire tendre  $\sigma$  vers 0 et on obtient finalement  $\mu(B(x_0, r)) \geq 2r$  ce qui nous donne bien  $\overline{\Theta}_1(\mu, x_0) \geq 1$  pour  $\mathcal{H}^1$ -presque tout  $x_0 \in C$ . On peut poser  $C' \subset C$  tel que :

$$\forall x \in C', \overline{\Theta}_1(\mu, x) \geq 1$$

On a de plus que  $\mu(C) \geq \mu(C')$  car  $C' \subset C$  et  $\mathcal{H}^1(C) = \mathcal{H}^1(C')$  car  $C'$  est défini à un ensemble de mesure nulle (pour la mesure de HAUSDORFF) près. On a donc finalement :

$$\mu(C) \geq \mu(C') \geq \mathcal{H}^1(C') = \mathcal{H}^1(C)$$

Et par les mêmes raisonnements que précédemment on a :

$$\mathcal{H}^1(C) \leq \mu(C) \leq \varliminf_{n \rightarrow \infty} \mu_n(E) \leq \varliminf_{n \rightarrow \infty} \mathcal{H}^1(C_n) = L$$

□

## 4 Applications

**Définition.** Soit  $f : E \rightarrow \mathbb{R}$ . On dit que  $f$  est semi-continue si et seulement si pour toute suite  $(x_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$  qui converge vers une valeur  $x$  on a que la valeur de l'image de  $x$  par  $f$  est inférieure à la limite inférieure de la suite  $(f(x_n))_{n \in \mathbb{N}}$  :

$$\forall (x_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}, x_n \rightarrow x \Rightarrow f(x) \leq \varliminf_{n \rightarrow \infty} f(x_n)$$

On peut de plus définir la semi-continuité supérieure et obtenir la propriété suivante : une fonction semi-continue supérieurement et inférieurement est continue.

Le théorème de GOLAB peut donc se ré-énoncé : La fonction “mesure de HAUSDORFF” allant de l'ensemble des compacts connexes muni de la distance de HAUSDORFF vers  $\mathbb{R}$  est semi-continue.

**Théorème.** Si  $f$  une fonction continue (qui représentera par exemple la densité de population) alors  $\min\{\int_E d_E(x, \Sigma) f(x) dx, \Sigma \in (K_E, d_H), \Sigma \text{ connexe } \mathcal{H}^1(\Sigma) \leq L\}$  existe.

**Théorème.** Si on se donne  $(x_i)_{0 \leq i \leq n}$  alors  $\min\{\mathcal{H}^1(\Sigma), \Sigma \in (K_E, d_H), \Sigma \text{ connexe}, \{x_0, \dots, x_n\} \subset \Sigma\}$  existe.

La démonstration de ces deux théorèmes repose notamment sur le fait que la mesure de HAUSDORFF est semi-continue (et donc sur le théorème de GOLAB). On remarque facilement que les problèmes de maximisation de réseaux sont de savoir par où faire passer la ligne (de train, à haute tension, ...) en fonction de la densité de la population, pour que le plus grand nombre puisse profiter au maximum de ces aménagements. Les théorèmes précédents nous disent qu'il existe un chemin optimal.

## 5 Annexe et théorème d'analyse

Cette partie d'analyse n'est pas nécessaire pour la théorie mais est indispensable pour des raisons techniques : certaines affirmations (qui apparaissent dans la démonstration du théorème

de GOLAB) n'ont pas été justifiées. Les arguments de ces affirmations sont à la limite entre l'analyse et la théorie de la mesure, mais surtout font intervenir des résultats complexes qu'il était délicat de développer entièrement au sein de ce projet, voici néanmoins quelques justifications.

## 5.1 A propos de la convergence faible d'une suite de mesures (finies) bornées

**Théorème (BANACH-ALAOGLU).** *Si  $X$  est un espace de Banach séparable et  $(\xi_n)_{n \in \mathbb{N}}$  une famille de formes linéaires équibornée (i.e :  $\exists C \in \mathbb{R}_+^*, \forall n \in \mathbb{N}$ ,*

$$|\xi_n| \leq C).$$

*Alors il existe une sous suite extraite qui converge simplement vers  $\xi \in X^*$ .*

*Démonstration.* On se donne  $D = \{x_i, i \in \mathbb{N}\}$  une partie dense dénombrable de  $X$ .

Pour tout entier naturel  $i$  on a que  $(\xi_n(x_i))_{n \in \mathbb{N}}$  est une suite bornée de  $\mathbb{R}$  (En effet comme la suite  $(\xi_n)_{n \in \mathbb{N}}$  est équibornée donc la suite considérée est bornée par  $C\|x_i\|_X$ ).

Ainsi il existe une extraction  $\phi_1$  telle que  $(\xi_{\phi_1(n)}(x_1))_{n \in \mathbb{N}}$  converge vers  $y_1 = \xi(x_1)$ .

On a alors que pour tout entier naturel  $i$   $(\xi_{\phi_1(n)}(x_i))_{n \in \mathbb{N}}$  est une suite bornée. Ainsi il existe une extraction  $\phi_2$  telle que  $(\xi_{\phi_1 \circ \phi_2(n)}(x_2))_{n \in \mathbb{N}}$  converge vers  $y_2 = \xi(x_2)$ .

On a alors que pour tout entier naturel  $i$   $(\xi_{\phi_1 \circ \phi_2(n)}(x_i))_{n \in \mathbb{N}}$  est une suite bornée.

On continue ainsi à construire successivement les extractions  $\phi_i$  pour tout entier naturel  $i$ .

On pose enfin l'extraction diagonale  $\phi$  telle que  $\forall i \in \mathbb{N}, \phi(i) = \phi_1 \circ \dots \circ \phi_i(i)$ . Et ainsi pour tout entier naturel  $i$   $(\xi_{\phi(n)}(x_i))_{n \in \mathbb{N}}$  est une suite convergente dont la limite est  $y_i = \xi(x_i)$ . La fonction  $\xi$  ainsi définie sur  $D$  est une fonction C-lipschitzienne comme limite de fonctions C-lipschitzienne. On peut ensuite prolonger  $\xi$  sur  $X$  en une fonction C-lipschitzienne. En effet, soit  $x \in X$  et  $(x_i)_{i \in \mathbb{N}} \in D^{\mathbb{N}}$  telle que  $\lim_{n \rightarrow \infty} x_n = x$  alors  $(\xi(x_i))_{i \in \mathbb{N}}$  est une suite de Cauchy, elle converge donc dans  $\mathbb{R}$ .

En effet :  $\forall \varepsilon \geq 0, \exists N_\varepsilon \in \mathbb{N}, \forall p \geq q \geq N_\varepsilon, |x_p - x_q| \leq \varepsilon$

donc :  $\forall \varepsilon \geq 0, \exists N_\varepsilon \in \mathbb{N}, \forall p \geq q \geq N_\varepsilon, |\xi(x_p) - \xi(x_q)| \leq C|x_p - x_q| \leq C\varepsilon$ .

On pose  $\lim_{n \rightarrow \infty} \xi(x_n) = \xi(x)$ . On montre tout de suite que cette définition est bien consistante puisqu'elle ne dépend pas de la suite choisie. Si on choisit deux suites de  $D$  qui convergent vers  $x$  alors les deux suites images (par la fonction  $\xi$ ) sont convergentes et par lipschitzianité de la fonction  $\xi$  sur  $D$  on a que les deux suites images sont aussi proches que souhaité pour un rang assez grand ; elles convergent donc vers la même limite  $\xi(x)$ .

On montre ensuite que  $\xi$  reste bien C-lipschitzienne sur  $X$ . Par les mêmes considérations discrètes et par passage à la limite on a naturellement le caractère lipschitzien qui est conservé. On veut montrer maintenant que  $\lim_{n \rightarrow \infty} \xi_{\phi(n)}(x) = \xi(x)$  pour tout  $x$  dans  $X$ .

$$|\xi_{\phi(n)}(x) - \xi(x)| \leq |\xi_{\phi(n)} - \xi_{\phi(n)}(x_i)| + |\xi_{\phi(n)}(x_i) - \xi(x_i)| + |\xi(x_i) - \xi(x)| \leq C\varepsilon + \varepsilon + C\varepsilon,$$

pour  $n$  assez grand et pour  $x_i$  bien choisit.

Il ne reste qu'à montrer que  $\xi$  est une forme linéaire, résultat qui se montre rapidement par linéarité de la limite.  $\square$

**Théorème (de représentation de RIESZ).** *Soit  $(X, d_X)$  un espace métrique compact et soit  $\mathcal{C}(X, \mathbb{R})$ , l'espace des fonctions continues sur  $X$ , et soit  $\xi$  une forme linéaire positive de cet espace dans  $\mathbb{R}$ , alors il existe une mesure  $\mu$  sur  $X$  telle que pour tout fonction  $f$  de  $\mathcal{C}(X, \mathbb{R})$  on ait :*

$$\xi(f) = \int_X f(u) d\mu(u)$$

Pour une démonstration on pourra consulter le Rudin[R].

**Remarque** : Cette mesure est de plus unique, en effet si tel n'était pas le cas alors on aurait deux mesures qui donnent la même valeur pour toutes les fonctions, on en déduirait alors que les mesures sont égales.

Traisons un peu le problème de convergence! Si je me donne une suite  $(\xi_n)_{n \in \mathbb{N}}$  de forme linéaire positive qui converge simplement vers  $\xi$  (qui est par conséquent une autre forme linéaire positive), la suite de mesure  $(\mu_n)_{n \in \mathbb{N}}$  associée à chacune des forme linéaire  $\xi_n$  converge-t-elle? Et pour quelle convergence?

Si on a une convergence simple alors :  $\forall f \in \mathcal{C}(X, \mathbb{R}), \lim_{n \rightarrow \infty} \xi_n(f) = \xi(f)$

Et ainsi  $\forall f \in \mathcal{C}(X, \mathbb{R}), \lim_{n \rightarrow \infty} \int_X f(u) d\mu_n(u) = \int_X f(u) d\mu(u)$ . Ce qui nous donne la convergence faible de  $(\mu_n)_{n \in \mathbb{N}}$  vers  $\mu$ . Le théorème de Banach-Alaoglu peut donc se ré-énoncer comme suit : Soit  $X$  un espace métrique compact et  $(\mu_n)_{n \in \mathbb{N}}$  une suite bornées de mesures finies de  $X$ , alors il existe une extraction  $(\mu_{n_k})_{k \in \mathbb{N}}$  telle qu'elle converge faiblement vers une mesure finie  $\mu$  de  $X$ .

## 5.2 A propos de la différentiabilité des fonctions lipschitziennes

**Théorème.** Soit  $f : [a, b] \rightarrow \mathbb{R}$  une fonction lipschitzienne alors  $f$  est différentiable presque partout.

*Démonstration.* On commence par faire remarquer qu'il est possible de généraliser ce résultat à des fonctions à valeurs dans  $\mathbb{R}^n$  en raisonnant composante par composante. On remarque également que si  $f$  est L-lipschitzienne alors la fonction  $g : x \rightarrow f(x) + Lx$  est également L-lipschitzienne et est de plus croissante ( $g$  est différentiable en  $a$  si et seulement si  $f$  l'est), on peut ainsi se restreindre aux fonctions croissantes.

La suite de la démonstration va reposer sur le fait que les limites inférieure et supérieure à gauche et à droite vont être égale (i.e : les quantités suivantes sont égales

$$D_d^+ f(x) = \overline{\lim}_{h \rightarrow 0^+} \frac{f(x+h) - f(x)}{h}, \quad D_d^- f(x) = \underline{\lim}_{h \rightarrow 0^+} \frac{f(x+h) - f(x)}{h},$$

$$D_g^+ f(x) = \overline{\lim}_{h \rightarrow 0^-} \frac{f(x+h) - f(x)}{h}, \quad D_g^- f(x) = \underline{\lim}_{h \rightarrow 0^-} \frac{f(x+h) - f(x)}{h} \quad )$$

On remarque que comme  $f$  est lipschitzienne les quantités qui viennent d'être posées sont toutes finies (car elles sont majorées par la constante de Lipschitz). Enfin on remarque que pour que toutes ces quantités soient égales il suffit d'avoir :  $D_d^+ f(x) \leq D_g^- f(x)$  presque partout, en effet en remplaçant  $f(x)$  par  $-f(-x)$  on a toujours  $D_d^+ - f(-x) \leq D_g^- - f(-x)$  presque partout (car  $-f(-x)$  est lipschitzienne et croissante) et on a  $D_d^+ - f(x) \leq D_g^+ f(x)$  et  $D_g^- f(x) \leq D_d^- f(x)$  ce qui nous donne  $D_g^+ f \leq D_d^- f$  et finalement on a la suite d'inégalités :

$$D_d^+ f(x) \leq D_g^- f(x) \leq D_g^+ f(x) \leq D_d^- f(x) \leq D_d^+ f(x)$$

Pour prouver le théorème il faut donc montrer que l'ensemble des éléments  $x$  qui ne vérifient pas cette inégalité est de mesure (de Lebesgue) nulle. On va en fait poser  $A_{r,s} = \{x, D_d^+ f(x) > s > r > D_g^- f(x)\}$  et montrer que pour tout  $s$  et  $r$  rationnels cette ensemble est de mesure de Lebesgue nulle. Donnons nous  $r$  et  $s$  rationnels et  $M = \mathcal{L}^1(A_{r,s})$  la mesure de notre ensemble. Pour tout  $\varepsilon$  positif on peut se donner un ouvert  $O$  tel que  $\mathcal{L}(O) < M + \varepsilon$  (en effet pour toute mesure borélienne  $\mu$  on a



$\mu(A) = \inf\{\mu(U), U \text{ est ouvert}, A \subset U\}$  lorsque  $A$  est un borélien). Ainsi pour tout point  $x \in A$  il existe une suite positive  $(h_i^x)_{i \in \mathbb{N}}$  telle que :  $\forall i \in \mathbb{N}, [x - h_i^x, x] \subset O$  et  $f(x) - f(x - h_i^x) \leq r h_i^x$  (la dernière inégalité provient du fait que  $r > D_g^- f$ ). De plus l'ensemble  $\mathfrak{F} = \{[x - h_i^x, x], x \in A, i \in \mathbb{N}\}$  forme un recouvrement fini de  $A$ . Je peux donc extraire un sous recouvrement fini et disjoint qui recouvre  $A$  à  $\varepsilon$  près. En enlevant les bornes (qui sont en quantité dénombrable et donc de mesure nulle) on a un ouvert  $O' (= \bigcup_{i=0}^n ]x_i - h_i, x_i[)$  tel que :  $\mathcal{L}^1(A \cap O') \geq M - \varepsilon$ .  
On a de plus que  $\sum_{i=0}^n h_i \leq M + \varepsilon$  car  $O' \subset O$ , ce qui nous donne donc :

$$\sum_{i=0}^n f(x_i) - f(x_i - h_i) \leq r \sum_{i=0}^n h_i \leq r(M + \varepsilon)$$

Par un raisonnement analogue on a pour tout élément  $y$  de  $A \cap O'$  l'existence d'une suite positive  $(k_i^y)_{i \in \mathbb{N}}$  telle que :  $\forall i \in \mathbb{N}, [y, y + k_i^y] \subset O'$  et  $f(y + k_i^y) - f(y) \geq s k_i^y$  (la seconde inégalité vient du fait que  $D_d^+ f > s$ ). Pour les mêmes raisons on a alors un recouvrement fini de  $A \cap O'$  et donc il existe un sous recouvrement disjoint et dénombrable. On a donc un ouvert  $O''$  tel que :  $\mathcal{L}^1(A \cap O' \cap O'') \geq M - 2\varepsilon$ . On a de plus que  $M - 2\varepsilon \leq \sum_{i=0}^m k_i \leq \sum_{i=0}^n h_i \leq M + \varepsilon$  ce qui nous donne finalement :

$$\sum_{i=0}^m f(y_i + k_i) - f(y_i) \geq s \sum_{i=0}^m k_i \geq s(M - 2\varepsilon)$$

De surcroît

$$s(M - 2\varepsilon) \leq \sum_{i=0}^m f(y_i + k_i) - f(y_i) \leq \sum_{i=0}^n f(x_i) - f(x_i - h_i) \leq r(M + \varepsilon)$$

(l'inégalité centrale est due au fait que  $]x_i, x_i + k_i[ \subset O'$  et que  $f$  est croissante). En faisant tendre  $\varepsilon$  vers 0 on obtient  $sM \leq rM$  et donc  $M = 0$ . On a donc ce qu'on souhaite.  $\square$

## Références

- [AT09] L. Ambrosio and P. Tilli. *Topics on Analysis in Metric Spaces*. Oxford University Press, 2009.  
[Rud98] W. Rudin. *Analyse réelle et complexe*. Dunod, 1998.



## Chapitre 5

# Reconstruction de mesures discrètes

# Table des matières

<b>1</b>	<b>Définition du problème et principaux résultats</b>	<b>2</b>
1.1	Un peu d'optique . . . . .	2
1.2	Reconstruction de mesures discrètes . . . . .	5
<b>2</b>	<b>En pratique</b>	<b>9</b>
2.1	Problème de minimisation, Lagrangien et problème dual . . . . .	9
2.2	reconstruction des minima . . . . .	10
<b>3</b>	<b>Pré-requis</b>	<b>11</b>

# 1 Définition du problème et principaux résultats

## 1.1 Un peu d'optique

Notre travail repose sur l'article [EC12] (nous aurons aussi besoin de [Rud98] et [LeG06]). On s'intéresse au problème suivant : Soit une source lumineuse  $\mu$  qui nous est inconnue subissant une diffraction au travers d'une fente.

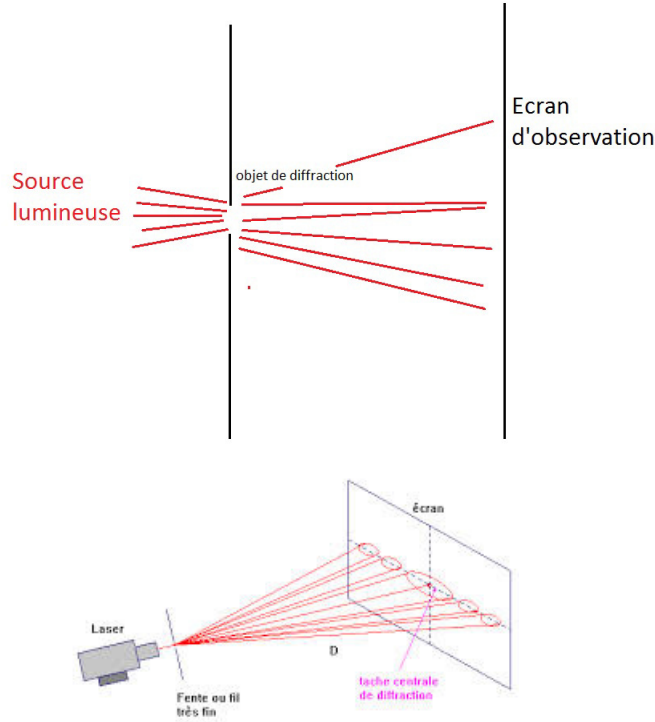


FIGURE 1 – Schémas de diffraction à travers une fente (unidimensionnelle)

**Question.** *A partir de l'observation de la diffraction subie peut-on retrouver la distribution de lumière  $\mu$  ?*

On se restreint, dans notre étude, aux cas où  $\mu$  est composée de sources ponctuelles (finies ou infinies) réparties dans un intervalle de distance définie (que l'on assimilera au compact  $[0, 1]$ ) et incohérentes (indépendantes les unes des autres).

Le principe de superposition est respecté (linéarité du système) : c'est à dire que si l'on connaît un ensemble de couples "entrée-sortie"  $((e_i(t), s_i(t)))_{i \in I}$  où le couple "entrée-sortie" représente l'intensité lumineuse de la source et de l'observation au cours du temps) alors la composition des entrées donne, en sortie, la même composition des sorties (i.e :  $(\sum_{i \in J} \lambda_i e_i(t), \sum_{i \in J} \lambda_i s_i(t))$  est également un couple entrée-sortie). Étant donnée la restriction apportée à notre étude et la remarque précédente, il nous est donc essentiel de connaître la sortie de notre système à une source lumineuse ponctuelle que l'on appellera solution fondamentale. En effet la réponse de notre système plus général ne sera qu'une combinaison linéaire de notre solution fondamentale.

**Étude de la solution fondamentale :** On se retrouve dans le cas d'une diffraction de FRAUNHOFER (i.e : une source ponctuelle éclaire une fente, la diffraction obtenue est observée sur un écran cf. FIGURE 1). On définit, pour un objet de diffraction  $D$ , l'intensité observée sur l'écran  $I_D(x, y)$ .

**Exemple.** 1.) Pour une fente unidimensionnelle la diffraction a lieu dans une seule direction cf. FIGURE 1 et l'expression de cette intensité est proportionnelle à

$$\text{sinc}^2\left(\frac{\pi ax}{\lambda d}\right)$$

(où  $a$  est la longueur de la fente,  $d$  la distance entre fente et écran,  $\lambda$  la longueur d'onde et  $x$  l'abscisse sur l'écran) cf. FIGURE 1.

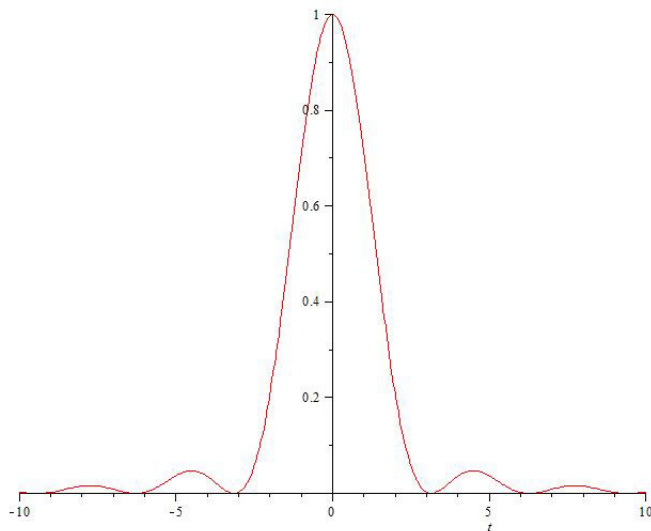


FIGURE 2 – Variation de la fonction intensité

2.) Dans le cas d'une fente rectangulaire on a que l'intensité est proportionnelle à

$$\text{sinc}^2\left(\frac{\pi ax}{\lambda d}\right)\text{sinc}^2\left(\frac{\pi by}{\lambda d}\right)$$

(où  $a$  et  $b$  sont respectivement la largeur et la hauteur de notre rectangle,  $d$  la distance à l'écran,  $\lambda$  la longueur d'onde et  $(x, y)$  les coordonnées sur l'écran) cf. FIGURE 3.

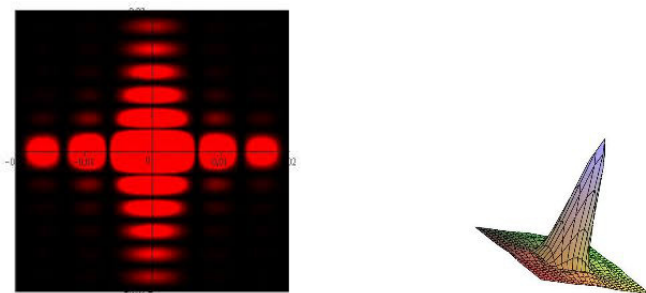


FIGURE 3 – Diffraction par une fente rectangulaire (gauche) et représentation de la fonction intensité lumineuse (droite)

3.) Enfin le cas d'une fente circulaire fait apparaître des fonctions de BESSEL sphériques cf. FIGURE 4.



FIGURE 4 – Diffraction par une fente circulaire ("Tache d'Airy")

La composition de plusieurs sources incohérentes nous donne un signal proportionnel à (la translation de la source ne provoque qu'un déphasage du signal) :

$$\sum_{i \in I} I_D(x - x_i, y - y_i)$$

où  $(x_i, y_i)$  est la localisation d'une source lumineuse ponctuelle. Et donc notre signal observé est la convolée de la solution fondamentale avec la source lumineuse :  $Y(x, y) = I_D * \mu(x, y)$ .

Nous pouvons alors faire des remarques sur les propriétés mathématiques de notre signal de sortie : la fonction  $Y$  a la régularité de la solution fondamentale, ainsi pour des régularités suffisantes (continument dérivable par exemple) la série de FOURIER associée à  $Y$  converge simplement vers  $Y$ . Pour des régularités moins importantes si les coefficients de la série de FOURIER de  $Y$  ne nous permettent pas une connaissance parfaite de  $Y$ , ils nous permettent cependant de savoir à quoi ne ressemble pas  $Y$ . Dans une grande majorité de cas la solution fondamentale est suffisamment régulière pour se retrouver dans le premier cas. On connaît ainsi notre signal de sortie, cependant les outils d'observation ont quelques imperfections et ne nous permettent qu'une connaissance partiel de notre signal (par exemple les hautes fréquences ne peuvent être observées : on a un filtre passe bas). Ainsi les observations nous permettent d'avoir accès qu'aux premiers termes de la série de FOURIER de la fonction sortie  $Y$ . Nos données sont donc les coefficients de FOURIER de notre observation :

$$\begin{aligned} \forall |k| \leq f_c, \quad \hat{y}(k) &= \int_{\mathbb{R}^2} \left( \int_{\mathbb{R}^2} I_D(t - u) \mu(u) du \right) e^{ikt} dt \\ &= \int_{\mathbb{R}^2} \sum_{j \in I} a_j I_D(t - u_j) e^{ikt} dt \\ &= \sum_{j \in I} a_j e^{iku_j} \int_{\mathbb{R}^2} I_D(t) e^{ikt} dt \\ &= \sum_{j \in I} a_j e^{iku_j} \hat{I}_D(k) \\ &= \hat{\mu}(k) \hat{I}_D(k) \end{aligned}$$

où  $f_c$  est la fréquence de coupure.

Voici enfin notre problème posé : à partir de l'observation de  $Y$  et de la connaissance de  $I_D$  retrouver la distribution de lumière  $S$ .

Ce problème est en fait beaucoup plus générale que le simple cas de la diffraction de FRAUNHOFER, on peut l'appliquer pour tous les types de signaux : lumineux (comme on vient de le voir), radio, sismique, et pour les ondes en général. Etudier un tel problème et le résoudre a donc une portée assez générale.

Une fois ces motivations posées on va étudier la résolution d'un tel problème. Pour cela on considéra le cas simplifié où les coefficients de FOURIER de notre intensité sont égaux à 1 pour tout  $k$  inférieur à  $f_c$  (le cas d'un filtre passe bas parfait), on se retrouve alors avec une intensité qui est en fait le noyau de DIRICHLET, le filtre agit alors comme une transformée de FOURIER.

Ce théorème de physique traite plus généralement du cas qui nous intéresse :

**Théorème.** *Un système physique peut être décrit par un opérateur de convolution (ou filtre) si et seulement si il est linéaire, continu et invariant par translation.*

*Démonstration.* Soit  $A$  un opérateur linéaire (ou filtre) continu et invariant par translation, i.e. :

$$\begin{aligned} \star A(\lambda f + \mu g) &= \lambda A(f) + \mu A(g), \\ \star A(f(t - \tau)) &= A(f)(t - \tau) \end{aligned}$$

La convolution correspond à la réponse du filtre à une entrée donnée (notée  $f(t)$ ). Le filtre est entièrement caractérisé par sa réponse impulsionnelle  $h(t)$ . La réponse du filtre est  $s(t) = A(f)(t) = h(t) * f(t)$ .

À partir de là, on peut construire la réponse du filtre linéaire et indépendant du temps à l'entrée  $f(t)$ . En effet, comme le filtre est linéaire, on peut décomposer le signal  $f(t)$  en parties indépendantes, à l'aide d'un ensemble de signaux  $f_i$  avec des supports disjoints compacts de telle sorte que  $f(t) = \sum_i f_i(t)$ . On injecte chaque partie du signal dans le filtre puis l'on somme les différentes réponses. Ainsi le filtrage donnera :  $A(f) = \sum_i A(f_i)$ . Cette décomposition temporelle de  $f(t)$  peut s'effectuer de manière récursive sur les signaux  $f_i(t)$ . À la fin, on obtient une suite de signaux dont le support se résume à un point. Ces signaux, élémentaires parce que non décomposables temporellement, correspondent chacun d'entre eux à la distribution de Dirac  $\delta_0(t - \tau)$  centrée en  $\tau$  avec une amplitude  $f(\tau)$ , l'impulsion s'écrit  $\delta_0(t - \tau) \cdot f(\tau)$ . Il suffit de sommer toutes les impulsions suivant la variable  $\tau$  pour obtenir le signal  $f(t)$  :

$$f(t) = \int \delta_0(t - \tau) \cdot f(\tau) d\tau$$

On applique l'opération de filtrage sur  $f(t)$ . Comme le filtre est linéaire et indépendant du temps, nous avons :

$$\begin{aligned} A(f)(t) &= \int A(\delta_0(t - \tau) \cdot f(\tau)) d\tau \\ &= \int f(\tau) \cdot A(\delta_0(t - \tau)) d\tau \quad (\text{linéarité}) \\ &= \int f(\tau) \cdot (A \circ \delta_0)(t - \tau) d\tau \quad (\text{indépendance du temps}) \end{aligned}$$

La réponse du filtre  $A$  à l'impulsion  $\delta_0(t)$  est nommée la réponse impulsionnelle du filtre  $h(t)$ . Finalement on a :

$$A(f) = \int f(\tau) \cdot h(t - \tau) d\tau = h(t) * f(t)$$

qui n'est qu'autre que le produit de convolution. □

## 1.2 Reconstruction de mesures discrètes

On se donne la mesure  $\mu$  discrète suivante :

$$\mu = \sum_{j \in J} a_j \delta_{t_j}$$

On pose  $A = \{a_j \in \mathbb{C}, j \in J\}$  et  $T = \{t_j, j \in J\} \subset [0, 1]$ .

On s'intéresse ensuite aux coefficients de Fourier de notre mesure. On pose donc :

$$\forall |k| \leq f_c, \quad y_\mu(k) = \int_0^1 e^{-2i\pi kt} d\mu(t) = \sum_{j \in J} a_j e^{-2i\pi kt_j}$$



Notre but est de retrouver la mesure  $\mu$  à partir de la donnée des  $y_\mu(k), |k| \leq f_c$ . Pour simplifier les notations nous allons commencer par noter  $\mathfrak{F}_n$  l'application linéaire définie comme suit :  $\mathfrak{F}_n(\tilde{\mu}) = y_{\tilde{\mu}}$ . On pose  $S = \{\tilde{\mu}, \mathfrak{F}_n(\tilde{\mu}) = y_\mu\}$ , on sait par ailleurs que  $\mu \in S$ , donc  $S$  est non vide.

On peut commencer par se convaincre que  $S$  n'est pas nécessairement composé d'une seule mesure (plusieurs mesures peuvent être réponse de notre problème) : en effet si on se donne la suite définie par

$$\begin{cases} u_k = 0 & \text{si } |k| \leq f_c \\ u_k = \frac{1}{k^2} & \text{sinon} \end{cases}$$

définie le terme général d'une série convergente (mais aussi de carré sommable). On en déduit que la fonction

$$f(t) = \sum_{k \geq 0} u_k e^{2i\pi kt}$$

est de carré intégrable et qu'elle est même continue par convergence normale (et elle n'est pas identiquement nulle). Ainsi on peut définir la mesure à densité  $f(t)dt$  sur  $[0, 1]$  et alors par définition on a  $f(t)dt \in \text{Ker}(\mathfrak{F}_n)$  et si  $\tilde{\mu}$  appartient à  $S$  alors  $\tilde{\mu} + f(t)dt$  appartient également à  $S$ . Tout ceci nous montre que la condition linéaire qui nous a permis de définir  $S$  ne détermine pas uniquement  $\mu$ .

**Question.** *Peut-on alors trouver des conditions supplémentaires pour que notre mesure "originelle" soit uniquement déterminée ?*

En pratique, en posant :  $\tilde{S} = \{\|\tilde{\mu}\|_{TV}, \mathfrak{F}_n(\tilde{\mu}) = y_\mu\}$  on aura, sous certaines conditions, que  $\mu$  réalise le minimum de  $\tilde{S}$ , et sous des conditions supplémentaires on aura que ce minimum est unique. Les conditions suffisantes nous sont données par le théorème qui suit.

**Définition.** *On appelle polynôme trigonométrique à basses fréquences un polynôme de la forme suivante :*

$$q(t) = \sum_{k=-f_c}^{f_c} c_k e^{2i\pi kt}$$

De plus on dira qu'un tel polynôme interpole  $T$  en  $v$ , où  $v \in \mathbb{C}^{|T|}$  et  $\forall j \in J, |v_j| = 1$ , si :

$$\begin{cases} q(t_j) = v_j & , \forall j \in J \\ |q(t)| < 1 & , \forall t \in [0, 1] \setminus T \end{cases}$$

Une condition nécessaire pour la reconstruction de mesure est l'existence pour tout vecteur  $v \in \mathbb{C}^{|T|}$  d'un polynôme à basse fréquence interpolateur en  $v$ . Ce polynôme est nécessaire pour la démonstration de l'unicité de la solution qui minimise  $\tilde{S}$  et pour la reconstruction algorithmique de la mesure  $\mu$ . Ce qui nous est donné par le théorème majeur suivant :

**Théorème** (CANDÈS & FERNANDEZ-GRANDA). *Soit  $v \in \mathbb{C}^{|T|}$  un vecteur quelconque vérifiant  $|v_j| = 1$  pour tout  $j$ , sous la condition que les éléments du support de  $\mu$  soit séparés d'une distance minimale de  $2/f_c$ . Alors, il existe un polynôme trigonométrique à basses fréquences qui interpole  $v$ .*

**Théorème.** *Supposons que pour tout vecteur  $v \in \mathbb{C}^{|T|}$  il existe un polynôme trigonométrique à basses fréquences qui interpole  $T$  en  $v$  alors  $\mu$  est l'unique solution qui minimise  $\tilde{S}$ .*

*Démonstration.* On commence par démontrer l'**unicité** de ce minimum.

On se donne  $\tilde{\mu} \in S$  et  $h = \tilde{\mu} - \mu$ . On peut décomposer  $h$  en une somme de deux mesures : la première discrète à support dans  $T$ , et l'autre à support dans  $T^c$ . On a donc :  $h = h_T + h_{T^c}$ .

---

1.  $|T|$  : cardinal de  $T$

Arrivé ici nous pouvons commencer à nous demander ce qui contredirait la minimalité de  $\tilde{\mu}$ , nous avons les résultats suivant :

$$\|\mu\|_{TV} \geq \|\tilde{\mu}\|_{TV} = \|\mu + h\|_{TV} = \|\mu + h_T\|_{TV} + \|h_{T^c}\|_{TV} \geq \|\mu\|_{TV} - \|h_T\|_{TV} + \|h_{T^c}\|_{TV}$$

Si nous avons, en supposant  $h \neq 0$ , que  $\|h_T\|_{TV} < \|h_{T^c}\|_{TV}$ , nous aboutirions à :  $\|\mu\|_{TV} > \|\mu\|_{TV}$  ce qui nous donnerait une contradiction et finalement nous aurions ce que l'on souhaite.

Une réciproque du théorème de RADON-NYKODYM nous donne l'existence d'une fonction  $\varphi$  telle que :  $h_T = e^{2i\pi\varphi(t)}|h_T|$ . Les hypothèses du théorème nous assure l'existence d'un polynôme trigonométrique à basses fréquences qui interpole  $T$  en  $e^{-2i\pi\varphi(t)}$ . On pose :

$$q(t) = \sum_{k=-f_c}^{f_c} c_k e^{-2i\pi kt}$$

ce polynôme, alors

$$\int_0^1 q(t) d\tilde{\mu}(t) = \sum_{k=-f_c}^{f_c} c_k \int_0^1 e^{-2i\pi kt} d\tilde{\mu}(t) = \sum_{k=-f_c}^{f_c} c_k \int_0^1 e^{-2i\pi kt} d\mu(t) = \int_0^1 q(t) d\mu(t)$$

car  $\tilde{\mu} \in S$ .

A partir de là on remarque que sous les conditions d'existence du polynôme et si  $h_{T^c} \neq 0$ , alors :

$$\left| \int_{[0,1]} q(t) dh_{T^c}(t) \right| \leq \int_{T^c} \underbrace{|q(t)|}_{<1} |dh_{T^c}(t)| < \int_{[0,1]} |dh_{T^c}(t)| \leq \|h_{T^c}\|_{TV}$$

Il s'ensuit :

$$0 = \int_0^1 q(t) dh(t) = \int_0^1 q(t) dh_T(t) + \int_0^1 q(t) dh_{T^c}(t) = \underbrace{\sum_{j \in J} e^{2i\pi(\varphi(t_j) - \varphi(t_j))} |h_T|(t_j)}_{=\|h_T\|_{TV}} + \int_0^1 q(t) dh_{T^c}(t)$$

et donc :

$$0 = \|h_T\|_{TV} + \int_0^1 q(t) dh_{T^c}(t) \geq \|h_T\|_{TV} - \left| \int_0^1 q(t) dh_{T^c}(t) \right| > \|h_T\|_{TV} - \|h_{T^c}\|_{TV}$$

Ce qui nous donne finalement :  $\|h_T\|_{TV} < \|h_{T^c}\|_{TV}$ .

On peut ensuite montrer l'**existence** de ce minimum. Les principaux arguments ont déjà été donnés dans la partie précédente : nous avons pour toute mesure  $\tilde{\mu} \in S$  le résultat précédent  $\|\mu\|_{TV} \leq \|\tilde{\mu}\|_{TV}$ , ainsi en se donnant une suite d'éléments de  $\tilde{S}$  qui tend vers l'infimum de  $\tilde{S}$  nous avons que chacun de ces éléments est supérieur à  $\|\mu\|_{TV}$ . Ensuite par passage à la limite on a que l'infimum est supérieur à la variation totale de  $\mu$ , donc on a bien l'existence de notre minimum. □

**Application.** Reprenons le système optique décrit au début de ce travail. Suite à la diffraction un appareil enregistre le signal lumineux situé sur l'écran. Cette appareil ne peut enregistrer la totalité du signal, il se contente de garder une valeur moyenne de la luminosité de l'écran ce qui revient à dire que c'est un filtre passe-bas dont la fréquence de coupure varie en fonction de la qualité du matériel. Les meilleurs appareils peuvent ainsi donner des fréquences de coupure de l'ordre de  $10^{10}$  hertz. En appliquant cela à notre système cela signifie que théoriquement il nous est possible de retrouver des sources dont les points lumineux sont au minimum écarté de  $2 \cdot 10^{-10}$  (tout cela manque quelque peu d'unité :

on peut prendre l'intervalle  $[0, 1]$  comme valeur pour un mètre[ou centimètre] pour fixer les idées et les ordres de grandeurs), soit environ la taille d'un atome. On se rend ainsi compte que malgré les potentielles restrictions théoriques qu'a ce théorème, il n'en est en fait tout autrement dans la réalité. Ceci est principalement dû à la qualité des appareils utilisés dont la précision est très importante.

## 2 En pratique

L'ensemble des considérations et des réflexions précédentes nous ont permis de proposer une description exacte du signal d'entrée de notre système, sous des conditions d'écartement minimum des points du support de la mesure. Cependant un autre problème se pose maintenant : celui de trouver la forme de la mesure qui réalise ce minimum.

Durant cette partie nous allons présenter le problème de minimisation lié à notre question et poser les problèmes primaux, duaux, Lagrangien qui doivent permettre de mieux comprendre les moyens de retrouver la mesure solution de notre problème à partir des données. Cependant certaines équivalences ne seront pas justifiées et pour l'application de ces méthodes nous renvoyons au programme écrit par EMMANUEL CANDÈS et CARLOS FERNANDEZ-GRANDA .

### 2.1 Problème de minimisation, Lagrangien et problème dual

En posant :  $\varphi = \varphi(t) = (e^{2i\pi kt})_{-f_c \leq k \leq f_c}$ .

Le problème de minimisation que nous voulons résoudre est :

$$\inf_{\tilde{\mu}} \|\tilde{\mu}\|_{VT} \quad \forall |k| \leq f_c, \int e^{2i\pi kt} d\tilde{\mu}(t) = \int \varphi(t) d\tilde{\mu} = y_\mu(k)$$

Nous passons au problème Lagrangien associé :

$$\inf_{\tilde{\mu}} \sup_a \|\tilde{\mu}\|_{VT} + \langle y_\mu - \int \varphi(t) d\tilde{\mu}(t), a \rangle = \inf_{\tilde{\mu}} \sup_a L(\tilde{\mu}, a)$$

Qui est bien équivalent à notre au premier problème, en effet

$$\sup_a L(\tilde{\mu}, a) = \begin{cases} \|\tilde{\mu}\|_{VT} & \text{si } y_\mu = \int \varphi d\tilde{\mu} \\ +\infty & \text{sinon} \end{cases}$$

et donc

$$\inf_{\tilde{\mu}} \sup_a L(\tilde{\mu}, a) = \inf_{\tilde{\mu}, y_\mu = \int \varphi d\tilde{\mu}} \|\tilde{\mu}\|_{VT}.$$

Nous pouvons alors regarder le problème dual qui y est lié à notre problème Lagrangien :

$$\sup_a \inf_{\tilde{\mu}} \|\tilde{\mu}\|_{VT} + \langle y_\mu - \int \varphi(t) d\tilde{\mu}(t), a \rangle = \sup_a \inf_{\tilde{\mu}} L(\tilde{\mu}, a).$$

On montre le résultat intermédiaire suivant

$$\left| \int \langle \varphi(t), a \rangle d\tilde{\mu}(t) \right| \leq \| \langle \varphi, a \rangle \|_\infty \|\tilde{\mu}\|_{VT}$$

En effet on a

$$\left| \int f d\tilde{\mu} \right| = \left| \int f d\tilde{\mu}^+ - \int f d\tilde{\mu}^- \right| \leq \left| \int f d\tilde{\mu}^+ \right| + \left| \int f d\tilde{\mu}^- \right| \leq \int |f| d\tilde{\mu}^+ + \int |f| d\tilde{\mu}^- = \int |f| d|\tilde{\mu}|$$

De plus  $|\langle \varphi(t), a \rangle| \leq \| \langle \varphi, a \rangle \|_\infty$  pour tout  $t$  dans  $[0, 1]$ . Par conséquent

$$\left| \int \langle \varphi(t), a \rangle d\tilde{\mu}(t) \right| \leq \int \| \langle \varphi, a \rangle \|_\infty d|\tilde{\mu}|(t) = \| \langle \varphi, a \rangle \|_\infty \|\tilde{\mu}\|_{VT}$$

Le problème dual devient donc :

$$\inf_{\tilde{\mu}} L(\tilde{\mu}, a) = \begin{cases} \langle y_\mu, a \rangle & \text{si } \| \langle \varphi, a \rangle \| \leq 1 \\ -\infty & \text{sinon} \end{cases}$$

Donc finalement le problème dual s'écrit :

$$\sup_a \inf_{\tilde{\mu}} L(\tilde{\mu}, a) = \sup_{a, \| \langle \varphi, a \rangle \| \leq 1} \langle y_\mu, a \rangle$$

## 2.2 reconstruction des minima

Une fois le problème de minimisation lagrangien posé et la valeur minimale connue (ce qui est possible de manière approché par un ordinateur), il nous reste à trouver la valeur pour laquelle ce minimum est atteint et le lien entre cet élément et la mesure à reconstruire.

Retrouver la valeur pour laquelle le problème atteint son optimum n'est pas toujours possible avec les techniques actuelles, mais il existe certaine classe de problème pour lesquels nous avons des méthodes pour l'obtenir, nous sommes dans une telle classe de problèmes. En effet, si on pose  $A = \{a \mid \langle \varphi, a \rangle \|_\infty \leq 1\}$  alors pour tout  $a$  dans  $A$  on a

$$\forall t \in [0, 1], \begin{cases} 1 - \langle \varphi, a \rangle \geq 0 \\ 1 + \langle \varphi, a \rangle \geq 0 \end{cases}$$

D'après le théorème de RIESZ-FÉJER ceci est équivalent à l'existence de deux polynômes  $P$  et  $Q$  tels que

$$\begin{cases} 1 - \langle \varphi, a \rangle = |P|^2 \\ 1 + \langle \varphi, a \rangle = |Q|^2 \end{cases} \Leftrightarrow \begin{cases} \langle \varphi, b \rangle = |P|^2 \\ \langle \varphi, c \rangle = |Q|^2 \end{cases}$$

Cela est encore équivalent au fait qu'il existe deux matrices semi-définie positive  $M_b$  et  $M_c$ . On peut alors définir une matrice  $M_a$  diagonale par bloc dont les éléments sont les deux matrices précédentes. Il nous est alors possible de retrouver la valeur prise par  $\tilde{a}$  à partir de cette matrice.

Nous sommes maintenant en possession de la valeur  $\tilde{a}$  qui réalise le minimum. Cette valeur est liée à notre mesure par le résultat suivant : le support de notre mesure est contenu dans l'ensemble des valeurs maximales du polynôme  $\langle \varphi(t), \tilde{a} \rangle$  (polynôme qui a la propriété d'être de norme inférieure ou égale à 1) et donc sur les points du support de la mesure solution le polynôme est de norme exactement 1 et il y a au plus  $2f_c + 1$  tels points. Il nous suffit maintenant de trouver les amplitudes de notre mesure solution sur chacun des points du support. Pour cela on trouve les maximum du polynôme dual (et donc le support de la mesure) que l'on note  $U = \{t, \mid \langle \varphi, \tilde{a} \rangle(t) = 1\}$  (et  $|U| \leq 2f_c + 1$ , cependant quitte à ajouter des valeurs on peut supposer que le support est de cardinal  $2f_c + 1$ ), on connaît ainsi la forme de l'application linéaire  $\mathfrak{F}_n$  ( $(e^{2i\pi kt})_{-f_c \leq k \leq f_c, t \in U}$  est une matrice de VANDERMONDE). On a ainsi une matrice de rang  $|U|$  et donc inversible, on en déduit l'intensité de la mesure aux points de  $U$  par inversion :

$$(e^{2i\pi kt})_{-f_c \leq k \leq f_c, t \in U} \cdot (\mu_i)_{1 \leq i \leq 2f_c + 1} = y_\mu$$

donc,

$$(\mu_i)_{1 \leq i \leq 2f_c + 1} = \mathfrak{F}_n^{-1} y_\mu$$

On obtient donc finalement l'ensemble des informations sur notre fameuse mesure, que l'on a pu retrouver.

### 3 Pré-requis

**Définition** (Mesure Signée). Soit  $\mu : (\Omega, \mathcal{A}) \rightarrow \mathbb{R}$  vérifiant les propriétés habituelles des mesures sigma-finies (i.e :  $\mu(\emptyset) = 0$ , sigma-additivité, sigma-finitude) alors cette "mesure" est appelée mesure signée. Une mesure signée peut être décomposée comme la différence de deux mesures positives notées  $\mu_-$  et  $\mu_+$ , on a donc :  $\mu = \mu_+ - \mu_-$  (c'est la décomposition de HAHN).

**Définition.** Soit  $\mu$  une mesure complexe, on définit alors la mesure positive  $|\mu|$  par l'égalité

$$|\mu|(E) = \sup \left\{ \sum_{n \in \mathbb{N}} |\mu(A_n)|, A = \bigsqcup_{n \in \mathbb{N}} A_n \right\}$$

pour tout ensemble  $\mu$ -mesurable  $A$ .

**Remarque :** On se convainc rapidement que  $|\mu|$  est une mesure positive. De plus on a pour tout élément  $A$  de la tribu :  $|\mu(A)| \leq |\mu|(A)$ .

Nous pouvons alors poser les fonctions  $\mu^+ = \frac{1}{2}(\mu + |\mu|)$  et  $\mu^- = \frac{1}{2}(|\mu| - \mu)$  qui définissent toutes deux des mesures positives et qui vérifient les égalités suivantes

$$\mu = \mu^+ - \mu^- \text{ et } |\mu| = \mu^+ + \mu^-$$

Ce qui nous permet d'avoir la décomposition citée.

**Définition** (Support d'une Mesure). Soit  $\mathfrak{F} = \{F \text{ fermés tel que } \mu({}^c F) = 0\}$ , on appelle support d'une mesure l'ensemble

$$\text{Supp}(\mu) = \bigcap_{F \in \mathfrak{F}} F$$

Conséquence : On fait remarquer que dans les cas où l'on considère des mesures sur l'espace  $\mathbb{R}^n$  le support de  $\mu$  a lui même un complémentaire de mesure nulle. En effet on a :

$${}^c \text{Supp}(\mu) = \bigcup_{F \in \mathfrak{F}} {}^c F$$

De plus comme l'espace considéré est séparable alors un théorème de LINDELHOF nous assure que l'on peut considérer le recouvrement de  ${}^c \text{Supp}(\mu)$  par des ouverts comme dénombrable. Il s'ensuit

$$\mu({}^c \text{Supp}(\mu)) \leq \sum_{i \in I} \mu({}^c F_i) = 0$$

ce qui nous donne le résultat souhaité. On a ainsi que tout sous ensemble contenu dans le complémentaire du support est de mesure nulle.

**Définition** (Norme en Variation Totale). Soit  $\mu$  un mesure signée, on défini la norme

$$\|\mu\|_{VT} = \sup \left\{ \sum_{i \in I} |\mu(A_i)|, \Omega \subset \bigsqcup_{i \in I} A_i \right\}.$$

*Démonstration.* Cette quantité est clairement positive.

Si  $\|\mu\|_{VT} = 0$  alors on a pour tout élément  $A$  de la tribu considérée on a  $\mu(A) = 0$  et donc la mesure est nulle.

Puis l'inégalité triangulaire :

$$\begin{aligned} \|\mu + \nu\|_{VT} &= \sup \left\{ \sum_{i \in I} |(\mu + \nu)(B_i)|, \Omega \subset \bigsqcup_{i \in I} B_i \right\} \\ &\leq \sup \left\{ \sum_{i \in I} |\mu(B_i)| + |\nu(B_i)|, \Omega \subset \bigsqcup_{i \in I} B_i \right\} \\ &\leq \|\mu\|_{VT} + \|\nu\|_{VT} \end{aligned}$$

□

**Lemme.** Soient  $\mu$  et  $\nu$  deux mesures sur  $\Omega$  à supports disjoints alors on a :  $\|\mu + \nu\|_{VT} = \|\mu\|_{VT} + \|\nu\|_{VT}$

*Démonstration.* On nomme  $T_\mu$  et  $T_\nu$  les supports de chacune des mesures. On raisonne par double inégalité, la première est due à l'inégalité triangulaire car on a une norme.

Pour la seconde inégalité on commence par montrer que pour avoir la valeur de la norme de  $\mu$  il suffit de prendre le suprémum sur un recouvrement du support de notre mesure. En effet

$$\begin{aligned} \|\mu\|_{VT} &= \sup \left\{ \sum_{i \in I} |\mu(A_i)|, \Omega \subset \bigsqcup_{i \in I} A_i \right\} \\ &= \sup \left\{ \sum_{i \in I} |\mu(A_i \cap T_\mu)|, \Omega \subset \bigsqcup_{i \in I} A_i \right\} \\ &\leq \sup \left\{ \sum_{i \in I} |\mu(C_i)|, T_\mu \subset \bigsqcup_{i \in I} C_i \right\} \end{aligned}$$

On a une inégalité, pour avoir la seconde égalité, il suffit de dire que tout recouvrement de  $T_\mu$  peut être "prolongé" en un recouvrement de  $\Omega$  et ainsi on a l'égalité souhaitée.

On se donne  $(A_i)_{i \in I_A}$  (respectivement  $(B_i)_{i \in I_B}$ ) une partition de  $T_\mu$  (respectivement  $T_\nu$ ) tels que

$$\|\mu\|_{VT} - \varepsilon \leq \sum_{i \in I_A} |\mu(A_i)| \text{ et } \|\nu\|_{VT} - \varepsilon \leq \sum_{i \in I_B} |\nu(B_i)|$$

Et on a donc

$$\|\mu\|_{VT} + \|\nu\|_{VT} - 2\varepsilon \leq \sum_{i \in I_A} |\mu(A_i)| + \sum_{i \in I_B} |\nu(B_i)| \leq \|\mu + \nu\|_{VT}$$

car les mesures sont à supports disjoints. Puis en faisant tendre  $\varepsilon$  vers 0 on obtient bien ce que l'on souhaite.  $\square$

**Théorème (RADON-NYKODYM-LEBESGUE).** Soient  $\nu$  une mesure positive  $\sigma$ -finie sur  $(X, \mathcal{A})$  et  $\mu$  une mesure positive  $\sigma$ -finie (resp. réelle, resp. complexe) sur  $(X, \mathcal{A})$ . Alors :

(i) Il existe un unique couple de mesures positives  $\sigma$ -finies (resp. réelles, resp. complexes)  $(\mu_1, \mu_2)$  tel que :

- \*  $\mu = \mu_1 + \mu_2$ ,
- \*  $\mu_1 \ll \nu$ ,
- \*  $\mu_2 \perp \nu$ .

Cette décomposition s'appelle la décomposition de LEBESGUE de  $\mu$ .

(ii) Il existe une unique (à égalité  $\nu$ -presque partout près) fonction  $h$   $\nu$ -intégrable mesurable positive (resp. réelle, resp. complexe), telle que pour tout  $A \in \mathcal{A}$  on ait :

$$\mu_1(A) = \int_A h \, d\nu = \int_X 1_A h \, d\nu.$$

Cette fonction s'appelle la dérivée de RADON-NYKODYM de  $\mu_1$  par rapport à  $\nu$ .

## Références

- [EC12] C. Fernandez-Granda E. Candès. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics* 67(6), 906–956, 2012.
- [LeG06] J-F. LeGall. Intégration, probabilités et processus aléatoires. *Cours de l'École Normale Supérieure*, Septembre 2006.
- [Rud98] W. Rudin. *Analyse réelle et complexe*. Dunod, 1998.





Troisième partie

Rapport de stage de Master 2





## Remerciements

A tous les stagiaires qui ont eu le courage de me supporter : Hélène, Catherine, Marguerite, Tony, Antoine et Farooq.

A Yannig pour ses cours de R, son encadrement et ses directives.

Mention spéciale pour Amandine qui m'a soutenu et surtout m'a pris en stage!

Au groupe R39 en général (j'aurai bien fait une dédicace spécifique à chacun d'entre vous mais il commence à se faire tard en ce dernier jour de stage et que je fatigue).

A M. PAULIN qui fera parti des trois ou quatre personnes à lire ces remerciements et ce mémoire.

"Bon allez... On bosse ?!"

# Table des matières

<b>1</b>	<b>Etude théorique des modèles de régression</b>	<b>5</b>
1.1	Modèle Statistique . . . . .	5
1.2	Validation croisée : Définitions . . . . .	8
1.2.1	Introduction . . . . .	8
1.2.2	Validation croisée pour la sélection . . . . .	8
1.2.3	Validation croisée généralisée . . . . .	9
1.3	Les modèles linéaires généralisés . . . . .	11
1.3.1	Présentation du modèle . . . . .	11
1.3.2	Estimation du modèle GLM . . . . .	11
1.4	Régression Ridge : présentation & plus . . . . .	13
1.4.1	Introduction . . . . .	13
1.4.2	& plus . . . . .	14
1.5	Les modèles additifs généralisés . . . . .	15
1.5.1	Présentation . . . . .	15
1.5.2	Estimation effective du modèle . . . . .	16
<b>2</b>	<b>Applications</b>	<b>17</b>
2.1	Remarques, informations sur les données et objectifs . . . . .	17
2.1.1	Vocabulaire . . . . .	17
2.1.2	Problématique et objectifs . . . . .	18
2.1.3	Les variables explicatives . . . . .	20
2.1.4	Estimation des modèles . . . . .	21
2.2	Comparaison entre une réestimation GAM et une régression linéaire sur les effets	22
2.2.1	Performances par instant . . . . .	23
2.2.2	Performances par jour . . . . .	24
2.2.3	Performances par mois . . . . .	25
2.2.4	Conclusion . . . . .	25
2.3	Les effets croisés . . . . .	25
2.4	Les effets fixés . . . . .	27
2.4.1	Performances par instant . . . . .	28
2.4.2	Performances par jour . . . . .	28
2.4.3	Performances par mois . . . . .	29
2.4.4	Conclusion . . . . .	30
2.5	Sélection du facteur d'oubli à priori . . . . .	31
2.6	Comportement des coefficients estimés . . . . .	32
2.7	Ouverture . . . . .	33

<b>3</b>	<b>ANNEXES</b>	<b>35</b>
3.1	les effets fixés . . . . .	35
3.1.1	Performances par jour . . . . .	35
3.1.2	Performances par mois . . . . .	36
3.2	Une perte de clients plus importante (30%) . . . . .	37
3.2.1	Performances globales . . . . .	38
3.2.2	Performances par instant . . . . .	39
3.2.3	Performances par jour . . . . .	39
3.2.4	Performances par mois . . . . .	40
3.2.5	Conclusion . . . . .	41
3.3	Une perte de clients progressive . . . . .	41
3.4	Facteur d'oubli inverse quadratique (perte de 10% au 1 <sup>er</sup> janvier 2012 . . . . .	43
3.5	RKHS : Introduction & plus . . . . .	44
3.5.1	Définition et notation . . . . .	44
3.5.2	& plus . . . . .	47
3.5.3	RKHS et apprentissage statistique : cas particulier des bases de splines . .	50

## Introduction-Problématique & autre forme de prélude compréhensible

Ce rapport fait suite à un stage de cinq mois effectué au sein d'EDF en Recherche et Développement (R&D), dans le département OSIRIS (Optimisation SIMulation RISques et Statistiques), prévision de consommation électrique court et moyen terme [groupe R39].

Cet axe de recherche s'inscrit dans un ensemble d'autres qui permettent à EDF d'avoir une maîtrise (et une expertise) sur l'ensemble des maillons de la chaîne du marché de l'énergie (aussi bien au niveau matériel [gestion et amélioration des installations,...] qu'intellectuel [prévision des consommations à court et long terme, prévision de l'évolution du marché de l'énergie,...]).

La prévision de consommation à court et moyen terme est une activité essentielle pour une entreprise qui possède la majorité des moyens de production électrique. Ces prévisions doivent fournir les informations nécessaires pour permettre une production en adéquation avec la consommation. Elles permettent aussi de comprendre et d'analyser les changements de comportements dans la consommation électrique ; par exemple la prévision à long terme (au delà de cinq ans) doit pouvoir analyser les changements de moyens de productions et de comportements des consommateurs.

Depuis quelques années le marché de l'électricité s'ouvre à la concurrence (depuis 2000 pour les entreprises et 2007 pour les particuliers). Ainsi EDF n'est plus seul fournisseur d'électricité en France. Dans les prochains mois certains tarifs réglementés vont disparaître pour permettre une plus grande concurrence. En conséquence EDF s'attend à des évolutions de son portefeuille clients.

Le modèle développé par le groupe R39 ne permet pas, sans l'intervention des prévisionnistes, d'avoir une estimation fiable de la consommation suite à des fluctuations trop importantes dans la consommation électrique. Il faudrait être capable d'adapter (ou de généraliser) le modèle utilisé pour prendre en compte ces variations et donner plus de souplesse aux méthodes de prévision.

Ce rapport résume les travaux qui ont été effectués et les résultats qui ont été obtenus sur ce sujet.

# Chapitre 1

## Etude théorique des modèles de régression

### 1.1 Modèle Statistique

Dans l'ensemble de ce document  $n$  représente le nombre d'observations (la taille des données),  $d$  le nombre de variables explicatives. L'utilisation de ces constantes pour nommer d'autres quantités pourra être considérée comme une faute (grave!).

**Notation.** Soient

$(Y_i)_{i \leq n} \in \mathbb{R}^n$ , les observations (on notera  $Y$  le vecteur colonne des observations).  
 $(X_i)_{i \leq n} \in (\mathbb{R}^d)^n$ , les variables explicatives (on notera  $X$  la matrice  $n \times d$  des variables explicatives [cette quantité est déterministe]).  
 $D_n = (Y_i, X_i)_{i \leq n} \in (\mathbb{R} \times \mathbb{R}^d)^n$ .

Le cadre est celui de la régression. On suppose qu'il existe un lien entre les observations et les variables explicatives donné par l'équation :

$$Y_i = f(X_i) + \varepsilon_i \quad , \quad \text{où } E(\varepsilon_i) = 0 \text{ et } Var(\varepsilon_i) = \sigma^2 \quad (1.1)$$

Le choix de la fonction  $f$  est naturellement notre principal problème. Ce choix doit se faire selon des critères précis permettant de répondre aux questions suivantes

**Question.** *Quelle est la "meilleure" fonction que je puisse choisir ?*

**Question.** *Comment puis-je, en pratique, la trouver ?*

**Exemples.** — *La famille linéaire.*

$$f(x) = ax \quad \text{dans le cas de univarié}$$

$$f(x_1, \dots, x_d) = a_1x_1 + \dots + a_dx_d \quad \text{dans le cas de multivarié}$$

— *La famille polynomiale.*

$$f(x) = \sum a_i x^i \quad \text{dans le cas de univarié}$$

$$f(x_1, \dots, x_d) = \sum a_{\alpha_i} x^{\alpha_i} \quad \text{dans le cas de multivarié, où } \alpha_i \text{ est un multi-indice}$$

— La famille harmonique.

$$f(x) = \sum a_i \cos(2\pi\omega_i x) + b_i \sin(2\pi\omega_i x)$$

— Une famille de splines.

$$f(x) = \sum b_i s_i(x) \quad \text{où } s_i \text{ est une base de splines}$$

— Une famille d'ondelettes.

— Une combinaison quelconque des familles précédentes.

La notion de "bon" ou de "meilleur" correspond en statistique à la notion de fonction de perte.

**Définition.** Une famille de fonctions de perte  $l_i : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}_+$  nous permet de quantifier la "perte" entre nos observations et la famille de valeurs  $(f(X_i))_{i \leq n}$  par la quantité  $l_i(Y_i, f(X_i))$ .

Plus généralement on définit la perte de notre régression par la quantité noté  $l(Y, f(X)) = \sum l_i(Y_i, f(X_i))$ .

On peut voir cette perte comme un éloignement des observations et des valeurs estimées (une quantification de l'erreur  $\varepsilon$ , par exemple dans le cas de la perte quadratique :  $l(x, y) = (x - y)^2$ ).

Le choix de la fonction de régression se fait par minimisation de cette fonction de perte et en ce sens la fonction de régression est la meilleure que l'on puisse trouver

$$\hat{f} \in \underset{f \in ?}{\operatorname{argmin}} \{l(Y, f(X))\},$$

([ndlr : le ? est là pour montrer que l'on n'a pas encore choisi l'espace fonctionnel]).

**Jargon.** Cet argument minimal est un estimateur de  $f$ .

Se donner une méthode de régression c'est choisir trois éléments : un espace fonctionnel, une perte  $l$  et un bruit (ou erreur)  $\varepsilon$ . Le choix de ces trois éléments est un problème inhérent au type de données dont l'on dispose et aux informations que l'on veut en tirer. Le choix de l'espace fonctionnel va naturellement influencer sur les performances de l'estimateur.

Notons un premier problème : on se donne la perte quadratique ( $l(y, x) = (y - x)^2$ ) et la famille de fonctions la plus large possible (disons les fonctions  $L^2$ ) alors la fonction nulle presque partout :

$$f(x) = \begin{cases} 0 & \text{si } x \notin \{x_i, i \leq n\} \\ y_i & \text{si } x = x_i \end{cases}.$$

a une perte nulle, donc c'est un argument minimal pour la perte quadratique. De plus dans cet espace il n'y a pas unicité de la solution. Le choix de l'espace fonctionnel ne doit pas se faire à la légère et les propriétés géométriques ne sont pas les seuls arguments en faveur de tel ou tel espace (comme le montre le cas de l'espace  $L^2$  qui est un espace de HILBERT mais ne présente aucun intérêt en régression).

**Exemples.** De modèles de régression

— Régression linéaire avec bruit gaussien

$$Y = X\beta + \varepsilon \quad \text{où } \varepsilon \text{ est un vecteur gaussien,}$$



$$\hat{\beta} \in \underset{\beta \in \mathbb{R}^d}{\operatorname{argmin}} \{ \|Y - X\beta\|^2 \}.$$

En dérivant la fonction de perte on obtient l'estimateur :  $\hat{\beta} = ({}^tXX)^{-1}XY$ , sous l'hypothèse que  ${}^tXX$  est inversible (ce qui est équivalent à demander que  $X$  soit de rang  $d$ , donc injective).

— Régression avec fonctions harmoniques et bruit gamma

$$Y_i = \sum \alpha_i \cos(\omega_i X_i) + \beta_i \sin(\omega_i X_i) + \varepsilon \quad \text{où } \varepsilon \text{ variables iid de loi gamma,}$$

réécrit sous forme matricielle :

$$Y = Z_\omega \xi + \varepsilon \quad \text{où } \varepsilon \text{ variables iid de loi gamma,}$$

$$\hat{\xi} \in \underset{\xi \in \mathbb{R}^{2d}}{\operatorname{argmin}} \{ \|Y - Z_\omega \xi\|^2 \}.$$

L'estimateur est de la même forme que précédemment.

— Régression linéaire du type LASSO (bruit quelconque)

$$Y = X\beta + \varepsilon_i \quad \text{où } E(\varepsilon_i) = 0 \text{ et } \operatorname{Var}(\varepsilon_i) = \sigma^2,$$

$$\hat{\beta}_\lambda^L \in \underset{\beta \in \mathbb{R}^d}{\operatorname{argmin}} \{ \|Y - X\beta\|^2 + \lambda \|\beta\|_1 \}.$$

Cette régression est très utilisée pour la grande dimension car parcimonieuse (pour un  $\lambda$  élevé j'aurai tendance à favoriser les vecteurs de norme  $\|\cdot\|_1$  faible [donc avec beaucoup de coefficients proche de 0] tandis qu'avec un  $\lambda$  proches de 0 je me rapprocherai de l'estimateur des moindres carrés).

— Régression linéaire du type Ridge (bruit quelconque)

$$Y = X\beta + \varepsilon_i \quad \text{où } E(\varepsilon_i) = 0 \text{ et } \operatorname{Var}(\varepsilon_i) = \sigma^2,$$

$$\hat{\beta}_\lambda^R \in \underset{\beta \in \mathbb{R}^d}{\operatorname{argmin}} \{ \|Y - X\beta\|^2 + \lambda \|\beta\|^2 \}.$$

Ce type de régression permet de pallier le problème d'inversibilité de la régression linéaire. En effet l'estimateur vaut  $({}^tXX + \lambda I)^{-1}XY$  et il est toujours bien défini pour  $\lambda$  non nul. De même que pour la régression LASSO, la norme de l'estimateur est pénalisée et favorise donc les éléments avec une faible norme. Dans le cas où la régression n'est plus linéaire mais sur un espace fonctionnel (on prend alors la norme associée), alors choisir une fonction de norme petite revient à choisir une fonction "plus lisse".

**Jargon.** Dans les deux derniers exemples nous appelons régularisation le fait d'introduire un terme supplémentaire dans la fonction de perte et  $\lambda$  le facteur de régularité

La régression linéaire par moindres carrés est très utilisée car simple à mettre en oeuvre. Cependant d'autres méthodes prennent des parts de plus en plus importantes dans la recherche et les applications (les régressions LASSO, Ridge, SVM et logistique). Elles proposent des alternatives intéressantes et moins évidentes que la régression linéaire avec moindres carrés.

Pour la modélisation de la consommation c'est une famille un peu particulière de régression que nous allons utiliser.

Le cas de la régression avec perte quadratique a été étudiée dans de nombreuses circonstances (famille de fonctions linéaire, projection sur des bases finies de fonctions, hölderiennes, splines, polynomiale, ... [Gasa]) et les propriétés de ces estimateurs sont connues (propriétés asymptotiques, bornes minimax, adaptativité de certaine classes, ... [Gasb], [RS12]).

Pour une revue des méthodes de régression voir aussi [Del].

Les méthodes de régression ont beaucoup évolué ces dernières années ; particulièrement au niveau de leur cadre d'application, de leur estimation effective et des résultats de convergence associés. Un nouveau point de vue est notamment apparu, complétant celui que nous avons utilisé pour présenter la régression. Ce point de vue (présenté en Annexe) est celui des RKHS (Reproducing Kernel HILBERT Space = espace de HILBERT à noyau reproduisant) qui, par l'introduction d'un "bon" espace de HILBERT (=notre espace fonctionnel), permet d'utiliser les propriétés d'algèbre sur ces espaces fonctionnels.

## 1.2 Validation croisée : Définitions

Pour cette partie nous utilisons principalement les cours PECCOT au collège de France de Sylvain ARLOT ([Arl11]), [AC10] et [CW79].

### 1.2.1 Introduction

Le travail de régression tel que présenté précédemment pose un problème conceptuel. Pour le moment nous avons proposé un modèle pour expliquer les observations à partir de certaines variables. Cependant notre but est d'être capable de prévoir ce qui arrivera, nous devons juger la qualité de notre modèle sur de nouvelles données. Pour ce faire il nous faut donc de nouvelles données qui nous permettrons d'effectuer des comparaisons.

Le principe de validation croisée repose sur la méthode suivante : dans un premier temps choisir un échantillon d'apprentissage qui permet l'estimation du modèle et ensuite (à partir de cette estimation) l'étude de la qualité du modèle à partir d'un échantillon de validation.

Plus formellement dans le cadre de la régression, on se donne  $D_n$ , une partition en deux sous ensembles  $I_1$  et  $I_2$  de  $\{1, \dots, n\}$ .

On note ensuite  $D_{I_1} = (Y_i, X_i)_{i \in I_1}$  et  $D_{I_2} = (Y_i, X_i)_{i \in I_2}$ , le premier ensemble  $D_{I_1}$  permettant de calculer notre estimateur, le second nous donnant ses performances sur un nouveau jeu de données.

On pose ainsi  $\hat{f}_{I_1}$  l'estimateur, et sa performance sur le nouveau jeu de données

$$Perf(\hat{f}_{I_1}) = \frac{1}{card(I_2)} \sum_{i \in I_2} l(Y_i, \hat{f}_{I_1}(X_i))$$

### 1.2.2 Validation croisée pour la sélection

Donnons nous plusieurs modèles  $(\mathcal{M}_p)$  (par exemple  $p$  peut correspondre au nombre de variables explicatives, ou à la taille de la base de projection) et  $(\hat{f}_p)_p$  leur estimateur. Notre objectif est de choisir le meilleur modèle parmi ces derniers (dans l'exemple cela reviendrait à choisir un nombre de variables explicatives).

Ainsi pour sélectionner un modèle nous choisissons plutôt de comparer leurs performances en validation croisée et de retenir le meilleur modèle.

Nous raffinons la méthode de validation croisée simple pour éviter une trop grande dépendance au choix de  $I_1$  et  $I_2$ . Soit  $(I_1^j, I_2^j)_{j \leq k}$  une suite de partition de  $\{1, \dots, n\}$ .  $I_1^j$  permettra d'estimer et  $I_2^j$  de calculer la performance de l'estimateur. Nous en faisons ensuite la moyenne :

$$\text{Perf}(\hat{f}_p) = \frac{1}{k} \sum_{j \leq k} \frac{1}{\text{card}(I_2^j)} \sum_{i \in I_2^j} l(Y_i, \hat{f}_{I_1^j}(X_i)),$$

$$\hat{p} \in \text{argmin} \left\{ \text{Perf}(\hat{f}_p) \right\}.$$

Il existe différentes méthodes de choix de cette partition.

**Exemples.** — *Leave-h-out.* On se donne une partition en deux éléments, un échantillon d'estimation de taille  $n - h$  et l'échantillon de calcul de performance est de taille  $h$ . Cette méthode est assez complexe à mettre en place puisqu'elle nécessite le calcul de  $\binom{n}{h}$  sous ensembles.

- *V-fold.* On résout le problème de complexité précédent en se donnant une partition de  $k$  éléments  $(I_j)_{j \leq k}$ . On choisit d'abord le premier élément pour l'estimation puis son complémentaire pour le calcul de performance, puis le second élément pour une nouvelle estimation et son complémentaire pour ... ainsi on obtient le calcul de performance

$$\text{Perf}(\hat{f}_p) = \frac{1}{k} \sum_{j \leq k} \frac{1}{n - \text{card}(I_j)} \sum_{i \notin I_j} l(Y_i, \hat{f}_{I_j}(X_i))$$

- Choisir les partitions de manière aléatoire (uniforme, méthodes de Monte-Carlo, ...)

Les techniques que nous venons d'énumérer pour le choix de modèles doivent aussi s'accompagner de résultats théoriques pour justifier leur utilisation en pratique. Ces résultats sont de deux ordres

- Asymptotique, en assurant que notre méthode de sélection donne une erreur qui tend vers l'erreur du meilleur des modèles

$$\hat{p} \in \text{argmin} \left\{ \text{Perf}(\hat{f}_p) \right\} \implies \lim_{n \rightarrow \infty} \frac{E(l(Y, \hat{f}_{\hat{p}}))}{\min_p E(l(Y, \hat{f}_p))} = 1.$$

- Non asymptotique (inégalité oracle), en assurant que le risque du modèle que nous avons choisi reste de l'ordre de grandeur du meilleur des modèles

$$\exists C > 0, \exists R_n \in \mathbb{R}, \hat{p} \in \text{argmin}_p \left\{ \text{Perf}(\hat{f}_p) \right\} \implies l(Y, \hat{f}_{\hat{p}}) \leq C \inf_p \left\{ l(Y, \hat{f}_p) + R_n \right\}.$$

Le premier résultat est essentiel pour justifier l'utilisation d'une de ces méthodes, la seconde est souvent plus délicate à obtenir, mais non moins intéressante. Pour plus de précision voir [AC10].

### 1.2.3 Validation croisée généralisée

La validation croisée généralisée (GCV en anglais) est un critère spécifique pour les méthodes d'estimation linéaire (i.e : l'estimateur s'écrit comme le produit des observations et d'une certaine matrice). Nous nous mettons dans le cas d'une régression Ridge :  $X \hat{\beta}_\lambda^R = M_\lambda Y$  ( $M_\lambda = (X^T X + \lambda I)^{-1} X^T$ ).

L'erreur d'approximation du modèle est donnée par

$$E(\|X\beta - X\hat{\beta}_\lambda^R\|^2) = E(\|X\beta - M_\lambda Y\|^2) = \|(I - M_\lambda)X\beta\|^2 + \sigma^2 \text{tr}(M_\lambda^t M_\lambda),$$

et l'erreur d'estimation par

$$E(\|Y - X\hat{\beta}_\lambda^R\|^2) = E(\|(I - M_\lambda)Y\|^2) = \|(I - M_\lambda)X\beta\|^2 + \sigma^2 n - 2\sigma^2 \text{tr}(M_\lambda) + \sigma^2 \text{tr}(M_\lambda^t M_\lambda).$$

Nous aimerions réduire l'erreur d'approximation mais ne connaissant pas la vraie valeur de  $\beta$  nous ne pouvons que travailler avec la seconde erreur. Cependant nous avons entre les deux erreurs le liens suivant

$$E(\|X\beta - M_\lambda Y\|^2) = E(\|(I - M_\lambda)Y\|^2) - \sigma^2 n + 2\sigma^2 \text{tr}(M_\lambda) - \sigma^2 \text{tr}(M_\lambda^t M_\lambda) + \sigma^2 \text{tr}(M_\lambda^t M_\lambda).$$

Il existe donc un critère de sélection naturel (celui de MALLOWS), en choisissant le modèle qui minimise l'erreur d'approximation

$$\hat{\lambda} \in \underset{\lambda > 0}{\text{argmin}} \left\{ \|(I - M_\lambda)Y\|^2 + 2\sigma^2 \text{tr}(M_\lambda) \right\}.$$

La connaissance de la valeur de  $\sigma$  n'est souvent pas acquise, il est donc aussi nécessaire de l'estimer

$$\hat{\sigma}^2 = \frac{1}{n-p} \|Y - \text{Proj}_X(Y)\|^2,$$

(où  $\text{Proj}_X(Y) = ({}^t X X)^{-1} {}^t X Y$ ).

Finalement le critère que nous voudrions nous donner est

$$\hat{\lambda} \in \underset{\lambda > 0}{\text{argmin}} \left\{ \|(I - M_\lambda)Y\|^2 + \frac{2}{n-p} \text{tr}(M_\lambda) \|(I - ({}^t X X)^{-1} X)Y\| \right\}.$$

Cela nous oblige à calculer de nouvelles quantités, pour éviter cela nous proposons un autre estimateur de  $\sigma$

$$\hat{\sigma}^2 = \frac{1}{n} \|Y - M_\lambda Y\|^2,$$

qui a l'avantage de ne pas faire apparaître de nouvelle quantité et de simplifier le critère de sélection

$$\hat{\lambda} \in \underset{\lambda > 0}{\text{argmin}} \left\{ \|(I - M_\lambda)Y\|^2 \left(1 + \frac{2}{n} \text{tr}(M_\lambda)\right) \right\},$$

de plus pour  $n$  suffisamment grand nous avons

$$1 + \frac{2}{n} \text{tr}(M_\lambda) \approx \left(1 - \frac{1}{n} \text{tr}(M_\lambda)\right)^{-2}.$$

Tout cela nous amène à proposer un nouveau critère, celui de la validation croisée généralisée

$$\hat{\lambda}_{GCV} \in \underset{\lambda > 0}{\text{argmin}} \frac{\|(I - M_\lambda)Y\|^2}{(1 - \text{tr}(M_\lambda/n))^2}.$$

**Remarque.** Ce critère de sélection a deux atouts majeur : il ne nécessite pas l'estimation du paramètre  $\sigma$  souvent inconnu, nous ne sommes pas obligé de découper notre famille de données en plusieurs pour faire de la validation croisée (ce qui est couteux en temps et complexe à mettre en oeuvre)

Pour une présentation plus théorique et rigoureuse de la GCV voir [CW79] (cet article justifie aussi l'appellation validation croisée généralisée). Pour des résultats sur l'efficacité asymptotique et non asymptotique de ce critère voir [CW79], [CG05] et [CG06]

## 1.3 Les modèles linéaires généralisés

Nous décrivons maintenant un type de régression légèrement différent du cadre général présenté dans la première partie.

Dans le cadre de la régression on se donne une famille de modèles de régression que l'on appelle les modèles linéaires généralisés (GLM : Generalized Linear Model).

### 1.3.1 Présentation du modèle

Supposons les observations  $(Y_i)_{i \leq n}$  indépendantes et suivants des distributions de la famille exponentielle (i.e :  $Y_i$  a une densité par rapport à la mesure de LEBESGUE,  $f_{\theta_i}(y) = \exp(\{y\theta_i - b_i(\theta_i)\})/a_i(\phi) + c_i(y, \phi)$ ). Nous voulons résoudre le problème

$$E(Y_i) = X_i\beta.$$

**Remarque.** Pour les lois faisant partie de la famille exponentielle et vérifiant la propriété

$$\forall a \in \mathbb{R}^d, \varepsilon \stackrel{\text{loi}}{\sim} L(E(\varepsilon) = 0, \text{Var}(\varepsilon) = \sigma^2) \implies a + \varepsilon \stackrel{\text{loi}}{\sim} L(a, \text{Var}(\varepsilon) = \sigma^2)$$

alors on peut modéliser (avec  $\varepsilon \sim L$ ) l'équation par

$$Y_i = X_i\beta + \varepsilon.$$

En réalité les modèles GLMs ont une représentation encore plus générale avec l'introduction d'une fonction de lien  $g$  et l'équation devient

$$g(E(Y_i)) = X_i\beta.$$

Pour plus d'information sur les modèles GLMs nous renvoyons au livre [Woo06] ou [Del].

**Question.** La forme de la formule de régression est différente des cas que nous avons pu voir précédemment, à quoi peut ressembler la fonction de perte pour une telle famille de régression ?

Nous avons vu la fonction de perte comme une manière de quantifier l'importance de l'erreur  $\varepsilon$ , qui n'est plus présente ici (ou bien seulement de manière implicite à travers la condition d'appartenance à la famille exponentielle). Pour choisir nos paramètres nous décidons finalement d'utiliser un critère de maximisation de la vraisemblance.

### 1.3.2 Estimation du modèle GLM

Nous allons estimer le modèle par maximum de vraisemblance. Nous faisons d'abord quelques rappels sur la vraisemblance et sur l'interprétation empirique de ce que représente son maximum.

Soit  $X_i$  une variable aléatoire réelle admettant une densité  $f_i$  par rapport à la mesure de LEBESGUE, alors on appelle  $f_i(X_i(\omega))$  la vraisemblance de cette variable [au point  $\omega$ ] (on peut aussi s'intéresser à la log-vraisemblance :  $\log(f(X(\omega)))$ ). Pour le cas d'une suite de variables aléatoires indépendantes la vraisemblance est  $\prod_i f_i(X_i(\omega))$  (on oubliera le  $\omega$  dans la suite).

Donnons nous une suite  $(X_i)_{i \geq 0}$  indépendantes de même loi normale centrée en  $\tilde{\theta}$  et  $f_{\tilde{\theta}}$  la densité d'une loi normale centrée en  $\theta$ . Les tirages  $(X_i(\omega))_{i \geq 0}$  auront tendances à se répartir autour de  $\tilde{\theta}$ .

Prenons  $\theta_0$  éloigné de  $\tilde{\theta}$ . la quantité  $f_{\theta_0}(X_i)$  sera plus faible que  $f_{\tilde{\theta}}(X_i)$ . L'idée est de prendre

$$\theta_n \in \underset{\theta \in \mathbb{R}}{\operatorname{argmax}} \left\{ \prod_{i \leq n} f_{\theta}(X_i) \right\},$$

comme estimateur de  $\tilde{\theta}$  (dans le cas gaussien  $\theta_n = \frac{1}{n} \sum_{i \leq n} X_i$ ).

Dans le cas général et sous de bonnes conditions nous espérons avoir le résultat suivant : Si  $(X_i)_{i \geq 0}$  sont indépendants et suivent une loi à densité  $f_{\tilde{\theta}}$  alors

$$\theta_n \in \underset{\theta \in \mathbb{R}}{\operatorname{argmax}} \left\{ \prod_{i \leq n} f_{\theta}(X_i) \right\} \implies \theta_n \xrightarrow{n \rightarrow +\infty} \tilde{\theta}$$

La famille des modèles exponentiels est dans ce cas de figure (voir [Gasa]).

Pour les modèles GLMs nous estimons  $\beta$  en maximisant la vraisemblance. Dans le cadre des familles exponentielles nous avons les relations suivantes :  $E(Y) = b'(\theta)$  et  $\operatorname{Var}(Y) = -b''(\theta)/a(\varphi)$  (supposons la fonction  $a$  linéaire [ $a(\varphi) = \varphi/\omega$ ], ce qui est souvent le cas en pratique).

Nous calculons le maximum de vraisemblance (en réalité de la log-vraisemblance). Pour cela nous dérivons dans chacune des directions et demandons à annuler la dérivée (les paramètres ne sont pas nécessairement les mêmes pour chaque observation)

$$l(\beta) = \sum_{i=1}^n \omega_i (y_i \theta_i - b_i(\theta_i)) / \varphi + c_i(\varphi, y_i),$$

$$\frac{\partial l}{\partial \beta_j} = \frac{1}{\varphi} \sum_{i=1}^n \omega_i \left( y_i \frac{\partial \theta_i}{\partial \beta_j} - \frac{\partial \theta_i}{\partial \beta_j} b'_i(\theta_i) \right).$$

Ecrivons cette dernière égalité avec des quantités que l'on connaît (faire disparaître  $\theta_i$  quitte à faire apparaître de nouvelles variables du modèle). Pour cela on nomme  $\mu_i = E(Y_i) = b'_i(\theta)$  et donc que  $\frac{d\mu_i}{d\theta_i} = b''_i(\theta)$ . On a la formule de dérivation suivante

$$\frac{\partial \theta_i}{\partial \beta_j} = \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \beta_j},$$

et

$$\frac{\partial \mu_i}{\partial \theta_i} = b''_i(\theta_i) \implies \frac{\partial \theta_i}{\partial \mu_i} = \frac{1}{b''_i(\theta_i)}.$$

Remplaçons ces quantités dans l'expression de la vraisemblance

$$\frac{\partial l}{\partial \beta_j} = \frac{1}{\varphi} \sum_{i=1}^n \frac{y_i - \mu_i}{b''_i(\theta_i)/\omega_i} \frac{\partial \mu_i}{\partial \beta_j}.$$

Finalement trouvons  $\beta$  tel que

$$\forall j, \sum_{i=1}^n \frac{y_i - \mu_i}{V(\mu_i)} \frac{\partial \mu_i}{\partial \beta_j} = 0.$$

En supposant la fonction  $V$  indépendante de  $\beta$  nous avons équivalence avec la minimisation de

$$S = \sum_{i=1}^n \frac{(y_i - \mu_i)^2}{V(\mu_i)}.$$

L'estimation effective se fait en minimisant récursivement cette quantité (on peut faire un parallèle avec l'algorithme Expectation-Minimization). Pour cela on suit le schéma suivant

- On part de valeurs initiales
- On optimise selon  $\beta$  la quantité  $S_0$ , ce qui donne  $\beta_1$
- On calcule les nouvelles valeurs des éléments ....
- On échange les valeurs précédentes avec les nouvelles calculées.

Il existe des méthodes pour faire converger plus rapidement ce type d'algorithme (telle que le développement de TAYLOR) voir [Woo06] ou [Del] pour plus de précision dans le cas des GLMs

## 1.4 Régression Ridge : présentation & plus

Avant de présenter le modèle développé par le groupe R39 pour la prévision de consommation électrique nous avons besoin de présenter le principe de la régression Ridge avec précision.

### 1.4.1 Introduction

Le cadre de régression Ridge linéaire ( $\lambda > 0$ )

$$Y = X\beta + \varepsilon,$$

$$\hat{\beta}_\lambda^R \in \underset{\beta \in \mathbb{R}^d}{\operatorname{argmin}} \{ \|Y - X\beta\|^2 + \lambda \|\beta\|^2 \}.$$

En dérivant la fonction de perte selon  $\beta$  et en annulant la dérivée (la fonction de perte est convexe) alors on obtient comme estimateur

$$\hat{\beta}_\lambda^R = ({}^tXX + \lambda I)^{-1} {}^tXY$$

Revenons sur l'introduction de cette perte. Nous avons l'estimateur des moindres carrés de la régression linéaire  $({}^tXX)^{-1} {}^tXY$  qui est défini si  ${}^tXX$  est inversible.

**Proposition.** *La matrice  $X \in M_{n \times p}$  est de rang égal à  $p$  si et seulement si  ${}^tXX$  est inversible (bijective)*

*Démonstration.* la matrice  $X^tX$  est symétrique positive et est inversible si et seulement si elle n'admet pas de valeurs propres nulles. On vérifie que cette dernière condition est équivalente à ce que  $X$  soit de rang  $p$ . En effet  $X$  est de rang  $p$  si et seulement si son noyau est de dimension nulle (théorème du rang) si et seulement si  $\langle X\beta, X\beta \rangle = 0 \iff \beta \in \operatorname{Ker}(X)$  si et seulement si  $(\beta^t X^t X \beta = 0 \iff \beta = 0)$ . On en déduit que  $X^tX$  est définie positive, et donc qu'elle est inversible, si et seulement si  $X$  est de rang  $p$   $\square$

**Remarque.** *Dans le cas où cette hypothèse n'est pas vérifiée nous n'avons plus l'unicité de la solution*

Nous aimerions pallier à ce problème d'inversibilité, pour cela nous remarquons que  ${}^tXX$  est une matrice symétrique positive et par conséquent pour tout  $\lambda$  supérieur strictement à zéro nous avons que  ${}^tXX + \lambda I$  est symétrique définie positive (donc inversible). Ainsi l'introduction de la régression Ridge donne un cadre où l'estimateur de régression linéaire est toujours bien défini (il existe et est unique).

Ce n'est pas le seul intérêt de cette régularisation : elle permet d'avoir des propriétés de décorrélation des différentes variables, théoriquement elle s'affranchit de la dépendance linéaire en introduisant une matrice définie positive ; en pratique on constate que lorsque deux variables

sont très "corrélés" la régression Ridge est plus efficace pour l'estimation des coefficients de régression.

La régression Ridge ne nous donne une famille d'estimateurs  $(\hat{\beta}_\lambda^R)_{\lambda>0}$  (dans le cadre effectif nous prendrons  $\lambda$  sur une grille). On est donc dans un cadre de sélection d'estimateur, il nous reste donc à nous donner un critère de sélection : la Validation Croisée Généralisée.

En posant  $A_\lambda = X({}^tXX + \lambda I)^{-1}{}^tX$

$$crit_{GCV}(X, Y, \lambda) = \frac{\|Y - A_\lambda Y\|^2}{(n - tr(A_\lambda))^2},$$

$$\hat{\lambda} \in \underset{\lambda>0}{argmin} \{crit_{GCV}(X, Y, \lambda)\}.$$

Pour justifier de l'utilisation de ce critère nous avons le résultat asymptotique

**Théorème** ([CW79]). *Dans le cas linéaire, nous avons pour  $\hat{\lambda}$  le choix par validation croisée généralisée.*

$$\lim_{n \rightarrow \infty} \frac{E\|Y - X\hat{\beta}_{\hat{\lambda}}^R\|^2}{\min_{\lambda>0} E\|Y - X\hat{\beta}_\lambda\|^2} = 1.$$

En fait ce résultat est démontré pour une classe de fonction plus générale que celle linéaire.

## 1.4.2 & plus

Le cadre de la régression Ridge s'étend aisément avec des espaces fonctionnels plus complexes.

Soit  $\mathcal{F}$  un espace fonctionnel munit d'une norme issu d'un produit scalaire alors on pose la régression Ridge

$$Y = f(X) + \varepsilon,$$

$$\hat{f}_\lambda \in \underset{f \in \mathcal{F}}{argmin} \{ \|Y - f(X)\|^2 + \lambda \|f\|_{\mathcal{F}}^2 \}.$$

Dans ce cas le facteur de régularisation  $\lambda$  peut être vu comme une pénalisation de la norme, c'est à dire que l'on handicape les fonctions qui ont une norme importante.

Dans le cas particulier où la fonction  $f$  peut s'écrire comme combinaison linéaire de fonction de base ( $f(x) = \sum_i \beta_i a_i(x)$ ) alors

$$\hat{f}_\lambda(x) = \sum_i \hat{\beta}_\lambda a_i(x),$$

$$\hat{\beta}_\lambda \in \underset{\beta}{argmin} \{ \|Y - A\beta\|^2 + \lambda {}^t\beta S \beta \},$$

où  $A = (a_j(X_i))_{i,j}$  et  $S = (\langle a_i, a_j \rangle_{\mathcal{F}})_{i,j}$ . C'est par exemple le cas pour les bases de splines, ou les projections sur des sous espaces de dimension finies. Pour la sélection du facteur de régularisation on peut encore le choisir par validation croisée généralisée. L'article [CW79] donne le théorème

**Théorème.** *Soit  $W_2 = \{f; f, f' \text{ absolument continue, } \int f''(t)^2 dt < \infty\}$ . Alors la validation croisée généralisée est asymptotiquement consistante pour l'espace de fonction  $W_2$*

$$\lim_{n \rightarrow \infty} \frac{E\|Y - \hat{f}_{\hat{\lambda}}(X)\|^2}{\min_{\lambda>0} E\|Y - \hat{f}_\lambda(X)\|^2} = 1$$



Voir l'Annexe sur les RKHS pour plus de précision sur l'espace  $W_2$ .

Dans le cas particulier où la famille de fonction est polynômiale ou à base de splines alors on a de plus les résultats non asymptotiques (de type inégalité oracle) suivants

**Théorème** ([CG05] et [CG06]). *Uniformément en  $f$  et en  $\gamma \in ]0, 1[$  on a*

$$R(\hat{f}_{\hat{\alpha}}, f) \leq \frac{1}{1-\gamma} \inf_{\alpha} \left\{ R(\hat{f}_{\alpha}, f) + \frac{C\sigma^2}{n\gamma} \right\}$$

**Théorème.** *Uniformément en  $f$  et en  $\gamma \in ]0, 1[$  on a*

$$R(\hat{f}_{\hat{\alpha}_G}, f) \leq \frac{1}{1-\gamma} \left\{ \left(1 - \frac{1}{\sqrt{n}}\right)^{-2} \inf_{\alpha, \text{tr}(S_{\alpha}) \leq \sqrt{n}} R(\hat{f}_{\alpha}, f) + \frac{C\sigma^2}{n\gamma} \right\}$$

## 1.5 Les modèles additifs généralisés

Nous avons vu les GLMs comme des régressions linéaires pour lesquelles on se donne un bruit d'une forme différente.

Il en est de même pour les modèles additifs généralisés : partant d'une régression Ridge avec une partie linéaire, une partie fonctionnelle (avec pour famille de fonctions les splines) et un bruit gaussien ; on peut généraliser ce modèle en prenant un bruit différent (pour cela on adopte le même formalisme que pour les modèles GLMs).

### 1.5.1 Présentation

Enfin nous nous intéressons à une dernière famille de modèle de régressions, qui représente un amalgame des deux familles précédentes (et c'est avec cette famille que nous allons travailler et simuler la consommation française) : ce sont les modèles additifs généralisés (GAM : Generalized Additive Model).

Dans le cas des ces modèles nous nous proposons de modéliser nos observations avec une partie linéaire (qu'on pourra aussi appeler paramétrique) et une partie non linéaire (ou non paramétrique) et l'espace fonctionnel sur lequel nous chercherons les fonctions de la partie non paramétrique est l'espace des fonctions splines. Ainsi

$$g(E(Y_i)) = X_i\beta + f(X_i),$$

où  $Y_i$  suit toujours une distribution de famille exponentielle et  $f$  est une fonction spline (le choix de cette famille de spline est aussi un sujet de discussion : voir Annexe sur les RKHS pour des précisions sur ces bases). Dans le cas gaussien on peut réécrire la formule de régression de la manière suivante (avec l'expression de l'estimateur)

$$Y = X\beta + f(X) + \varepsilon,$$

$$(\hat{\beta}, \hat{f}) \in \underset{\beta \in \mathbb{R}^d, f \in \mathcal{F}}{\text{argmin}} \left\{ \|Y - X\beta - f(X)\| + \lambda \int f''(x)^2 dx \right\}.$$

On a donc un modèle de régression qui mélange une partie linéaire et une partie fonctionnel qui se sert de la perte Ridge pour pénaliser la partie non paramétrique.

### 1.5.2 Estimation effective du modèle

Il y a une analogie claire entre les modèles GLMs et GAMs. Le lien entre la régression linéaire par moindres carrés et les GLMs est la même qu'entre la régression splines avec pénalité Ridge et les GAMs.

Par conséquent l'estimation des paramètres du modèles GAM va se faire de manière analogue à GLM (i.e : par maximisation de la fonction de vraisemblance). Les calculs ayant déjà été fait pour le cas GLM, il reste à adapter cette forme d'estimation pour le GAM.

Nous pouvons donc donner le même schéma d'estimation que précédemment

- On part de valeurs initiales
- On optimise selon  $\beta$  la quantité  $S_0 + \lambda^t \beta S \beta$ , ce qui donne  $\beta_1$
- On calcule les nouvelles valeurs des éléments ....
- On échange les valeurs précédentes avec les nouvelles calculées.

Nous choisissons la base de splines suivant le type d'observations que nous avons. Il faut néanmoins avoir en tête le résultat suivant

**Proposition.** *Sur  $W_2$  la famille des splines cubique minimise la quantité  $\int f''(t)^2 dt$  (voir [Woo06])*

Les différentes bases de splines sont en partie présentées en Annexe. Choisir la taille de la base est aussi un problème, pour cela il est nécessaire de choisir le nombre de noeuds des splines (qui sera aussi la taille de la base). Nous avons beaucoup plus d'observations que nous allons choisir de noeuds, par conséquent les noeuds seront répartis de manière équi-quantile.

# Chapitre 2

## Applications

Les conclusions que nous tirerons restent empiriques et spécifiques à nos jeux de données, toute généralisation étant non triviale.

### 2.1 Remarques, informations sur les données et objectifs

Commençons ce chapitre d'application par l'étude des données qui vont nous permettre de travailler.

#### 2.1.1 Vocabulaire

Dans un premier temps il faut rappeler que le marché de l'énergie s'est ouvert à la concurrence depuis les années 2000. Nous devons donc séparer en deux groupes la population française : les clients EDF et ceux qui ne le sont plus.

**Jargon.** *Le portefeuille clients EDF correspond à l'ensemble des clients EDF*

De plus au sein du portefeuille EDF nous pouvons décomposer en plusieurs autres groupes les consommateurs : les SUP 7, les 32 000, les profilés

**Jargon.** *L'agrégat des SUP 7 est composé de très grosses entreprises dont chacune consomme plus de 7GWh/an (ex : industrie)*

**Jargon.** *L'agrégat des 32 000 est composé de petites et moyennes entreprises (environ 32 000 au total)*

**Jargon.** *L'agrégat des profilés est composé des résidentiels et autres petits clients dont les consommations ne sont pas mesurés à partir des courbes de charge mais par index.*

Cette séparation en sous-groupes est nécessaire pour modéliser avec précision les comportements des consommateurs. Les industries ne consomment pas de la même manière que les particuliers (les chaînes de production doivent fonctionner quelle que soient la température, le vent, la couverture nuageuse tandis qu'un particulier va consommer beaucoup plus en fonction de ces aléas).

**Jargon.** *La courbe de charge d'une entreprise correspond à sa consommation électrique au cours de la journée (au pas demi-horaire)*

La connaissance des courbes de charge de chacun des clients EDF permet une prévision plus fine de la consommation.

Avoir de telles données est un enjeu industriel qui a le défaut d'être sensible aux données manquantes.

Pour l'agrégat des 32 000, EDF a accès aux courbes de charge

**Jargon.** *On dit que les données sont empilées lorsque la consommation d'un agrégat est calculée en sommant les courbes de charge de chacun des clients.*

**Jargon.** *On dit que les données sont dépilées lorsque la consommation d'un agrégat est calculée en faisant la différence entre la consommation totale et celle des autres agrégats.*

L'intérêt des données dépilées est de ne plus avoir de données manquantes au niveau des courbes de charge, cependant cela donne moins de souplesse et les erreurs sur les mesures de consommation des autres agrégats se retrouvent dans cet agrégat.

EDF a développé des tarifs spéciaux EJP (= Effacement de Jour de Pointe), dont l'objectif est de faire payer moins cher aux consommateurs leur électricité mais de décupler les prix 22 jours entre le 1<sup>er</sup> novembre et le 1<sup>er</sup> mars. Ceci permet à EDF de faire diminuer la consommation lors de ces 22 journées d'hiver (la période de consommation la plus tendue).

### 2.1.2 Problématique et objectifs

Pour favoriser la concurrence, au début de l'année 2016, le marché de l'énergie verra la fin de certains tarifs réglementés (jaune et vert). Par conséquent EDF prévoit d'importantes variations dans le nombre de ses clients (particulièrement pour le secteur des 32 000). Or les modèles actuels ont besoin, pour produire une estimation fiable, d'avoir des observations qui ont un comportement sensiblement équivalent en fonction des variables explicatives. La variation du portefeuille EDF entre en contradiction avec cette hypothèse de stabilité.

Dans la suite nous nous donnons le cas d'étude suivant (voir Annexe pour des généralisations de ce cas d'étude) : nous possédons les données de consommations du 1<sup>er</sup> janvier 2009 au 31 décembre 2012, le 1<sup>er</sup> janvier 2012 nous simulons une perte 10% de la consommation des 32 000.

Une mise à jour est nécessaire pour prendre en compte cette évolution. Cependant les données reflétant les consommations réelles (celles après la perte de clients) sont peu nombreuses.

**Question.** *Comment faire évoluer le modèle GAM pour avoir une estimation fiable malgré des variations de périmètre clients ?*

L'idée est d'ajouter un paramètre d'oubli dans les modèles pour donner plus d'importance aux données récentes.

Dans le cas d'un modèle GAM avec bruit gaussien (= régression Ridge avec splines), pour jouer sur l'importance des observations en fonction de l'époque où elles se produisent, nous avons une possibilité : changer la fonction de perte de notre régression

$$Y = X\beta + f(X) + \varepsilon,$$

pour une régression Ridge elle ressemble à

$$(\hat{\beta}, \hat{f})_{\lambda} \in \underset{\beta \in \mathbb{R}^d, f \in \mathcal{F}}{\operatorname{argmin}} \left\{ \|Y - X\beta - f(X)\|^2 + \lambda \int f''(x)^2 dx \right\},$$

si on se donne une famille de poids  $(\omega_i)_{i \leq n}$  positive décroissante on choisit la nouvelle fonction de perte

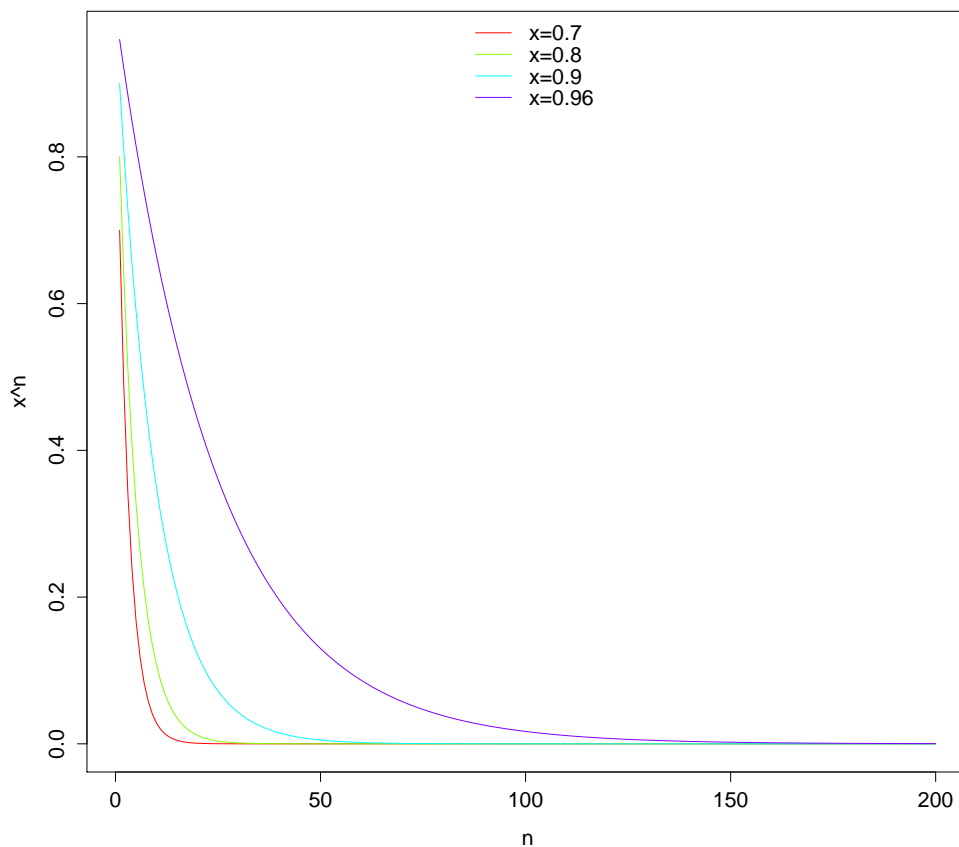
$$(\hat{\beta}, \hat{f})_{\lambda, (\omega_i)} \in \underset{\beta \in \mathbb{R}^d, f \in \mathcal{F}}{\operatorname{argmin}} \left\{ \sum_{i \leq n} \omega_i (Y_i - X_i \beta - f(X_i))^2 + \lambda \int f''(x)^2 dx \right\}.$$

**Question.** *En quoi cette nouvelle fonction de perte répond à notre problème ?*

Un poids  $\omega_i$  important (proportionnellement aux autres) "obligera" à choisir les paramètres de régression de manière à avoir une faible différence quadratique à l'observation  $i$ . Dans le même temps un poids  $\omega_j$  faible permettra d'avoir une différence quadratique importante sans que cela n'augmente de beaucoup la fonction de perte.

Ce nouveau modèle permet plus de souplesse dans l'estimation du modèle sans ajouter plus de complexité.

Dans la suite nous choisissons un modèle où les poids décroissent de manière exponentielle : soit  $\omega < 1$  on pose  $\omega_i = \omega^{n-i}$  (pour une pondération différente voir l'Annexe).



Le graphique précédent nous permet de faire une première remarque sur le facteur d'oubli. A partir d'un certain temps ( $n = 20$  pour  $x = 0.7$ ,  $n > 100$  pour  $x = 0.98$ ) les poids associés aux

données plus éloignées restent sensiblement les mêmes. Cela signifie que les données d'importances (celles qui permettent l'estimation du modèle) sont celles qui précèdent cette valeur (pour  $x = 0.7$  cela fait très peu de données d'importance). On a dès maintenant l'intuition que pour un facteur d'oubli faible (éloigné de 1) le modèle n'aura pas suffisamment d'information pour être estimé et donc qu'il sera mauvais en prévision, de même un facteur d'oubli élevé (proche de 1) ne va pas permettre au modèle de faire de différence entre les données récentes et anciennes et le modèle va prévoir sans prendre en compte la perte de clients. Le choix du facteur d'oubli est entre ces deux extrêmes.

### 2.1.3 Les variables explicatives

Nous sommes dans le cadre de la prévision par agrégat (ici celui des 32 000).

L'équation GAM des 32 000 tient compte de :

- La température : calculée comme moyenne pondérée de 32 villes françaises (on utilise une température pour toute la France),
- La position du jour dans l'année (*DayPosan*) : comprise entre 0 et 1 elle est calculée comme le rapport entre le nombre de jours écoulés depuis le début de l'année et le nombre de jours dans l'année. Elle doit modéliser les changements de comportement de consommation au cours de l'année (été, hiver, ruptures dues aux vacances),
- Le niveau EJP : calculé en fonction de la région de France (Nord, Sud, Ouest, PACA) et de la présence d'un jour EJP dans l'une de ces régions,
- Le type de jour (*DayType*) : différencie cinq types de jour (dimanche, lundi, vendredi, samedi, mardi-mercredi-jeudi),
- Le type de jour férié (*JF*) : différencie les veilles de jours fériés, les jours fériés, leur lendemain des jours normaux,
- La rupture (*Rupture*) : symbolise la rupture dans la consommation lors du mois d'août et des vacances de décembre,
- La tendance : représente la croissance économique.

La variable tendance pose certains problèmes de modélisation. Cette variable doit représenter une évolution de la consommation électrique inhérent à notre mode de consommation (= la part de (dé)croissance économique). Dans notre modèle on la considère comme une droite affine dont la régression doit estimer les paramètres. Cette croissance n'est pas toujours identique, nous aimerions représenter cet élément comme affine par morceaux (et définir les points de changement de tendance). L'étude d'Audrey PICHAVANT a montré que cette modélisation de la tendance donne de bonnes performances.

On modélise la consommation des 32 000 par l'équation suivant

$$\begin{aligned}
 Y = & s_1(\textit{Temperature}) + s_2(\textit{DayPosan}, by = \textit{DayType2}) \\
 & + \textit{NiveauEJP} + \textit{tendance} \\
 & + \textit{DayType5} : \textit{JF} : \textit{Rupture} + \varepsilon
 \end{aligned}$$

où

- $s_1$  et  $s_2$  sont des splines (on doit choisir une base de splines, voir la partie sur les modèles GAM et l'Annexe sur les RKHS),
- $DayType2$  vaut 1 pour un jour de semaine et 0 pour un jour de week-end,
- $DayType5$  forme la partition : 1(Dimanche), 2 (Lundi), 3 (Mardi Mercredi-Jeudi), 4 (Vendredi), 5 (Samedi),
- La notation *by* signifie que l'on estime deux splines pour la variable  $DayPosan$ , l'une pour la semaine, la seconde pour la fin de semaine,
- La notation  $DayType5 : JF : Rupture$  signifie que l'on croise les effets des différentes variables.

## 2.1.4 Estimation des modèles

### Les données

Nous travaillons avec les données 32 000 empilées entre le 1<sup>er</sup> janvier 2009 et le 31 décembre 2012. Nous construisons, à partir des courbes de charge, un échantillon d'entreprises (que nous appellerons "contestables") qui représentera une famille de clients susceptibles de partir à la concurrence.

Nous faisons le choix de modéliser le départ à la concurrence par une perte dans la consommation le 1<sup>er</sup> janvier 2012.

### Estimation d'un modèle de référence

Nous commençons par estimer un modèle GAM, que nous appellerons le modèle de référence, sur la période du 1<sup>er</sup> janvier 2009 au 31 décembre 2011. Dans le cas où il n'y aurait pas eu de départ de clients, ce modèle aurait servi pour la prévision de consommation pour l'année 2012.

Sous l'hypothèse qu'il n'y ait pas eu de perte de clients nous aurions les performances suivantes :

	RMSE	MAPE
Estimation (2007-2011)	110	1.7%
Prévision (2012)	158	2.8%

Ensuite avec la prise en compte de la perte de 10% de la consommation des 32 000

	RMSE	MAPE
Estimation (2007-2011)	110	1.7%
Prévision (2012)	540	13.1%

### Mise à jour du modèle de référence

Après le 1<sup>er</sup> janvier 2012 et la perte de clients, il faut mettre à jour le modèle de référence pour prendre en compte la variation du portefeuille EDF. L'arrivée des nouvelles données se fait chaque semaine le samedi (le samedi 14 juillet nous recevons les données de la semaine allant du samedi 1<sup>er</sup> juillet au vendredi 7 juillet : il y a donc une semaine de latence pour laquelle il n'y a pas de données de consommation) et permet une prévision pour la semaine qui suit.

### Première méthode : réestimation d'un modèle GAM avec facteur d'oubli

La mise à jour peut se faire en réestimant un modèle GAM sur l'historique de 2009-2011 et des consommations reçues lors des semaines suivant la rupture

$$Y = X\beta + f(X) + \varepsilon,$$

$$(\hat{\beta}, \hat{f})_{\lambda, (\omega_i)} \in \underset{\beta \in \mathbb{R}^d, f \in \mathcal{F}}{\operatorname{argmin}} \left\{ \sum_{i \leq n} \omega_i (Y_i - X_i \beta - f(X_i))^2 + \lambda \int f''(x)^2 dx \right\}.$$

### Seconde méthode : régression linéaire sur les effets avec facteur d'oubli

Nous proposons un second modèle pour la mise à jour de notre modèle de référence qui consiste à garder les effets estimés par le modèle de référence et de faire une régression linéaire avec facteur d'oubli. Formellement, si  $(\hat{\beta}, \hat{f})$  sont les paramètres estimés par le modèle de référence, alors on résout le problème de régression linéaire :

$$Y_i = \alpha_i X_i \hat{\beta} + \tilde{\alpha}_i \hat{f}(X_i) + \varepsilon,$$

$$(\hat{\alpha}, \hat{\tilde{\alpha}}) \in \underset{\alpha \in \mathbb{R}^n, \tilde{\alpha} \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \sum_{i \leq n} \omega_i (Y_i - \alpha_i X_i \hat{\beta} - \tilde{\alpha}_i \hat{f}(X_i))^2 \right\}.$$

Ainsi les splines, qui ont été estimées sur plusieurs années, gardent l'information liée au modèle d'origine (saisonnalité de la consommation). De plus ce modèle propose plus de souplesse en prenant en compte les données récentes et en modulant la forme des splines.

### Critère de performance

Après avoir estimé nos modèles il faut trouver un critère pour les comparer.

$$\text{RMSE (root mean square error)}(Y, \hat{Y}, 1, n) = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}$$

$$\text{MAPE (mean absolute percentage error)}(Y, \hat{Y}, 1, n) = \frac{100}{n} \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{Y_i}$$

Un handicap du RMSE est que sa valeur dépend de la consommation moyenne (plus elle est élevée, plus le RMSE est fort) : ceci est assez contraignant pour une modélisation où nous prenons en compte une perte de clients (variation de la consommation moyenne). Dans cette étude l'utilisation du MAPE est plus pertinente.

## 2.2 Comparaison entre une réestimation GAM et une régression linéaire sur les effets

Les modèles de prévision de consommation sont calés une fois par an (la valeur des coefficients est valable pour l'année). Ces estimations sont donc très sensibles (en terme de performance) à des changements de comportement rapides, ou à des crises. Ainsi le modèle estimé avant la perte



de clients (qui sera notre référence), ne sera pas mis à jour pour correspondre aux habitudes (tandis que les modèles adaptatifs seront mis à jour chaque semaine).

Nous nous intéressons à une perte de l'ordre de 10% du portefeuille EDF sur les 32 000. Dans notre situation nous prenons le parti de modéliser cette perte de clients par un départ brusque des clients au 1<sup>er</sup> janvier 2012 (la généralisation à des pertes graduées est traité en Annexe). Le facteur d'oubli est choisi exponentiellement décroissant ( $\omega_i = \omega^{n-i}$ ), nous prenons  $\omega$  sur une grille  $\Omega$ . Dans la suite le trait horizontal rouge représente les performances du modèle de référence.

Pour chaque facteur d'oubli de la grille  $\Omega$  nous faisons une prévision, avec mise à jour chaque semaine, sur l'année 2012 [figure 2.1].

Le graphe des performances (en prévision) montre la différence de comportement entre la réestimation GAM et la régression linéaire sur les effets.

- La réestimation GAM est très instable : la meilleure performance est localisée entre 0.99 et 1 et les performances se dégradent rapidement en dessous de 0.99,
- La régression linéaire sur les effets est très stable : de meilleures performances que le modèle précédent et restant bornées entre 3% et 4% (avec une performance minimale sur un large intervalle de la grille  $\Omega$ ).

Ces premiers résultats montrent un certain avantage pour la réestimation par régression linéaire sur les effets (nous verrons que ces résultats se confirment en regardant les performances par instant, type de jour, mois). On peut justifier cet avantage par le fait que la réestimation GAM demande le calcul de nouvelles fonctions splines, ce qui nécessite d'avoir un historique sur plusieurs années ce qui n'est pas le cas lorsque le facteur d'oubli est inférieur à 0.98.

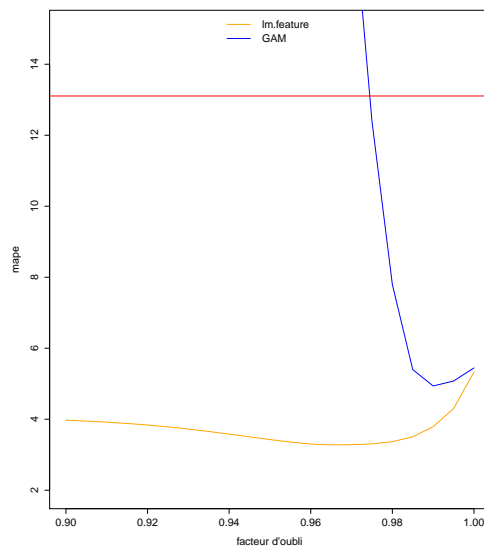


FIGURE 2.1 – Performances des modèles GAM et régression linéaire sur les effets

## 2.2.1 Performances par instant

Travaillant avec un pas de temps demi-horaire, nous avons à estimer 48 modèles pour une journée. Les performances des deux méthodes par instant dans la journée sont données sur les figures 2.2 (nous avons uniquement pris trois instants de la journée ; à 5h, 12h et 22h)

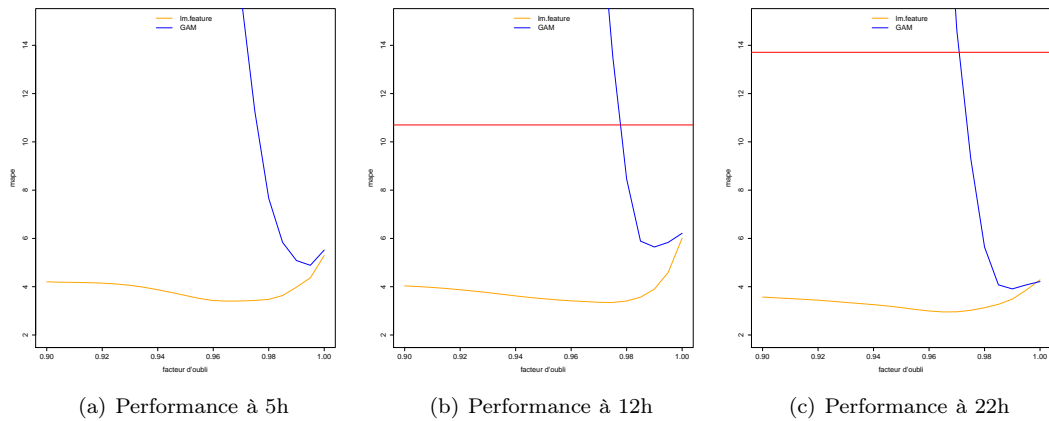


FIGURE 2.2 – Performance par instant

Nous faisons les mêmes constatations que précédemment : le modèle GAM souffre d'une grande instabilité ce qui n'est pas le cas de la régression linéaire sur les effets.

Nous distinguons aussi des courbes de performances très différente en fonction de l'instant de la journée (la dégradation des performances est plus ou moins rapide lorsque facteur d'oubli est loin de 1).

## 2.2.2 Performances par jour

En classifiant les journées en 9 types différents on peut analyser l'efficacité des méthodes en fonction du type de jour. La nomenclature pour les types de jours est

- Type 1, 2, ..., 7 : Dimanche, Lundi, ..., Samedi,
- Type 8 : jour fériés,
- Type 9 : jours EJP.

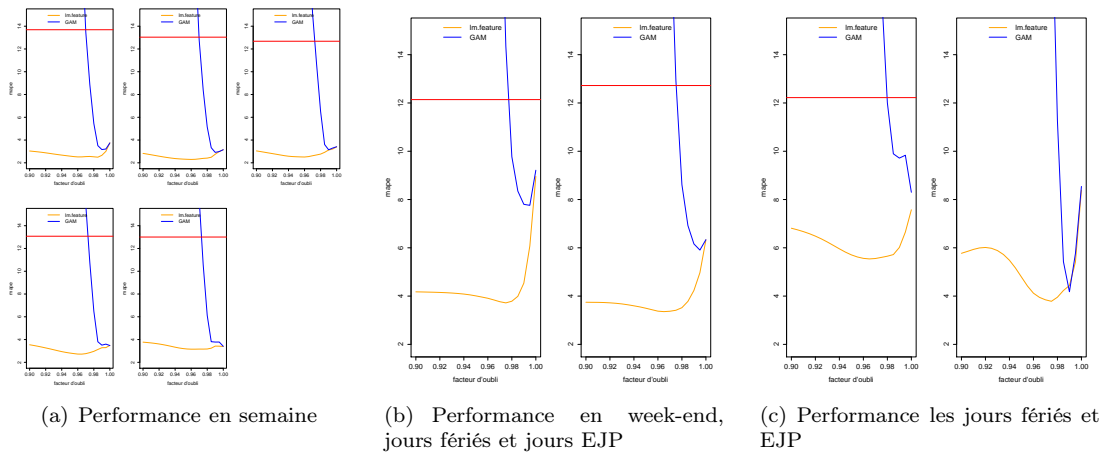


FIGURE 2.3 – Performance en semaine

Nous avons des graphes assez différents suivant le type de jour : les jours de fin de semaine ont des courbes assez semblables, les jours de semaine sont aussi de même nature, en revanche il y a une nette différence entre ces deux classes (ce qui paraît assez naturel puisque les 32 000 correspond à un agrégat très sensible à l'effet week-end).

Nous tirons les mêmes conclusions pour la comparaison entre les deux types de mise à jour.

### 2.2.3 Performances par mois

La prévision a été effectuée sur une année (52 semaines), le temps de latence dû à l'arrivée des données de la semaine précédente fait que nous réestimons les modèles sur 50 semaines (les deux premières semaines de janvier sont estimées avec le modèle de référence, en conséquence les performances de janvier sont mauvaises).

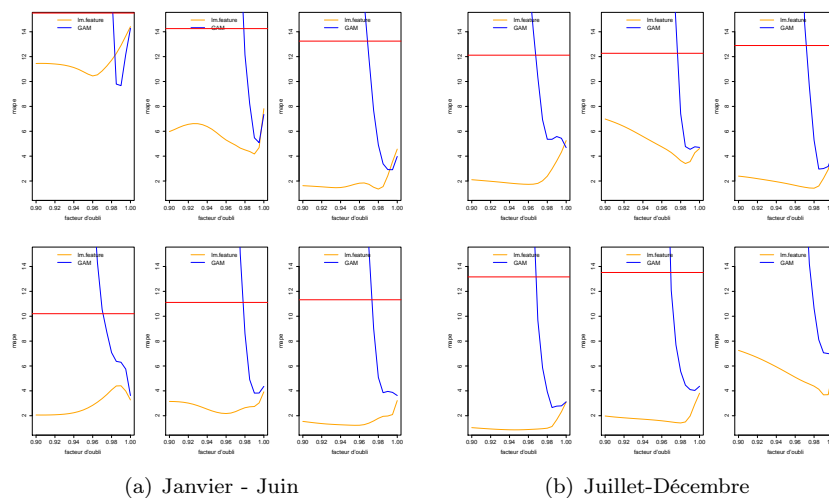


FIGURE 2.4 – Performance par mois

### 2.2.4 Conclusion

Les considérations précédentes nous permettent de faire un premier choix dans la modélisation de la consommation avec forte variation de périmètre : celle de la régression linéaire sur les effets plutôt que la réestimation d'un modèle GAM.

On peut expliquer les principaux défauts de la modélisation GAM par le fait que l'estimation des splines nécessite un historique important ce qui n'est pas vérifié pour des facteurs d'oubli trop faible.

## 2.3 Les effets croisés

Notre modèle de référence fait apparaître des effets croisés "DayType5 : JF : Rupture", qui lors de la mise à jour sont considérés comme un seul effet, par conséquent on estime qu'un seul coefficient  $\alpha$  devant cet effet.

$$\begin{aligned}
Y = & \alpha_1 \hat{s}_1(Temperature) + \alpha_2 \hat{s}_2(DayPosan) \\
& + \alpha_3 \hat{\beta}_1 \cdot NiveauEJP + \alpha_4 \hat{\beta}_2 \cdot tendance \\
& + \alpha_5 (DayType : JF : Rupture) \cdot \hat{\beta}_3.
\end{aligned}$$

Il pourrait être judicieux de regrouper certains de ces effets croisés et d'estimer un coefficient pour chacun.

Nous pouvons prendre des sous-groupes des effets croisés :  $((DayType : JF : Rupture)_i)_{i \leq o}$ . Et remplacer la partie bleue de l'équation précédente par

$$\sum_{5 \leq i \leq o+4} \alpha_i (DayType : JF : Rupture)_i$$

Nous proposons les sous-groupes suivant :

- Regrouper les types de jour en deux groupes : semaine et week-end,
- Regrouper les jours fériés en deux groupes : jour férié et hors jour férié.

Cette décomposition donne quatre sous-groupes (et donc quatre coefficients à réestimer). Les performances de ce nouveau modèle est mis en comparaison du précédent sur la figure 2.5

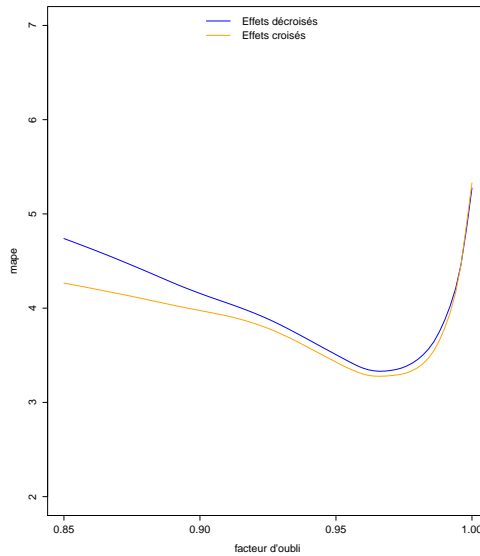


FIGURE 2.5 – Comparaison entre le modèle avec effets croisés considéré comme un tout et celui dont on a décroisé les effets

Ce nouveau modèle est moins bon que le précédent (les performances restent néanmoins très proches). Ceci peut s'expliquer par le fait que l'introduction de nouvelles variables induit une instabilité dans l'estimation, ainsi les résultats en prévision ne sont pas aussi concluants que souhaités.

Nous faisons le choix de ne plus utiliser ce nouveau modèle (décroisant les effets).

## 2.4 Les effets fixés

Nous travaillons dorénavant uniquement avec la régression linéaire sur les effets.

Un nouvel aspect de la modélisation est de savoir s'il est pertinent d'estimer un coefficient linéaire pour chaque effet. Certains effets n'ont peut être pas besoin d'être réestimés

$$\begin{aligned}
 Y - \hat{s}_1(\text{Temperature}) = & \alpha_2 \hat{s}_2(\text{DayPosan}) \\
 & + \alpha_3 \hat{\beta}_1 \cdot \text{NiveauEJP} + \alpha_4 \hat{\beta}_2 \cdot \text{tendance} \\
 & + \alpha_5 (\text{DayType} : \text{JF} : \text{Rupture}) \cdot \hat{\beta}_3.
 \end{aligned}$$

Nous étudions, dans la suite, la qualité des différents modèles en fonction des effets qui sont réestimés ou non

- *NiveauEJP*,
- *Tendance*,
- $s_1(\text{Temperature})$ ,
- $s_2(\text{DayPosan}, \text{by} = \text{DayType2})$  : si l'on décide de ne pas réestimer cet effet alors on ne réestime aucune des deux splines.

Cela nous donne 16 modèles possibles. Les courbes de performance de ces 16 modèles ne sont pas très agréables à représenter. Nous en donnons néanmoins quelques unes : les courbes seront classées en deux groupes

- Premier groupe, les effets fixés sont (les graphes seront rangés de gauche à droite et de haut en bas) : Rien, Température, DayPosan, NiveauEJP, Température-DayPosan, Température-NiveauEJP, DayPosan-NiveauEJP, Température-DayPosan-NiveauEJP,
- Second groupe, les effets fixés sont (les graphes seront rangés de gauche à droite et de haut en bas) : Tendance, Température-Tendance, DayPosan-Tendance, NiveauEJP-Tendance, Température-DayPosan-Tendance, Température-NiveauEJP-Tendance, DayPosan-NiveauEJP-Tendance, Température-DayPosan-NiveauEJP-Tendance.

Les performances sur la période de prévision de ces différents modèles gardent les avantages de stabilité que nous avons observés précédemment [figure 2.6].

Le premier groupe de graphique représente les modèles pour lesquels on réestime la tendance, le second groupe ceux pour lesquels on ne réestime pas la tendance

Nous pouvons essentiellement faire les mêmes analyses pour les courbes appartenant au même groupe

- Premier groupe : des performances contrastées, avec des minima localisés mais un MAPE faible. Remarquons aussi que pour chacun des modèles les minima ne sont pas atteints pour les mêmes facteurs d'oubli, cela laisse penser que chaque modèle peut avoir un comportement différent suivant le type de jour, de mois ou de l'instant dans la journée.
- Second groupe : des performances moins bonnes, les modèles sont plus stables avec une première partie relativement plate et une dégradation vers un facteur d'oubli égal à 1.

Nous verrons dans la suite que ces résultats se confirment lorsque l'on regarde les performances par instant, jour, mois : ne pas réestimer la tendance est très mauvais pour la prévision.

Il risque d'être difficile de choisir un seul modèle parmi les huit qui resterons.

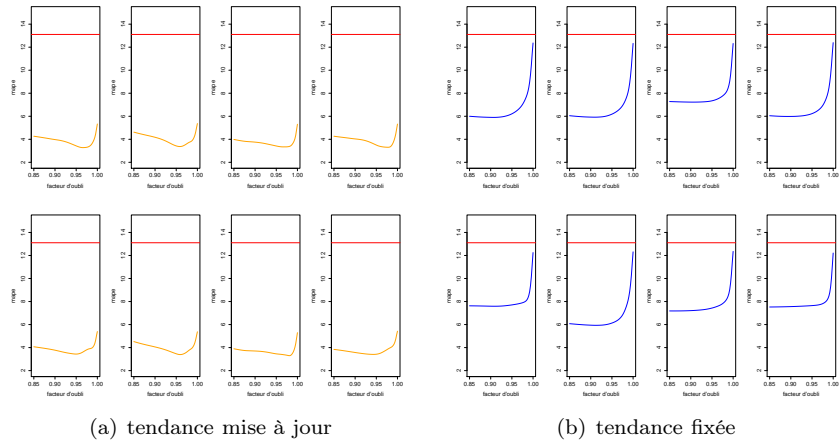


FIGURE 2.6 – MAPE en prévision

### 2.4.1 Performances par instant

Comme précédemment pour trois instants dans la journée (5h, 12h et 22h)

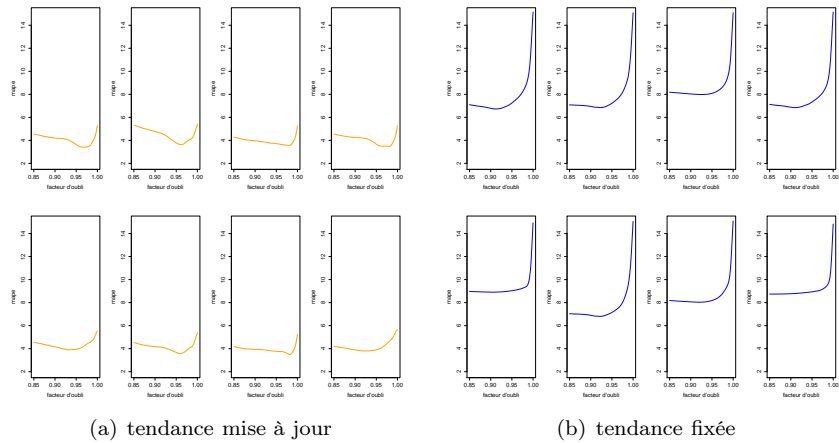


FIGURE 2.7 – MAPE à 5h

La principale remarque reste la scission entre les deux familles de courbes. Mettre à jour la tendance fournit de meilleures performances.

### 2.4.2 Performances par jour

Nous travaillons avec les mêmes type de jour que défini précédemment : un jour de week end, un jour de semaine, un jour férié et un jour EJP.

Nous remarquons que les écarts entre les deux familles de modèles sont moins importants pour les jours de semaine (pour le lundi il reste significatif, pour d'autres jours de la semaine cela est moins évident : voir l'Annexe)

Pour les jours de week-end on a un avantage de la première famille de modèle.

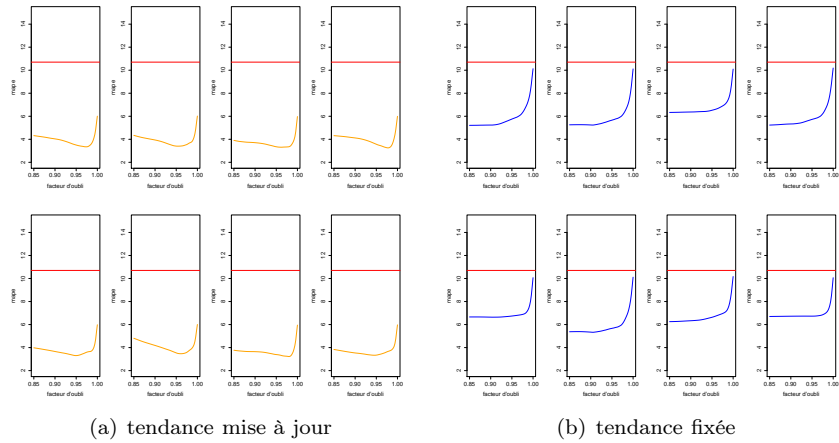


FIGURE 2.8 – MAPE à 12h

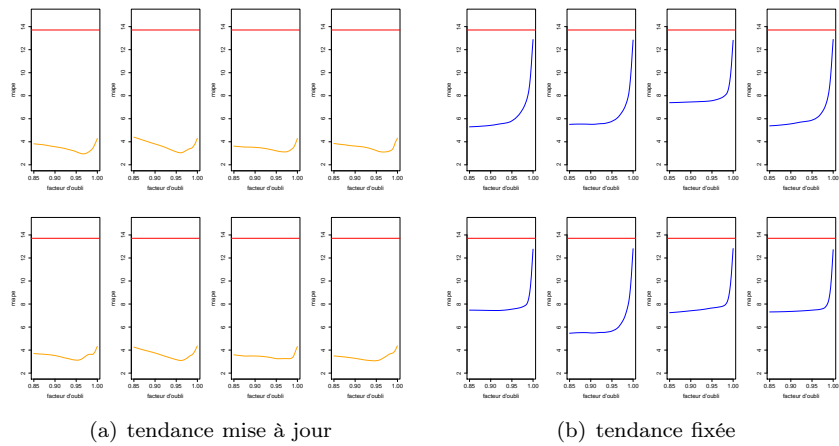


FIGURE 2.9 – MAPE à 22h

Pour le cas des jours fériés et des jours EJP voir l'Annexe.

### 2.4.3 Performances par mois

Les performances des différents modèles sont très variables en fonction du mois de l'année. En effet les deux premiers mois sont assez mauvais, ensuite des mois assez stables se succèdent. Les mois d'août et de décembre sont plus mauvais [voir Annexe].

Pour le mois de mai, on distingue la grande stabilité du second groupe et les meilleures performances du premier groupe.

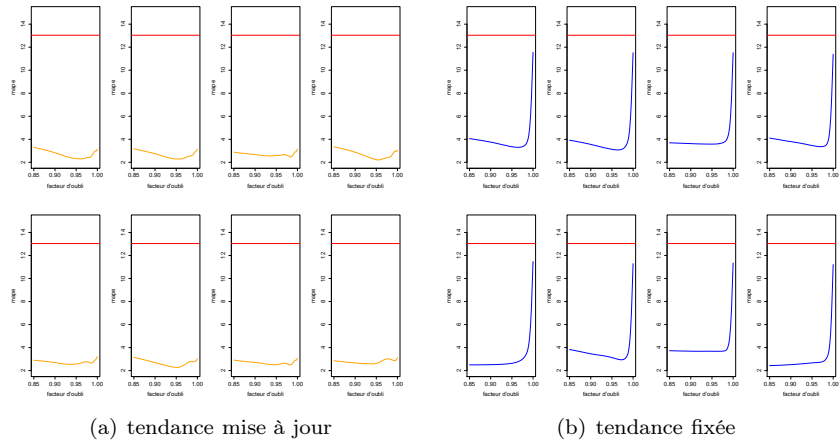


FIGURE 2.10 – MAPE le lundi

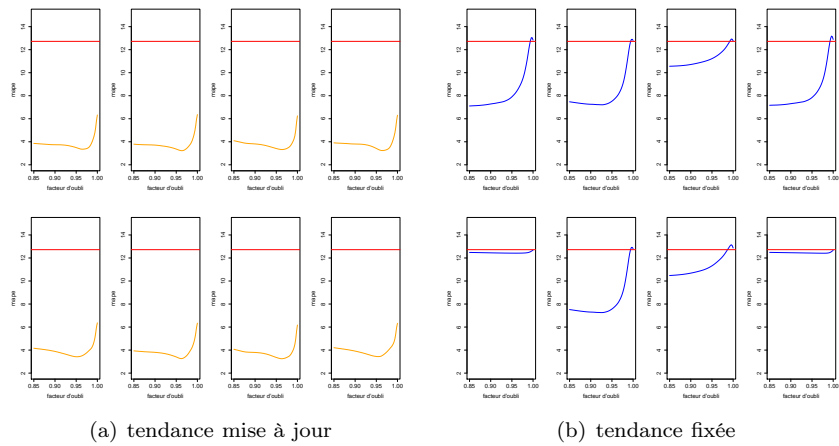


FIGURE 2.11 – MAPE le samedi

## 2.4.4 Conclusion

Nous avons vu qu'il est compliqué de choisir un modèle pour la prévision de consommation, chacun ayant ses avantages et ses inconvénients. Cependant nous constatons que globalement fixer la tendance (second groupe de courbes) dégrade les performances des modèles. De même au sein du premier groupe de courbes nous pourrions étudier avec plus de minutie le comportement de chaque modèle, nous donnons quelques pistes d'analyses

- Il y a un ensemble de courbes "plus stables" que les autres, dans le sens où lorsque le facteur d'oubli diminue, les performances se dégradent moins vite (les courbes 3-5-7-8). Ce sont les courbes pour lesquelles la spline de position de jour dans l'année est fixée.
- Les courbes 2-5-6-8 ont une plage de facteur d'oubli optimal relativement large par rapport aux autres courbes. Ce sont les modèles pour lesquels la spline de température est fixée.
- Une observation précise des courbes permet de voir que pour un facteur d'oubli proche de 1 la croissance de la performance est différente selon que certains effets sont fixés ou non.



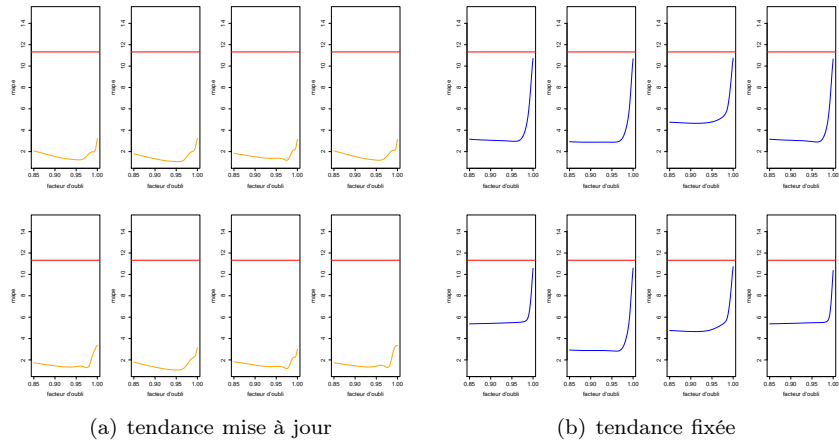


FIGURE 2.12 – MAPE de mai

Malgré ces remarques on constate que les courbes possédant une bonne stabilité ont une performance moins bonne que les courbes moins stables

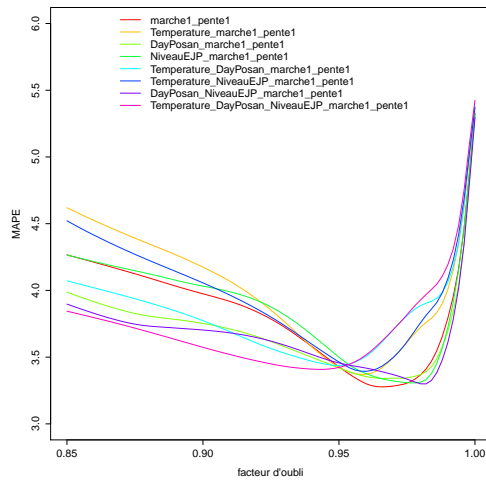


FIGURE 2.13 – Comparaison de performance des huit modèles

## 2.5 Sélection du facteur d'oubli à priori

Les résultats précédents décrivent des modèles pour lesquels le facteur d'oubli est fixé sur toute la période de prévision. Cette méthode nous donne déjà une certaine souplesse dans la mise à jour de nos modèles. Cependant il serait intéressant d'essayer de tirer parti des différentes valeurs du facteur d'oubli pour essayer d'améliorer les performances des prévisions.

Nous proposons deux méthodes de sélection du facteur d'oubli :

- Méthode naïve : le facteur d'oubli est fixé pour toute la période de prévision (la méthode utilisée depuis le début). Il faut choisir le facteur d'oubli à l'avance (en fonction des informations que l'on peut avoir sur la future variation de périmètre),
- Seconde méthode (prévision a priori) : choisir le facteur d'oubli en fonction des performances sur la dernière semaine

Ces deux méthodes sont à comparer à la meilleure prévision possible (prévision a posteriori) qui choisit chaque semaine le meilleur facteur d'oubli.

Nous constatons d'importantes différences entre ces trois méthodes. Pour chacun des huit modèles nous avons les performances suivantes

	1	2	3	4	5	6	7	8	facteur d'oubli
Méthode naïve	3.30	3.37	3.35	3.38	3.51	3.39	3.42	3.52	0.96
	3.28	3.50	3.34	3.32	3.70	3.50	3.37	3.71	0.97
	3.37	3.72	3.37	3.31	3.89	3.76	3.30	3.93	0.98
	3.79	4.00	3.69	3.72	4.06	4.09	3.61	4.19	0.99
Méthode a priori	3.51	3.63	3.89	3.50	3.89	3.50	3.86	3.74	
Méthode a posteriori	2.44	2.42	2.50	2.35	2.49	2.40	2.37	2.37	

**Remarque.** *Remarquons que les performances pour un facteur d'oubli optimal est largement en dessous des deux autres choix de facteur d'oubli (ce n'est pas nécessairement étonnant). Cependant nous voyons que la méthode a priori est celle qui donne les plus mauvaises performances. Nous pouvons expliquer cela par le temps de latence entre la prévision et l'arrivée de nouvelles données (il y a deux semaines d'écart entre la semaine qui permet de choisir le facteur d'oubli et la semaine de prévision).*

## 2.6 Comportement des coefficients estimés

Regardons le comportement des coefficients de la régression linéaire sur les effets. Rappelons que nous estimons un coefficient par effet, par instant de la journée, par mise à jour et par facteur d'oubli. Pour proposer un graphique lisible nous allons faire une moyenne des coefficients sur la journée (= un coefficient pour la journée et par effet).

Pour voir comment évolue les coefficients nous regardons le modèle pour lequel tous les effets sont réestimés. De plus nous choisissons le facteur d'oubli qui assure la meilleure performance chaque semaine.

où marche et pente représente respectivement l'ordonnée à l'origine et la pente de la tendance.

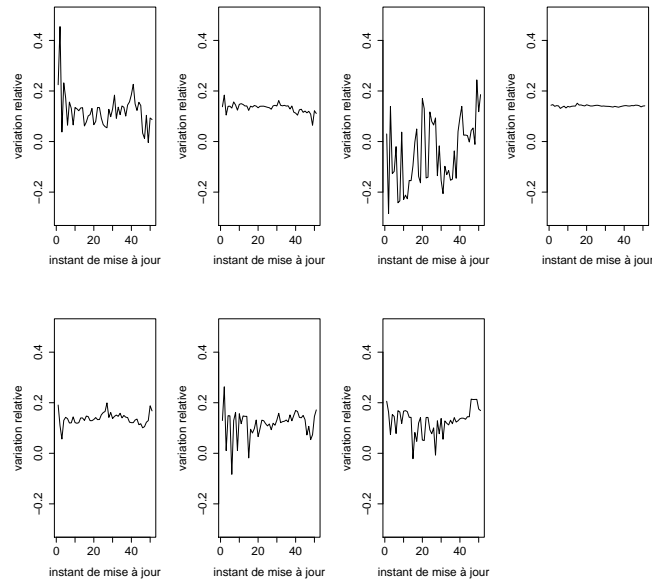


FIGURE 2.14 – Evolution des coefficients en fonction des mises à jour (NiveauEJP, marche, pente, DayType :JF :Rupture,  $s(\text{Température})$ ,  $s(\text{DayPosan } 1)$ ,  $s(\text{DayPosan } 2)$ )

## 2.7 Ouverture

L'ensemble des résultats énoncés peuvent gagner encore en précision, et l'étude de certains paramètres peut permettre l'amélioration sensible des résultats. Nous donnons quelques pistes possibles d'améliorations :

- Etudier des formes différentes de facteur d'oubli : quelques résultats ont été obtenus (voir Annexe) sur des formes inverse quadratique ayant une grande stabilité. Utiliser aussi ce facteur d'oubli (et en proposer d'autres) pour les prochaines études.
- Les méthodes de choix du facteur d'oubli : trouver des méthodes pour améliorer les performances (choisir le facteur d'oubli ayant la meilleure performance sur les trois derniers mois au lieu d'une semaine). Faire de la GCV en choisissant le facteur d'oubli en même temps que le paramètre de régularisation  $\lambda$  (valable pour la réestimation des modèles GAM). Faire du mélange de prédicteur entre les différents facteurs d'oubli.
- Les méthodes de réestimation : proposer des méthodes différentes de réestimation (régression robuste, régression LASSO, régression Ridge).
- Amélioration calculatoire : proposer des réestimations qui prennent en compte le fait qu'on a déjà estimé un modèle de référence GAM (mise en jour récursive comme pour `bam` et `bam.update`).
- Généralisation : tester sur d'autres bases de données les algorithmes et étudier les performances.
- Evolution des estimations : étudier plus précisément l'évolution des coefficients lors des réestimations pour s'assurer de la stabilité du modèle.

- Mélange de prédicteurs : utiliser des algorithmes de mélanges pour améliorer la prévision finale.
- Etudier plus précisément le cas des pertes progressives (= avec des départs de clients moins brusques)

# Chapitre 3

## ANNEXES

Les résultats qui sont illustrés dans cette partie ne seront pas commentés, les analyses étant identiques aux précédents. Nous invitons le lecteur à regarder les mêmes aspects de stabilité, de performance, de croissance et décroissance de courbes en fonction du facteur d'oubli, et de comparer ces aspects entre les modèles.

### 3.1 les effets fixés

#### 3.1.1 Performances par jour

Quelques figures complémentaires sur le comportement des modèles en fonction du jour

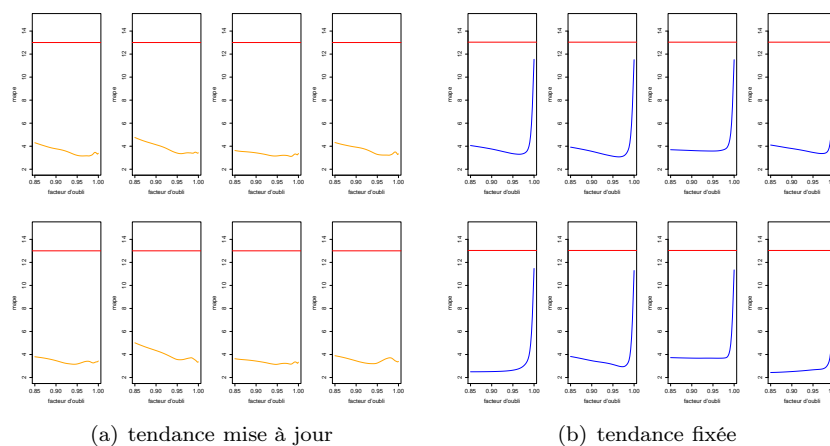


FIGURE 3.1 – MAPE du mardi

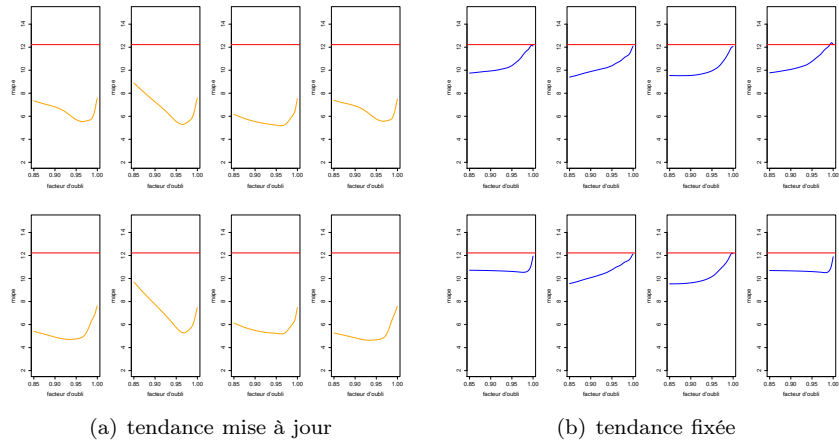


FIGURE 3.2 – MAPE des jours fériés

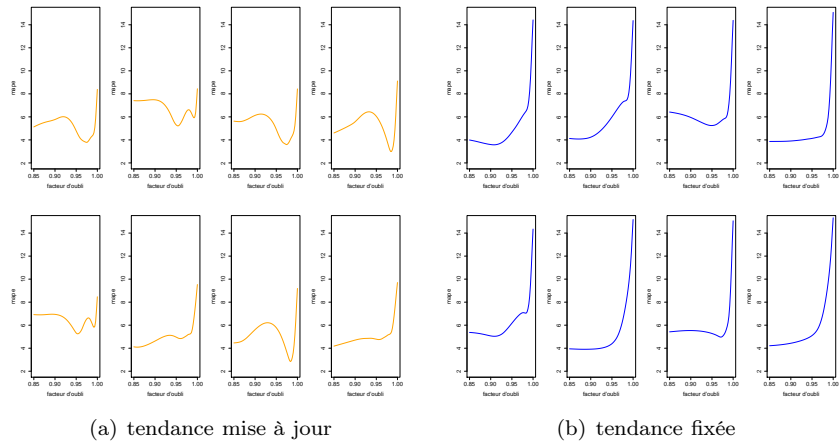


FIGURE 3.3 – MAPE des jours EJP

### 3.1.2 Performances par mois

On voit que la situation du mois de février n'est pas bonne : pour justifier cela on regarde la consommation réalisée au cours du mois. La courbe noire est le réalisé tandis que la courbe orange est la prévision qui est faite (les traits rouges correspondent à la fin de chaque mois). On distingue clairement un changement dans la consommation pour les deux premières semaines de février (avec des week-ends très différents des précédents, ceci explique la forme particulière des performances pour le mois de février).

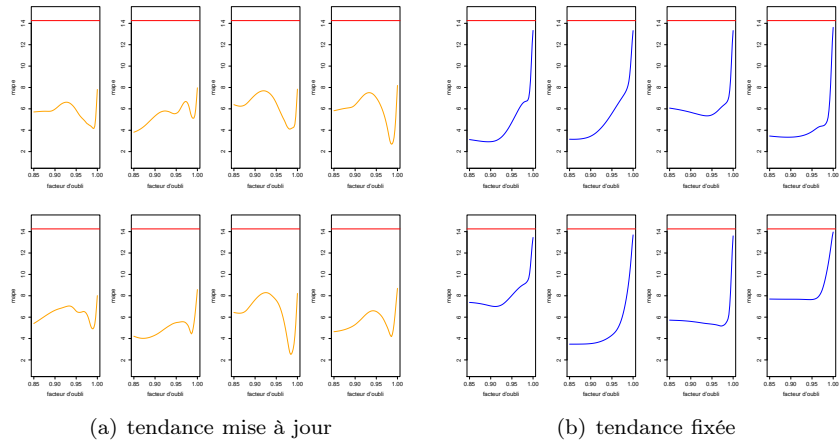


FIGURE 3.4 – MAPE de février

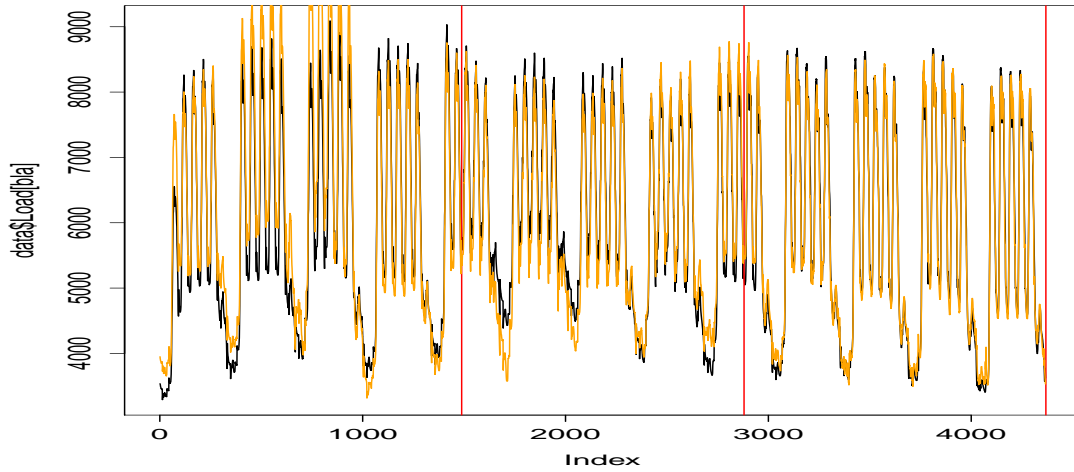


FIGURE 3.5 – Prévision en janvier-février-mars (orange)

### 3.2 Une perte de clients plus importante (30%)

Les études ont aussi porté sur une variation plus importante du périmètre du portefeuille EDF. Pour cette étude nous nous plaçons dans le même cadre que précédemment (une perte de clients brutale au 1<sup>er</sup> janvier 2012, avec le même historique d'apprentissage). Les performances, naturellement plus mauvaises, n'ont pas un comportement différent de ce que l'on a pu voir pour le cas de la rupture précédente.

Nous avons vu précédemment que nous aurions tendance à ne pas fixer la tendance (la seconde famille des courbes étant globalement plus mauvaise). Pour avoir des graphes plus lisibles nous ne faisons plus apparaître ces graphes ci.

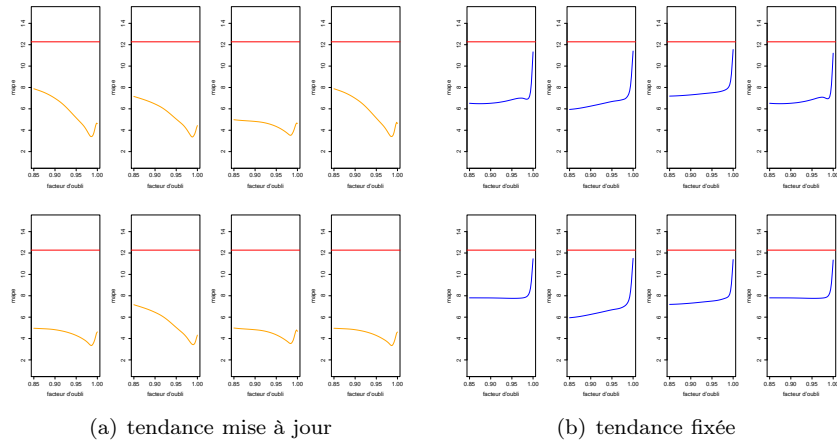


FIGURE 3.6 – MAPE de aout

### 3.2.1 Performances globales

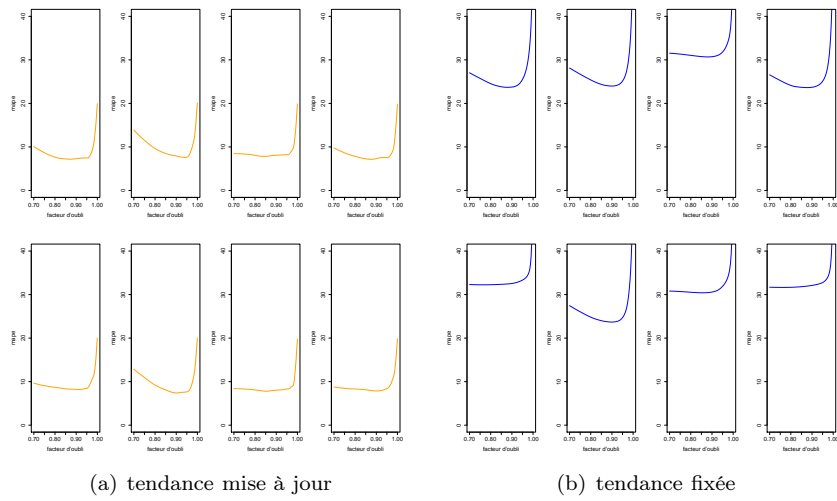


FIGURE 3.7 – Performances globales



### 3.2.2 Performances par instant

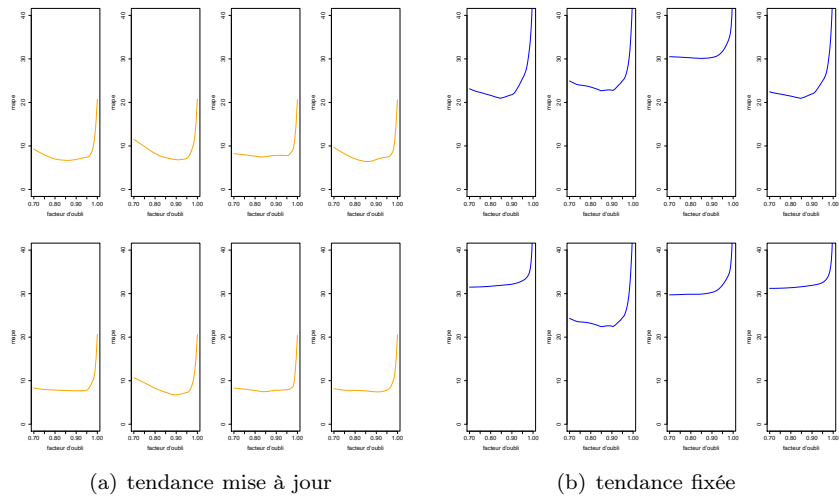


FIGURE 3.8 – Performances à 12h

### 3.2.3 Performances par jour

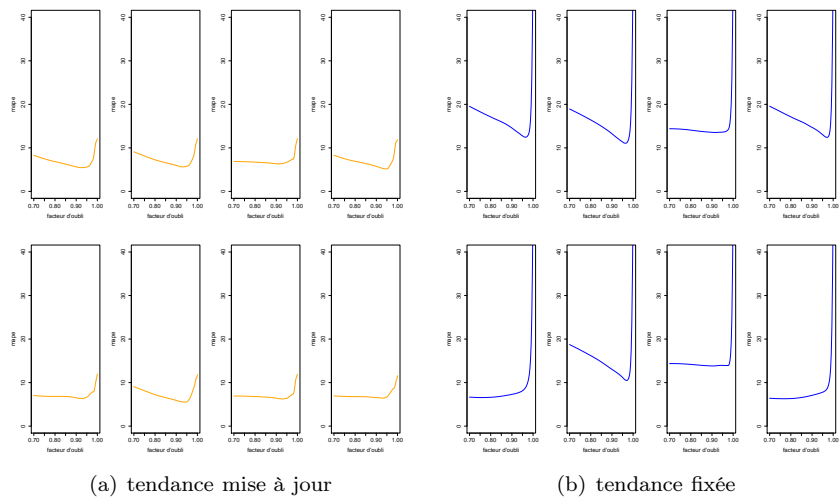


FIGURE 3.9 – Performances le mardi

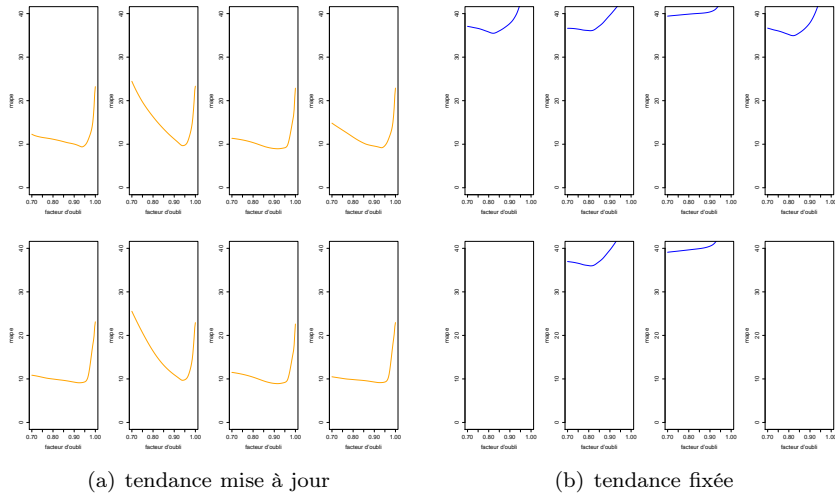


FIGURE 3.10 – Performances les jours fériés

### 3.2.4 Performances par mois

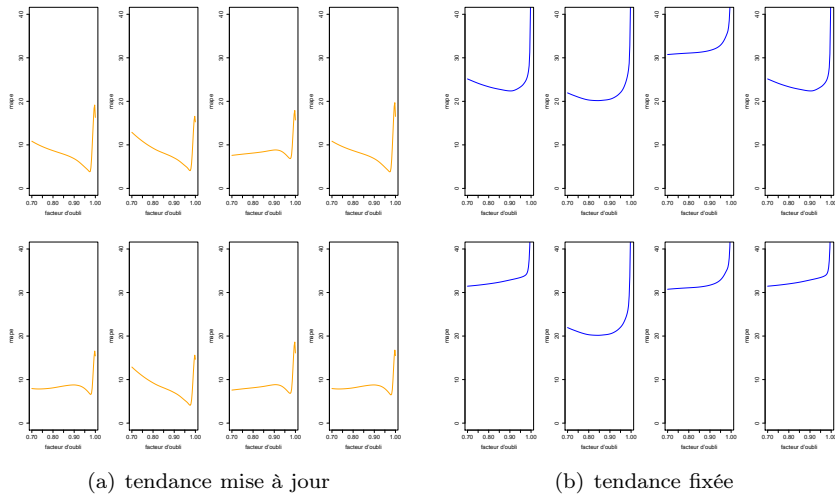


FIGURE 3.11 – Performances en août

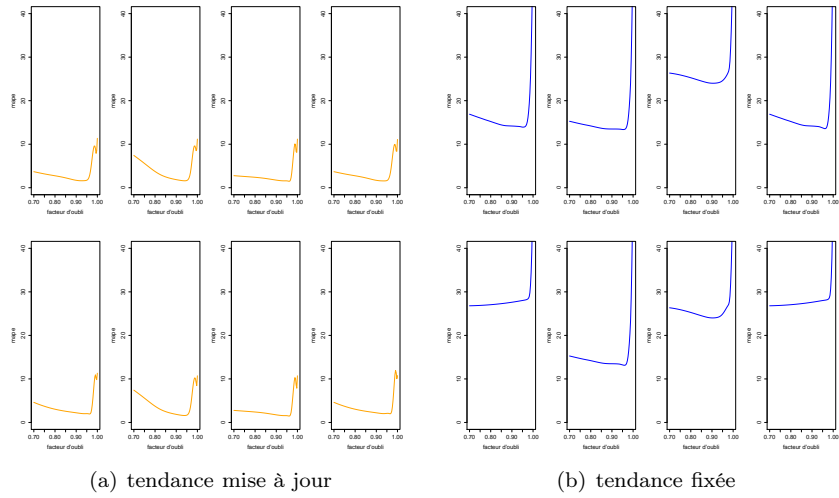


FIGURE 3.12 – Performances en juin

### 3.2.5 Conclusion

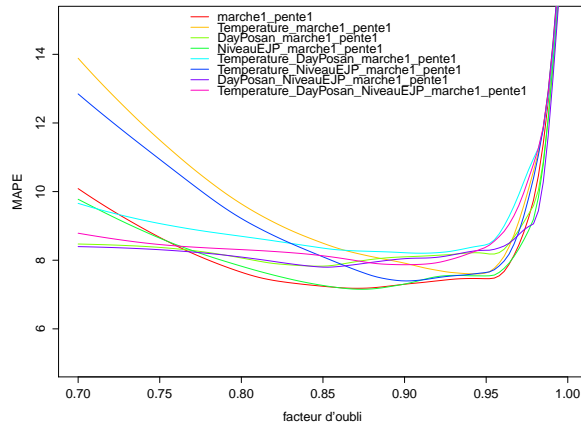


FIGURE 3.13 – Comparaison des performances des 8 modèles

Les conclusions sont essentiellement les mêmes que pour la perte de 10%.

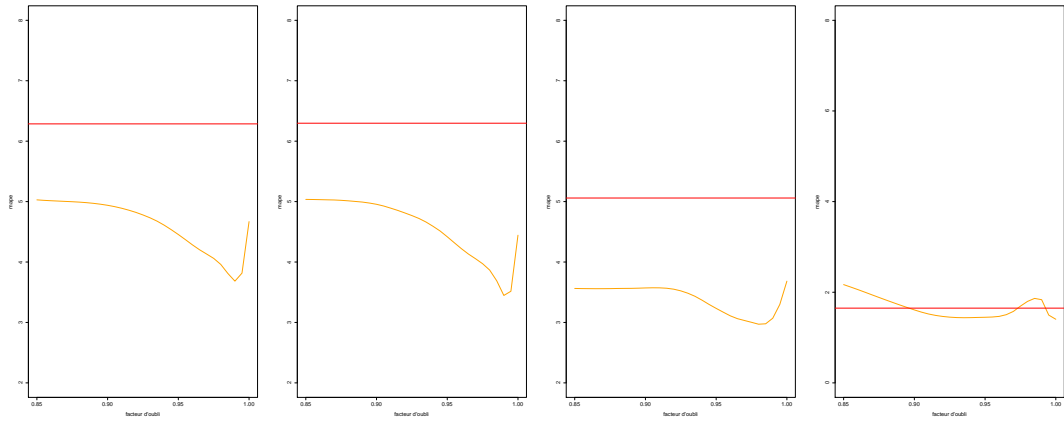
### 3.3 Une perte de clients progressive

Une modélisation plus réaliste des variations du portefeuille EDF est de proposer des évolutions mensuelles mais moins brusques (avec la possibilité d'avoir des retours de clients). L'analyse des données reste identique aux cas précédents, cependant les résultats peuvent différer, essentiellement pour deux raisons :

- Le modèle de référence peut redevenir compétitif dans le cas où une perte puis un gain de clients peuvent se compenser,
- L'instabilité de la réestimation des modèles dont l'historique d'apprentissage est oublié (il est parfois plus intéressant d'avoir un modèle stable et légèrement moins bon).

Pour plus de lisibilité on étudie uniquement le modèle pour lequel on réstime la totalité des effets.

Pour une perte de la forme :  $-70\%$  en janvier,  $+60\%$  en février,  $+10\%$  en avril,  $+50\%$  en juillet.



(a) Performance globale (b) Performance à 12h (c) Performance le mardi (d) Performance en juin

FIGURE 3.14 – Performances

### 3.4 Facteur d'oubli inverse quadratique (perte de 10% au 1<sup>er</sup> janvier 2012)

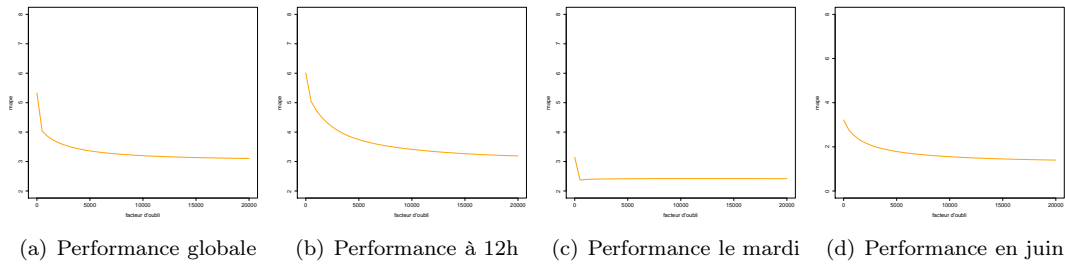


FIGURE 3.15 – Performances

## 3.5 RKHS : Introduction & plus

### 3.5.1 Définition et notation

**Définition.** On dit qu'un espace  $\mathcal{H} \subset \mathbb{R}^X$  est un RKHS (Reproducing Kernel HILBERT Space=Espace de HILBERT à Noyau Reproductible) lorsqu'il vérifie les conditions suivantes :

- $\mathcal{H}$  est un espace de HILBERT (muni de son produit scalaire  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ ),
- Pour tout  $x$  dans  $X$  l'application projection

$$\begin{aligned} p_x : \mathcal{H} &\rightarrow \mathbb{R} \\ f &\mapsto f(x), \end{aligned}$$

est continue.

Nous noterons le produit scalaire de  $\mathcal{H}$   $\langle \cdot, \cdot \rangle$ , dans la mesure où cette notation n'induit aucune ambiguïté.

**Proposition.** Soit  $\mathcal{H}$  un RKHS sur  $X$  alors il existe une application  $\Phi : X \rightarrow \mathcal{H}$  que l'on appelle caractéristique (feature map) telle que :

$$\forall x \in X, \forall f \in \mathcal{F}, f(x) = \langle f, \Phi(x) \rangle.$$

*Démonstration.* C'est une application immédiate du théorème de représentation de RIESZ puisque l'application projection est une forme linéaire continue □

**Définition.** Si  $\mathcal{H}$  est un RKHS on définit le noyau de  $\mathcal{H}$  comme étant l'application :

$$\begin{aligned} k : X \times X &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \langle \Phi(x), \Phi(y) \rangle. \end{aligned}$$

**Proposition.** Le noyau d'un RKHS est une forme quadratique positive

$$(i.e. : \forall (\alpha_i, x_i)_{i \leq p} \in (\mathbb{R} \times X)^p, \sum_{i, j \leq p} \alpha_i \alpha_j k(x_i, x_j) \geq 0)$$

*Démonstration.* si on se donne deux familles finies  $(x_i)_{i \leq p} \in X^p$  et  $(c_i)_{i \leq p} \in \mathbb{R}^p$  alors

$$\sum_{i=0, j=0}^p c_i c_j k(x_i, x_j) = \left\| \sum_{i=0}^p c_i \Phi(x_i) \right\|^2 \geq 0.$$

□

**Remarque.** Le RKHS en tant qu'espace de HILBERT et les propriétés de symétrie du noyau  $k$  nous laisse espérer l'utilisation possible des théorèmes fonctionnelles.

**Proposition.** (Propriétés Reproductibles)

$$\begin{aligned} \forall (x, y) \in X^2, \forall f \in \mathcal{F}, f(x) &= \langle f, k(\cdot, x) \rangle \\ k(x, y) &= \langle k(\cdot, x), k(\cdot, y) \rangle. \end{aligned}$$

*Démonstration.* Ces formules reposent sur les égalités suivantes (remarquons que  $\Phi(x)$  est une fonction de  $\mathcal{H}$ )

$$\begin{aligned} \forall y, k(y, x) = \langle \Phi(y), \Phi(x) \rangle = \Phi(x)(y) &\iff \forall y, \Phi(x)(y) = k(y, x) \\ &\iff \Phi(x) = k(\cdot, x). \end{aligned}$$

□

**Théorème** (de ARONSZAJN 1950). *Les propositions suivantes sont équivalentes :*

- $k : X \times X \rightarrow \mathbb{R}$  est un noyau symétrique positif.
- Il existe un RKHS tel que  $k$  est le noyau du RKHS.
- Il existe  $\mathcal{H}$  et  $\Phi : X \rightarrow \mathcal{H}$  tel que  $k(x, y) = \langle \Phi(x), \Phi(y) \rangle$ .

*Démonstration.* La majorité des implications sont triviales, on se contente donc de montrer qu'à partir d'un noyau symétrique positif on peut construire un RKHS.

On pose  $\mathcal{H}_0$  l'espace vectoriel engendré la famille de fonctions  $(k(\cdot, x))_{x \in X}$ . On définit sur cet espace le produit scalaire suivant :

$$\left\langle \sum_{i \leq n} \alpha_i k(\cdot, x_i), \sum_{j \leq n} \beta_j k(\cdot, y_j) \right\rangle = \sum_{i, j \leq n} \alpha_i \beta_j k(x_i, y_j).$$

On a ainsi construit un espace pré-hilbertien, on pose alors  $\mathcal{H}$  comme étant le complété de  $\mathcal{H}_0$ . Nous obtenons bien un RKHS. □

**Exemples.** — [Noyau linéaire] On se donne  $\mathcal{H} = \text{Vect} \{ \langle \alpha, \cdot \rangle, \alpha \in \mathbb{R}^d \}$  munit du produit scalaire  $\langle f_\alpha, f_\beta \rangle = \langle \alpha, \beta \rangle$ . On vérifie rapidement que l'application caractéristique de cet espace est  $\Phi(x)(\cdot) = x \cdot$  et que le noyau associé est  $k(x, y) = \langle x, y \rangle$ ,

- [Noyau quadratique] Le noyau de  $\mathbb{R}^2$

$$k(x, y) = (\langle x, y \rangle)^2 = (x \cdot y)^2 = x_1^2 y_1^2 + 2x_1 x_2 y_1 y_2 + x_2^2 y_2^2 = \langle \Phi(x), \Phi(y) \rangle,$$

avec  $\Phi(x) = (x_1^2, \sqrt{2}x_1 x_2, x_2^2)$  (Ce n'est pas une réelle égalité mais plutôt une bijection) et  $\Phi(x)(y) = k(x, y)$ . Le RKHS ainsi formé est celui des polynômes à deux inconnus homogènes de degré 2

$$\mathcal{H} = \text{Vect} \{ k(\cdot, x), x \in \mathbb{R}^2 \} = \text{Vect} \{ x_1^2, x_1 x_2, x_2^2, x \in \mathbb{R}^2 \},$$

- [Noyau polynomiale] Le noyau

$$\forall (x, y) \in (\mathbb{R}^d)^2, k(x, y) = (\langle x, y \rangle)^p.$$

L'application caractéristique est alors en bijection avec un espace de dimension  $\binom{d+p-1}{d-1}$  (et correspond à une généralisation du précédent avec les coefficients multi-indices). Le RKHS est celui des polynômes à  $d$  inconnus homogènes de degré  $p$ ,

- [Noyau exponentiel] Le noyau

$$\forall (x, y) \in \mathbb{R}^d, k(x, y) = \exp(-\lambda|x - y|),$$

— [Noyau gaussien] Le noyau

$$\forall (x, y) \in \mathbb{R}^d, k(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right),$$

**Remarque.** La démonstration précédente nous donne, en plus d'un RKHS, une famille dense de cet espace (dans le cas fini nous avons même une famille génératrice)

**Théorème.** Soient  $k_1$  et  $k_2$  deux noyaux alors  $k = k_1 + k_2$ ,  $k = k_1 \cdot k_2$  et  $ck_1$  (pour  $c > 0$ ) sont encore des noyau (et définissent les RKHS  $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ ,  $\mathcal{H}_1 \otimes \mathcal{H}_2$  et  $\sqrt{c}\mathcal{H}_1$ ) Soient  $(k_i)_{i \geq 0}$  une famille de noyau telle que

$$\exists k, \forall (x, y) \in X^2, k(x, y) = \lim_{n \rightarrow +\infty} k_n(x, y),$$

alors  $k$  est un noyau

**Exemples.** — [Noyau linéaire général] le noyau de  $\mathbb{R}^d$

$$k(x, y) = \langle x, y \rangle + c$$

— [Noyau polynomiale général] le noyau de  $\mathbb{R}^d$

$$k(x, y) = (\langle x, y \rangle + c)^p$$

*Démonstration.* Le cas de la somme n'est pas difficile, il suffit de vérifier les propriétés des deux noyaux restent vraies pour leur somme (symétrie et positivité). le RKHS définie ainsi est donc inclus dans la somme directe des deux RKHS qui définissent les deux noyaux. Réciproquement la somme directe est un RKHS dont le noyau est la somme des deux et le produit scalaire est le produit scalaire canonique sur un tel espace.  $\square$

**Théorème (du Représentant).** Soient  $X$  non vide (et non nécessairement fini),  $\mathcal{H}$  un RKHS et  $k$  son noyau associé. On se donne  $g$  et  $\Omega$  deux fonctions à valeurs dans  $\mathbb{R} \cup \{\infty\}$  (et  $\Omega$  décroissante) alors l'ensemble

$$\operatorname{argmin}_{f \in \mathcal{H}} \{g(f(x_1), \dots, f(x_n)) + \Omega(\|f\|)\},$$

admet un élément (un représentant) de la forme :  $\sum_{i=1}^n \alpha_i k(\cdot, x_i)$ .

Il semble presque superflus de faire une remarque sur l'intérêt de ce théorème. Nous pouvons noter qu'un résultat purement formel sur les espaces de HILBERT puisse fournir une méthode (ou du moins une piste de recherche très précise) de résolution d'un problème d'optimisation. Ce théorème nous permet de chercher un minimum sur un sous-espace vectoriel de dimension fini d'un espace fonctionnel (qui sont la plus part du temps infini voire même non séparable...).

*Démonstration.* On se donne un élément  $f \in \mathcal{H}$  et soit  $\sum_{i=1}^n \alpha_i k(\cdot, x_i)$  la projection de  $f$  sur  $\operatorname{Vect} \{\Phi(x_i), i \leq n\}$  et  $v$  la projection sur l'orthogonal.

$$f(x_j) = \left\langle \sum_{i=0}^n \alpha_i \Phi(x_i) + v, \Phi(x_j) \right\rangle = \sum_{i=0}^n \alpha_i k(x_i, x_j).$$

Ainsi la valeur de  $g(f(x_1), \dots, f(x_n))$  ne dépend pas de la projection de  $f$  sur  $\operatorname{Vect} \{\Phi(x_i), i \leq n\}^\perp$ . De plus on a



$$\begin{aligned}\Omega(\|f\|) &= \Omega\left(\sqrt{\left\|\sum_{i=1}^n \alpha_i \Phi(x_i)\right\|^2 + \|v\|^2}\right) \\ &\geq \Omega\left(\left\|\sum_{i=1}^n \alpha_i \Phi(x_i)\right\|\right).\end{aligned}$$

On en déduit donc qu'on peut choisir un minimum dans le sous-espace  $\text{Vect}\{\Phi(x_i), i \leq n\}$ .  $\square$

### 3.5.2 & plus

Nous donnons quelques résultats supplémentaires, d'ordre théoriques (non nécessaire pour le sujet de ce mémoire).

**Théorème** (BÖCHNER). *Soit  $k$  un noyau invariant pas translation (i.e :  $k(x, y) = q(x - y)$  où  $q : \mathbb{R}^d \mapsto \mathbb{R}$ ) alors  $k$  est défini positif si et seulement si  $q$  est la transformée de FOURIER d'une mesure de BOREL finie positive (nous avons également que si  $q$  est continue en 0 alors cette mesure admet une densité par rapport à celle de LEBESGUE)*

*Démonstration.* On ne démontre ici que l'implication de la gauche vers la droite et on ne donne qu'une esquisse de preuve pour la réciproque ([CGCDM05] et [Che]) Si  $q$  est la transformée de FOURIER d'une mesure positive on a (où l'on appelle  $\hat{q}$  la mesure associée)

$$q(x) = \frac{1}{(2\pi)^d} \int \exp(it\omega x) d\hat{q}(\omega)$$

On peut donc écrire

$$\begin{aligned}\sum_{i,j \leq p} \alpha_i \alpha_j k(x_i, x_j) &= \sum_{i,j \leq p} \alpha_i \alpha_j q(x_i - x_j) \\ &= \frac{1}{(2\pi)^d} \int \sum_{i,j \leq p} \alpha_i \alpha_j \exp(it\omega(x_i - x_j)) d\hat{q}(\omega) \\ &= \frac{1}{(2\pi)^d} \int \left| \sum_{j \leq p} \alpha_j \exp(it\omega(x_j)) \right|^2 d\hat{q}(\omega) \geq 0\end{aligned}$$

Schéma de preuve. (référence :[CGCDM05] et [Che])

**Définition.** *Une fonction  $f$  est dite de type positif si et seulement si :*

$$\forall (\alpha_i, x_i)_{i \leq p} \in (\mathbb{C} \times \mathbb{R}^d)^p, \sum_{i,j \leq p} \overline{\alpha_i} \alpha_j f(x_i - x_j) \geq 0.$$

Ce type de fonctions possède des propriétés intéressantes (voir [CGCDM05]) : la valeur en 0 est toujours positive, la valeur en 0 majore toutes les autres, elles sont paires ( $f(t) = \overline{f(-t)}$ ), si  $f$  est  $g$  sont des fonctions à valeurs complexe et  $k$  une fonction de type positif alors dès que la quantité existe on a le résultat  $\int_{\mathbb{R}^d} f(y)g(x)k(x - y)dx dy \geq 0$ (pour montrer cela on peut utiliser le théorème des points de LEBESGUE ou le théorème de convergence dominée).

1<sup>er</sup> cas : la fonction  $q$  est intégrable.

Comme pour de nombreuses démonstrations sur les critères d'inversion de transformée de FOURIER (d'ondelettes, de Ridgelettes) nous nous donnons l'approximation de l'unité par le noyau gaussien (qui à la propriété d'être presque invariant par transformation de FOURIER et qui de plus appartient à l'espace de SCHWARTZ [voir la référence...]) pour lequel nous savons que la transformation de FOURIER est une isométrie).

On commence par montrer l'égalité (grâce au théorème de PARSEVAL et les propriétés de convolution)

$$\int_{(\mathbb{R}^d)^2} f(x)\overline{f(y)}k(x-y)dx dy = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{k}(\xi)|\hat{f}(\xi)|^2 d\xi,$$

en remplaçant  $|\hat{f}|^2$  par le noyau gaussien  $h_\sigma(x) = \exp\left(-\frac{\|x\|^2}{2\sigma^2}\right)$  (ce qui est possible dans l'espace de SCHWARTZ par bijection),

$$\int_{\mathbb{R}^d} \hat{k}(\xi)h_\sigma(\xi)d\xi = \hat{k} * h_\sigma(0),$$

puis en faisant tendre  $\sigma$  vers 0 nous obtenons  $\hat{k}(0) \geq 0$ .

En multipliant par  $\exp(i \langle t, \cdot \rangle)$  nous obtenons  $\hat{k}(t) \geq 0$ .

Nous montrons ensuite que  $\hat{k}$  est intégrable (nous pourrions alors utiliser la formule inverse de la transformée de FOURIER et avoir le résultat souhaité).

Pour le cas où  $q$  n'est pas intégrable voir les références données.  $\square$

**Remarque.** Ceci nous permet de définir un RKHS à partir de mesure finie positive et d'élargir un peu plus le glossaire de nos RKHS. Mais aussi d'avoir la propriété suivante.

**Proposition.** Dans le cas où les bonnes hypothèses sont vérifiées on obtient, en définissant la mesure.

$$d\mu = \frac{1}{\hat{q}(x)} dx$$

et l'on obtient alors

$$\mathcal{H} = \left\{ f, \hat{f} \in L^2(\mathbb{R}, \mu) \right\}$$

que l'on peut munir de son produit scalaire (canonique sur un espace  $L^2$ )

$$\langle f, g \rangle_{\mathcal{H}} = \int \hat{f}(x)\hat{g}(x)d\mu(x)$$

**Proposition.**

$$\|f\|_{\mathcal{H}} = \frac{1}{2\pi} \int \frac{|\hat{f}(\omega)|^2}{\hat{q}(\omega)} d\omega$$

*Démonstration.* Voir [Che]  $\square$

Le théorème de BÖCHNER nous donne un critère pour qu'une fonction soit de type positif.

### Cas fini

On suppose ici que l'espace  $X$  est fini, nous avons un espace vectoriel de dimension fini, par conséquent le noyau est diagonalisable dans une base orthonormée.

Nous connaissons alors une base orthonormale du RKHS, cela revient à l'étude d'un espace vectoriel de dimension finie.

## Cas infini

Nous aimerions proposer une construction analogue à celle du cas fini (notamment par l'utilisation du théorème spectral). La construction de la base de MERCER se place dans ce cadre.

Soit une mesure borélienne  $\mu$  finie sur  $X$  compact, de support  $X$ . Supposons qu'on ait un noyau positif et continu  $k$  alors l'application linéaire

$$L_k : L^2(X, \mu) \rightarrow L^2(X, \mu)$$

$$f \rightarrow \int k(\cdot, t) f(t) d\mu(t)$$

est bornée, auto-adjointe, compact et positif.

*Démonstration.* Commençons par montrer que cette application est bien définie (donc qu'elle est bien à valeurs dans  $L^2(X, \mu)$ ).

$$|L_k f(t) - L_k f(u)| \leq \|k(t, \cdot) - k(u, \cdot)\|_2 \|f\|_2 \quad (3.1)$$

$$\leq \|f\|_2 \max_{x \in X} |k(t, x) - k(u, x)| \sqrt{\mu(X)}. \quad (3.2)$$

Par compacité de  $X$  on en déduit l'uniforme continuité de  $k(t, \cdot)$  et donc la continuité de  $L_k f(\cdot)$ . D'où  $L_k f \in L^2(X, \mu)$ .

On montre ensuite la continuité de l'application linéaire  $L_k$  (il suffit qu'elle soit bornée). On remarque d'abord que  $k(\cdot, \cdot)$  est une fonction bornée (car continue sur un compact), ainsi

$$\|L_k f\|_2 = \left( \int_X \left( \int_X k(x, t) f(t) d\mu(t) \right)^2 d\mu(x) \right)^{1/2} \stackrel{CS}{\leq} \left( \int_X (\sqrt{\mu(X)} \max_{t \in X} k(x, t) \|f\|_2)^2 d\mu(x) \right)^{1/2} \quad (3.3)$$

$$\leq \mu(X) \max_{x, t \in X} k(x, t) \|f\|_2 \quad (3.4)$$

Pour montrer que  $L_k$  est un opérateur compact on se donne  $(f_n)$  une suite bornée de  $(L^2(X, \mu), \|\cdot\|_2)$ , la suite  $(L_k f_n)$  est une suite de fonctions continues uniformément bornées (i.e  $\|L_k f_n\|_\infty$  est bornée), cela se montre de la même manière que l'inégalité (2). De plus elle définit une famille équicontinue (l'inégalité (1) nous permet d'avoir ce résultat). On en déduit par le théorème d'ASCOLI-ARZELA, que l'on peut extraire une sous suite convergente et sa limite sera dans l'ensemble des fonctions continues donc dans  $L^2(X, \mu)$ .

$L_k$  est auto-adjoint (c'est une utilisation du théorème de FUBINI, dont on vérifie aisément les hypothèses).

Enfin pour la positivité on a

$$\begin{aligned} \langle f, L_k f \rangle &= \int_{X^2} f(t) f(x) k(t, x) d\mu(t) d\mu(x) \\ &= \lim_{k \rightarrow \infty} \frac{\mu(X)}{k^2} \sum_{i, j \leq k} k(x_i, x_j) f(x_i) f(x_j) \geq 0. \end{aligned}$$

La dernière égalité se justifie par le théorème de différentiation de LEBESGUE qui assure  $(\forall f \in L^1 \implies \text{p-s, } \frac{1}{\mu(B_x(r))} \int_{B_x(r)} f(t) dt \xrightarrow{r \rightarrow 0} f(x))$   $\square$

Nous pouvons donc justifier l'existence d'une base hilbertienne  $(\psi_i)_{i \geq 0}$  de  $L^2(X, \mu)$  formée de vecteurs propres associés à la famille de valeurs propres  $(\lambda_i)_{i \geq 0}$ .

### 3.5.3 RKHS et apprentissage statistique : cas particulier des bases de splines

#### Apprentissage

Nous nous donnons pour premier exemple la régression Ridge (linéaire dans le premier cas)

$$Y = X\beta + \varepsilon,$$

$$\hat{\beta} \in \underset{\beta \in \mathbb{R}^d}{\operatorname{argmin}} \{ \|Y - X\beta\|^2 + \lambda \|\beta\|^2 \}.$$

Soit le RKHS des applications linéaires (le noyau linéaire  $k(x, y) = \langle x, y \rangle$ ), alors

$$Y = f_\beta(X) + \varepsilon,$$

cherchons  $f_\beta$  de la forme  $\sum \alpha_i k(\cdot, x_i)$ , par conséquent le problème d'optimisation revient à chercher

$$f_{\hat{\beta}} \in \underset{\alpha}{\operatorname{argmin}} \left\{ \|Y - \sum \alpha_i k(X, x_i)\|^2 + \lambda \alpha^t K \alpha \right\}.$$

Ce qui nous donne, après dérivation (et la remarque que la fonction est convexe donc atteint son minimum où elle s'annule), l'estimateur

$$\hat{\alpha} = (K + \lambda I)^{-1} Y.$$

L'avantage de cet exemple, outre qu'il est très simple, est qu'il est généralisable. En effet nous avons uniquement utilisé la structure d'espace vectoriel du RKHS.

Soit un RKHS quelconque nous voulons faire une régression Ridge

$$Y = f(X) + \varepsilon,$$

$$\hat{f} \in \underset{f \in \mathcal{H}}{\operatorname{argmin}} \{ \|Y - f(X)\|^2 + \lambda \|f\|_{\mathcal{H}}^2 \}.$$

Le théorème du représentant nous invite à choisir  $\hat{f}$  de la forme  $\sum \alpha_i k(\cdot, x_i)$  et par la même suite de calculs nous obtenons finalement

$$\hat{\alpha} = (K + \lambda I)^{-1} Y.$$

Nous avons une méthode pour trouver, sur un espace potentiellement complexe, le meilleur estimateur de la perte Ridge.

Nous proposons une autre manière de pénaliser les estimateurs dans le cadre de la régression Ridge. Soit

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1,$$

$$Y = f(X) + \varepsilon,$$

$$f = f_0 + f_1,$$

$$\hat{f} \in \underset{f \in \mathcal{H}}{\operatorname{argmin}} \{ \|Y - f(X)\|^2 + \lambda \|f_1\|_{\mathcal{H}}^2 \}.$$

où  $\mathcal{H}_0$  est appelé l'espace nul. Nous choisissons de pénaliser la projection de  $f$  sur  $\mathcal{H}_1$  ; or  $\|f_1\|_{\mathcal{H}}^2 = \|f - f_0\|_{\mathcal{H}}^2$ , ainsi la pénalisation représente l'éloignement de la fonction  $f$  à l'espace nul.

## RKHS et bases de splines

On utilise ici comme référence [Wan11] et [Wah90].

Nous allons voir comment la théorie des RKHS peut s'inscrire naturellement pour une estimation Ridge avec des bases de splines.

**Définition.** On dit que  $f$  est absolument continue s'il existe une fonction  $g$  intégrable telle que

$$f(x) - f(a) = \int_a^x g(t)dt.$$

**Notation.** Par abus de langage nous noterons cette fonction  $g$  avec la notation de dérivation  $f'$ .

On se donne l'espace (que nous appellerons par abus de langage espace de SOBOLEV) suivant

$$W_2^m([a, b]) = \left\{ f : [a, b] \rightarrow \mathbb{R}, \forall k \leq m-1, f^{(k)} \text{ est absolument continue, } \int_a^b f^{(m)}(t)^2 dt < \infty \right\}.$$

Nous allons munir cet espace d'une structure d'espace de HILBERT (en réalité nous allons donner plusieurs structures différentes et en déduire des estimateurs différents pour la régression Ridge)

## Les splines polynomial

Avant d'introduire notre premier produit scalaire nous montrons que le développement de TAYLOR reste valable dans cet espace de SOBOLEV, grâce à l'hypothèse d'absolue continuité.

$$\begin{aligned} f(x) &= \int_a^x f'(t_1)dt_1 + f(a) \\ &\vdots \\ &= \int_a^x \int_a^{t_1} \dots \int_a^{t_{m-1}} f^{(m)}(t_m)dt_m + \sum_{i=1}^{m-1} \int_a^{t_i} f^{(i)}(a)dt_i + f(a) \\ &= \int_a^x \frac{(x-u)^{m-1}}{(m-1)!} f^{(m)}(u)du + \sum_{i=0}^{m-1} f^{(i)} \frac{(x-a)^i}{i!} \end{aligned}$$

On se donne le produit scalaire sur  $W_2^m([a, b])$

$$(f, g) = \sum_{k=0}^{m-1} f^{(k)}(a)g^{(k)}(a) + \int_a^b f^{(m)}(t)g^{(m)}(t)dt = \langle f, g \rangle_{\mathcal{H}_0} + \langle f, g \rangle_{\mathcal{H}_1}$$

Cette quantité donne une décomposition en deux espaces vectoriels qui sont en somme directe

$$\begin{aligned} \mathcal{H}_0 &= \text{Vect} \{1, (x-a), \dots, (x-a)^{m-1}/(m-1)!\}, \\ \mathcal{H}_1 &= \left\{ f, \forall k \leq m-1, f^{(k)}(a) = 0, \int_a^b f^{(m)}(t)^2 dt < \infty \right\}. \end{aligned}$$

Ces deux espaces sont des RKHS pour les noyaux

$$k_0(x, y) = \sum_{i=0}^m \frac{(x-a)^{i-1}}{(i-1)!} \frac{(y-a)^{i-1}}{(i-1)!},$$

$$k_1(x, y) = \int_a^b \frac{(x-u)_+^{m-1}}{(m-1)!} \frac{(y-u)_+^{m-1}}{(m-1)!} du.$$

On justifie ces deux résultats à partir du développement de TAYLOR et en trouvant une fonction  $\phi(x)$  telle que  $f(x) = \langle f, \phi(x) \rangle$ . Ainsi nous avons que  $W_2^m([a, b])$  munit de ce produit scalaire est un RKHS dont le noyau est  $k = k_0 + k_1$ .

Si  $P_1$  représente l'opérateur de projection sur  $\mathcal{H}_1$  alors

$$\|P_1 f\|^2 = \int_a^b f^{(m)}(t)^2 dt.$$

Ainsi lorsqu'on se donne une pénalité de la forme

$$\|Y - f(X)\|^2 + \lambda \int_a^b f^{(m)}(t)^2 dt,$$

De plus la norme de la projection de  $f$  sur  $\mathcal{H}_1$  représente aussi la distance entre  $f$  et l'espace  $\mathcal{H}_0$  donc une norme importante pour la projection  $P_1 f$  signifie que  $f$  est éloigné de  $\mathcal{H}_0$  et la fonction est moins "régulière".

**Jargon.** On pourra appeler espace nul  $\mathcal{H}_0$ . Cette appellation fait référence au fait que sur  $W_2^m([a, b])$  la semi-norme définie par  $\|f\|^2 = \int_a^b f^{(m)}(t)^2 dt$  est nulle pour toute fonction de  $\mathcal{H}_0$

### Une autre construction de RHKS

On propose un autre produit scalaire pour l'espace  $W_2^m([0, 1])$

$$(f, g) = \sum_{i=0}^{m-1} \left( \int_0^1 f^{(i)}(t) dt \right) \left( \int_0^1 g^{(i)}(t) dt \right) + \int_0^1 f^{(m)}(t) g^{(m)}(t) dt$$

et alors l'espace de SOBOLEV est la somme directe de

$$\mathcal{H}_0 = \text{Vect} \{B_0(x), \dots, B_{m-1}(x)/(m-1)!\}$$

$$\mathcal{H}_1 = \left\{ f, \forall k \leq m-1, \int_0^1 f^{(k)}(t) dt = 0, \int_a^b f^{(m)}(t)^2 dt < \infty \right\}$$

où  $B_i$  est le  $i$ -ème polynôme de BERNOULLI.

Les noyaux correspondant sont

$$k_0(x, y) = \sum_{i=0}^{m-1} \frac{B_i(x) B_i(y)}{i! i!}$$

$$k_1(x, y) = \frac{B_m(x) B_m(y)}{m! m!} + (-1)^{m-1} \frac{B_{2m}(|x-y|)}{2m!}$$

**Remarque.** *Nous n'avons pas changé d'espace fonctionnel pour cette nouvelle estimation et pourtant si nous devons trouver en pratique un nouvel estimateur nous n'aurions pas le même que précédemment, alors que la formule et la perte de régression sont les mêmes. Ceci peut paraître contre intuitif mais s'explique par le fait que le changement de noyau entraîne naturellement un changement dans l'estimateur choisit. Terminons par cette réflexion : le choix de l'espace fonctionnel est aussi important que le choix du produit scalaire.*

On peut ainsi construire un grand nombre de somme directe de notre espace et proposer des familles assez larges de splines (splines périodiques, splines sphériques, splines partielles, Thin-plate splines, ...) pour plus d'exemples voir le livre [Wan11].

Nous pouvons voir la variété des familles de splines que nous pouvons estimer avec une mise en oeuvre théorique assez faible grâce à la vision généraliste des RKHS. Et les avantages des estimations par des théorèmes d'optimisations facile à mettre en oeuvre.

### Les Thin-Plate Splines

Les familles de splines que nous avons étudiées précédemment nous ont permis de proposer plusieurs variante de la régression Ridge avec des familles de fonction à une variable, donnons nous maintenant la possibilité d'avoir une spline à plusieurs variables. Pour toute fonction  $f$  à  $d$  variables on se donne la quantité

$$J_m^d(f) = \sum_{\alpha_1 + \dots + \alpha_d = m} \frac{m!}{\alpha_1! \dots \alpha_d!} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \left( \frac{\partial^m f}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \right)^2 dx_1 \dots dx_d$$

et on définit

$$W_2^m(\mathbb{R}^d) = \{f, J_m^d(f) < \infty\}$$

que l'on munit du produit

$$(f, g) = \sum_{\alpha_1 + \dots + \alpha_d = m} \frac{m!}{\alpha_1! \dots \alpha_d!} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \left( \frac{\partial^m f}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \right) \left( \frac{\partial^m g}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \right) dx_1 \dots dx_d$$

Ce produit n'est pas scalaire puisqu'il n'est pas définit (les polynômes d'ordre inférieur à  $m - 1$  ont une norme nulle), ainsi  $J_m^d$  est une semie norme. On se donne  $\mathcal{H}_0$  l'espace vectoriel des fonctions dont la semie norme est nulle et on quotiente  $W_2^m(\mathbb{R}^d)$  par  $\mathcal{H}_0$ , on peut munir ainsi ce nouvel espace d'un produit scalaire.

On peut vérifier (par double inclusion) que tout élément de

$$\mathcal{H}_0 = Vect \left\{ (x_1, \dots, x_d)^{(\alpha_1, \dots, \alpha_d)}, \sum_{i \leq d} \alpha_i \leq m - 1 \right\}$$

est de norme nulle.

Une condition nécessaire et suffisante pour que l'application projection soit continue (et que cet espace soit un RKHS) est que  $2m > d$  (Voir Duchon (1977) et Meinguet (1979)), pour un énoncé de cette affirmation : p.129, <https://books.google.fr/books?id=9XHsCAAAQBAJ&pg=PA141&lpg=PA141&dq=>

# Bibliographie

- [AC10] S. Arlot and A. Celisse. A survey of cross-validation procedures for model selection. *Statistics surveys*, 2010.
- [Ar11] S. Arlot. Sélection de modèles et sélection d'estimateurs pour l'apprentissage statistique. <http://www.di.ens.fr/~arlot/peccot.htm>, 2011. Cours Peccot.
- [CG05] Y. Cao and Y. Golubev. On oracle inequalities related to a polynomial fitting. *mathematical methods of statistics*, 2005.
- [CG06] Y. Cao and Y. Golubev. On oracle inequalities related to smoothing splines. *mathematical methods of statistics*, 2006.
- [CGCDM05] M. Cottrell, V. Genon-Catalot, C. Duhamel, and T. Meyre. *Exercices de probabilités*. Cassini, 2005.
- [Che] C. Chesseboeuf. recalage d'images médicales. <http://petille.univ-poitiers.fr/notice/view/33342>.
- [CW79] P. Craven and G. Wahba. Smoothing noisy data with spline functions, estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 1979.
- [Del] B. Delyon. Régression, cours de deuxième année de master. <https://perso.univ-rennes1.fr/bernard.delyon/regression.pdf>. Université Rennes I.
- [Gasa] E. Gassiat. Statistiques, notes de cours, m1. <http://www.math.u-psud.fr/~gassiat/M1Stat.pdf>. Université Paris Sud.
- [Gasb] E. Gassiat. Statistiques, notes de cours, m2. <http://www.math.u-psud.fr/~gassiat/notesdecours-M2Stat.pdf>. Université Paris Sud.
- [RS12] V. Rivoirard and G. Stoltz. *Statistiques en action*. Vuibert, seconde édition, 2012.
- [Wah90] G. Wahba. *Spline Models for Observational Data*. University of Wisconsin at Madison, 1990.
- [Wan11] Y. Wang. *Smoothing Splines, methods and applications*. Chapman & Hall, 2011.
- [Woo06] S.N. Wood. *Generalized Additive Models, an introduction with R*. Chapman & Hall, 2006.