Mémoire de Magistère

Romain Mismer

Introduction

Vous trouverez dans ce document une description de mon parcours universitaire au sein du Magistère de Mathématiques de la faculté d'Orsay, ainsi qu'une présentation du domaine de recherche en statistiques dont ma thèse fait actuellement l'objet. En annexe sont joints les comptes-rendus des divers travaux effectués au long de mon parcours.

Table des matières

Chapitre 1

Parcours au sein du Magistère

J'ai étudié à la faculté d'Orsay de 2011 à 2015. Souhaitant depuis fort longtemps devenir professeur de mathématiques, c'est dès le début de ma première année en classe préparatoire MPSI au lycée Henri Poincaré à Nancy que je me suis intéressé au Magistère de Mathématiques d'Orsay. C'est au cours de ce cursus de 4 années (dont une année de césure dans le parcours magistère pour préparer l'agrégation) que mon intérêt et ma passion pour les probabilités et les statistiques se sont forgées. Vous trouverez dans ce qui suit une description plus détaillée de mon parcours.

L3 MFA

L'année de L3 a été l'occasion pour moi de consolider mes bases en mathématiques et d'en acquérir de nouvelles. J'ai suivi les modules de graphes et d'informatique au premier semestre et d'informatique théorique au second. ces modules m'ont particulièrement intéressé, tout comme les cours obligatoires et en particulier les cours d'algèbre et de topologie au premier semestre et de théorie de Fourier et de probabilités au second. Les cours de magistère, qui nous ont été dispensés par Dominique Hulin et Wendelin Werner, m'ont permis d'apprendre nombre de choses à la fois intéressantes et utiles en topologie générale, avec notamment des théorèmes et des démonstrations classiques pour qui prépare l'agrégation, mais j'ai aussi découvert les ensembles fractals et la dimension de Hausdorff.
Mon stage d'apprentissage hors mur a eu lieu dans un laboratoire de recherche en physique au sein du Laboratoire Environnement et Minéralurgie de Nancy. Il portait sur la supraconduction, sujet qui en lui-même m'avait toujours intéressé, et donc sur des techniques de résolution d'équations aux dérivées partielles appliquées à l'électromagnétisme. Ce stage a été l'occasion pour moi de découvrir le fonctionnement d'un laboratoire de recherche.

M1 MFA

C'est au cours de cette année que j'ai vraiment commencé à me spécialiser vers les probabilités et les statistiques, bien que j'ai tenté de suivre le plus de cours possibles, étant donné qu'ils étaient nombreux à me sembler passionnants. J'ai suivi avec une grande joie les cours de probabilité et de statistiques ainsi que le MAO qui allait avec, les cours d'analyse et mathématiques générales pour avoir des bases plus solides pour l'agrégation, j'ai découvert les distributions. J'ai également suivi le début du cours d'arithmétique, mais j'ai décidé de l'arrêter en cours d'année parce que les horaires se chevauchaient avec ceux du cours de statistiques et que je me suis aperçu que l'arithmétique n'était pas vraiment ma tasse de thé. En magistère nous avons eu une introduction à l'analyse spectrale par Frédéric Paulin (à l'humour et à la bonne humeur inégalés -et je réfute d'ailleurs immédiatement toute "tentative de corruption du lecteur"-) et à la théorie de l'information de Shannon par Pierre Pansu, deux cours particulièrement captivants et notamment pour moi ce dernier qui m'a fait découvrir un domaine assez particulier des probabilités et des statistiques. Mon sujet de TER portait sur les fonctions holomorphes et méromorphes, avec une approche assez topologique que j'ai particulièrement appréciée et un lien avec l'analyse harmonique dont j'avais vu les prémices dans le cadre du cours de Monsieur Paulin. J'ai choisi ce sujet car il était aussi l'occasion de consolider mes fondamentaux sur les fonctions holomorphes et la topologie générale.

M2 MFA : Probabilités et statistiques

Cette année m'a permis de me spécialiser en statistiques, notamment en grande dimension et en bayésien. Au début de l'année j'ai été quelque peu décontenancé par le fait que nous devions en théorie savoir si l'on allait s'orienter plutôt probabilités ou plutôt statistiques, en effet tout me paraissait intéressant ! Pour le magistère nous devions suivre un cours supplémentaire, et c'est aussi en partie ce qui m'a motivé à suivre un grand nombre de cours au premier semestre. J'ai compris lors de ce semestre que les statistiques étaient vraiment ce qui me tenait à coeur, et le sujet de stage et de thèse proposé par Ismaël Castillo réunissait toutes les notions qui me plaisaient le plus ! Au second semestre j'ai donc uniquement suivi les cours de statistiques qui, bien que n'étant pas vraiment en lien direct avec mon sujet de stgae, m'ont permis d'acquérir des connaissances standard en statistiques et notamment en algorithmique. Cette année a forgé pour moi une base solide pour aborder sereinement ma thèse encadrée par Ismaël Castillo, dont je vous présente en quelque sorte les prémices dans la grande partie qui suit.

Chapitre 2

Estimation bayésienne non-paramétrique et seuillage dans un cadre parcimonieux

En statistiques, de nombreux problèmes s'intéressent à un paramètre en très grande dimension sur lequel on ne dispose que d'une seule observation, par exemple après avoir moyenné, et sujette à un bruit. Le modèle auquel nous allons nous intéresser dans ce document est le suivant : on considère les observations à valeurs réelles $X = (X_1, \ldots, X_n)$, avec n un entier naturel que l'on fera tendre vers l'infini, telles que, pour tout $i \in \{1, \ldots, n\}$:

$$X_i = \theta_i + \varepsilon_i$$

où les ε_i représentent le bruit, on les considèrera donc comme des variables aléatoires de loi normale standard. On notera $\theta = (\theta_1, \ldots, \theta_n)$ notre paramètre d'intérêt (et donc inconnu), qui est donc un vecteur fixé de $\mathbb{R}^n$.
Il apparaît tout de suite que, sans connaître un minimum d'informations sur θ, il est compliqué de l'estimer efficacement. Dans ce document, on supposera l'existence d'une certaine parcimonie dans cette suite θ.

Cette hypothèse de parcimonie se rencontre dans de nombreux contextes, par exemple en astronomie, en traitement de l'image, en sélection de modèle, en data mining ou encore lorsque l'on fait de l'estimation non-paramétrique de fonction en utilisant des ondelettes.

Une approche naturelle à ces problèmes parcimonieux peut être le seuillage : si la valeur absolue d'un certain X_i dépasse un certain seuil t, alors on suppose qu'il correspond à un θ_i non nul et on l'estime alors généralement par X_i lui-même ; sinon, on l'estime comme étant zéro.

Mon domaine de recherche porte sur un cadre bayésien : on considère une loi a priori(aussi appelée *prior*) sur les θ_i qui permettra d'induire de la parcimonie, et on étudie certains aspects de la loi a posteriori de θ sachant X. Dans [1], Johnstone et Silverman mettent en valeur le fait que, sous certaines hypothèses, certains aspects de l'a posteriori sont de "bons" estimateurs en un certain sens. En effet, la médiane a posteriori est alors naturellement un estimateur de type seuillage et qui a en plus une propriété de rétrécissement, et la moyenne a posteriori, bien que n'ayant pas la propriété de seuillage, est tout de même un "bon" estimateur de par sa propriété de rétrécissement.

2.1 Cadre général

Le modèle de suite gaussienne s'écrit, avec $X = (X_1, \dots, X_n)$ le vecteur de $\mathbb{R}^n$ observé :

$$X_i = \theta_{0,i} + \varepsilon_i, \quad i = 1, \dots n, \tag{2.1}$$

où $\varepsilon_1, \dots, \varepsilon_n$ sont des variables aléatoires réelles indépendantes identiquement distribuées (iid) de loi $N(0,1)$ (de densité notée ϕ), et le paramètre $\theta_0 = (\theta_{0,1}, \dots, \theta_{0,n})$ appartient à la classe $\ell_0[p_n]$; définie par

$$\ell_0[p_n] = \{\theta \in \mathbb{R}^n, \ |\{i \in \{1, \dots n\}, \ \theta_i \neq 0\}| \leq p_n\},$$

pour $0 \leq p_n \leq n$, où $|A|$ est la cardinalité de l'ensemble A.
On supposera $p_n = o(n)$ quand $n \to \infty$.

On note E_{θ_0} l'espérance par rapport à la loi sur les données engendrée par (2.1)(autrement dit, sous P_{θ_0}, chaque X_i suit une loi normale de moyenne $\theta_{0,i}$ et de variance 1). On note $\|\cdot\|$ la norme euclidienne, $\|v\|^2 = \sum_{i=1}^n v_i^2$ pour $v = (v_1, \dots, v_n) \in \mathbb{R}^n$. On notera $p_{n,\theta}(X) = (2\pi)^{-n/2} e^{-\|X-\theta\|^2/2}$.

Un résultat dû à Donoho, Johnstone et al [6] sur la vitesse minimax est le suivant :

Théorème 1. *[Vitesse minimax] On note* $R_{n,2}(\ell_0[p_n]) = \inf_{\hat{\theta}} \sup_{\theta \in \ell_0[p_n]} \frac{1}{n} \sum_{i=1}^n E_\theta(\hat{\theta}_i - \theta_i)^2$*, où la borne inférieure est prise sur tous les estimateurs.*
On a alors $R_{n,2}(\ell_0[p_n]) = \frac{2p_n}{n}\big(\log \frac{n}{p_n}\big)(1 + o(1))$ *quand* $n \to \infty$*.*

2.2 Une loi a priori naturelle : le Spike and Slab *prior*

On considère l'a priori suivant sur θ, un a priori produit qui traduit la parcimonie du modèle :

$$\theta \sim \Pi = \bigotimes_{i=1}^n (1-\alpha)\delta_0 + \alpha\gamma \tag{2.2}$$

avec δ_0 la masse de Dirac en 0 sur $\mathbb{R}$, γ une densité de probabilité sur $\mathbb{R}$ à choisir et un paramètre $\alpha \in [0,1]$ à fixer. Plus ce paramètre est proche de 0 et plus le modèle est parcimonieux.
La loi a posteriori s'écrit elle aussi comme un produit. En effet, en notant $g = \phi * \gamma$:

$$\Pi(\theta|X) = \prod_{i=1}^n \frac{[(1-\alpha)\delta_0(\theta_i) + \alpha\gamma(\theta_i)]\phi(X_i - \theta_i)}{(1-\alpha)\phi(X_i) + \alpha g(X_i)}.$$

On obtient donc :

$$\Pi(\theta|X) = \prod_{i=1}^{n}(1-\tilde{\alpha})\delta_0(\theta_i) + \tilde{\alpha}\psi_{X_i}(\theta_i), \tag{2.3}$$

avec $\tilde{\alpha} = \frac{\alpha g(X_i)}{(1-\alpha)\phi(X_i)+\alpha g(X_i)}$ et la densité $\psi_{X_i} = \frac{\phi(X_i-\cdot)\gamma(\cdot)}{g(X_i)}$.

Johnstone et Silverman montrent alors que la médiane a posteriori est un estimateur de type seuillage, mais que ce n'est pas le cas malheureusement pour la moyenne a posteriori.

2.2.1 Johnstone et Silverman et la médiane et la moyenne a posteriori

Dans [1], il est mis en avant qu'il est bien plus avantageux pour la densité γ d'être une densité à queues suffisamment lourdes, d'où aussi le choix dans la suite d'une densité de Laplace plutôt qu'une loi normale standard. Le fait que ceci soit nécessaire est prouvé dans [3] (pour la mesure a posteriori complète).
Précisément, on suppose pour ce paragraphe que la densité γ dans la formule (2.2) vérifie

$$\sup_{u>0} |\frac{d}{du}\log\gamma(u)| = \Lambda < \infty.$$

On obtient alors que $\forall u > 0$, $\log\gamma(u) \geq \log\gamma(0) - \Lambda u$ et donc, $\forall u > 0$: $\gamma(u) \geq \gamma(0)e^{-\Lambda|u|}$, ce qui élimine le choix de γ gaussienne.

On supposera de plus que $u^2\gamma(u)$ est bornée pour tout u et qu'il existe $\kappa \in [1,2]$ tel que, quand $y \to \infty$:

$$\frac{1}{\gamma(y)}\int_y^\infty \gamma(u)du \asymp y^{\kappa-1}.$$

(pour deux suites u_n et v_n, on note $u_n \asymp v_n$ s'il existe a et b des réels strictement positifs tels que pour n assez grand on ait $av_n \leq u_n \leq bv_n$)

Définition 1. *Une fonction $\delta(x,t)$ de $\mathbb{R}\times\mathbb{R}^+$ dans $\mathbb{R}$ est une règle de rétrécissement si pour tout $t \geq 0$ elle est croissante et antisymétrique en x et que pour tout $x \geq 0$ on a $0 \leq \delta(x,t) \leq x$. C'est une règle de seuillage de seuil t si et seulement si :*

$$\delta(x,t) = 0 \text{ si et seulement si } |x| < t.$$

Elle a de plus la propriété de rétrécissement borné si et seulement s'il existe une constante b telle que pour tous x et t, $|\delta(x,t) - x| \leq t + b$.

Johnstone et Silverman ont montré que la médiane a posteriori dans ce cadre était une règle de seuillage avec la propriété de rétrécissement borné. Bien que n'étant pas une règle de seuillage, la moyenne a posteriori a la propriété de rétrécissement borné.

Comment choisir le paramètre α de la formule (2.2) ?

Il y a plusieurs moyens de choisir α. Tout d'abord, si l'on considère une approche seuillage, on aimerait utiliser un seuil $t_n = \sqrt{2\log(n/p_n)}$. Pour cela, si l'on utilise la médiane a posteriori, en établissant le lien entre α et le seuil t de la médiane, on prendrait α de l'ordre de p_n/n. Ceci n'est pas possible étant donné que le paramètre de sparsité p_n est inconnu !

La première possibilité serait de faire le choix $\alpha = 1/n$, qui donne un seuil en $\sqrt{2\log n}$. C'est ce que j'ai fait pendant mon stage de M2, on obtient des résultats satisfaisants dans le sens où les estimateurs considérés atteignent la vitesse minimax à un facteur p_n près dans le $\log$, mais ce qui est quand même correct en prenant p_n sous la forme d'une puissance de n.

La seconde possibilité serait d'utiliser une approche complètement bayésienne, c'est-à-dire de considérer qu'α suit lui aussi une certaine loi a priori.

La troisième possibilité est d'estimer α par maximum de vraisemblance, c'est-à-dire de prendre $\alpha = \hat{\alpha}$ comme étant la valeur maximisant la quantité suivante (la vraisemblance du modèle) : $\prod_{i=1}^{n}((1-\alpha)\phi(X_i)+\alpha g(X_i))$. C'est ce que font Johnstone et Silverman. Ils obtiennent alors le résultat suivant.

Théorème 2. *Soit $\delta(x,t)$ une règle de seuillage avec la propriété de rétrécissement borné. On note $\hat{t}$ le seuil de la médiane a posteriori correspondant au maximiseur $\hat{\alpha}$ de la vraisemblance. Alors l'estimateur $\hat{\theta}$ tel que $\theta_i = \delta(X_i, \hat{t})$ pour tout $i \in \{1, \ldots, n\}$ atteint la vitesse minimax.*

Ce résultat est dit adaptatif au sens où, même sans connaître la valeur p_n intrinsèque au modèle, notre estimateur se rapproche asymtotiquement de θ à la vitesse minimax qui dépend pourtant de p_n !

Perspective : On peut alors tenter de prouver des résultats similaires mais avec un a priori légèrement différent : on ne prend plus γ comme une densité de probabilité, mais on prend γ la mesure de Lebesgue. La loi a priori est alors impropre puisque n'est plus une loi de probabilité, mais la loi a posteriori est, elle, parfaitement définie et ce sont les aspects de la loi a posteriori qui nous fournissent des estimateurs. On peut alors espérer utiliser des raisonnements similaires pour en dériver des résultats.

2.2.2 Veronika Rockova et un Spike and Slab légèrement modifié

Dans [7] est utilisé un a priori légèrement différent, on remplace la masse de Dirac par une densité de probabilité :

$$\theta \sim \Pi := \bigotimes_{i=1}^{n}(1-\alpha)\gamma_1 + \alpha\gamma_2 \tag{2.4}$$

où les densités γ_1 et γ_2 sont celles de loi de Laplace (de densité $\lambda_i/2\exp(-\lambda_i|x|)$) dont les paramètres λ_1 et λ_2 sont différents : le premier est beaucoup plus grand que le second, la première densité ressemble donc à un dirac (en une version continue en quelque sorte) et la seconde à une loi plus diffuse.

Les modes de la loi a posteriori sont alors bien définis (valeurs où la loi a posteriori est maximale), et il y a même un mode unique sous une certaine condition. Veronika Rockova montre alors un résultat de convergence minimax (et donc lui aussi adaptatif) pour ce mode unique.

Perspective : On peut alors tenter de prouver des résultats en utilisant cet a priori pour d'autres aspects de cet a posteriori, comme la médiane. Mais surtout on peut tenter une approche par le maximum de vraisemblance sur cet a priori !

2.3 Un a priori plus général

On commence par tirer une dimension $s \in \{0, 1, \dots, n\}$ suivant une certaine loi $\pi(s)$ sur l'ensemble $\{0, 1, \dots, n\}$.
On tire ensuite un support $S \subset \{1, \dots, n\}$ uniformément sur les ensembles de cardinal $s : \Pi(S) = \frac{\pi(s)}{\binom{n}{s}}$.
On a alors l'a priori général suivant sur θ :

$$\theta \sim \sum_{S \subset \{1,\dots,n\}} \Pi(S) \bigotimes_{i \in S} \gamma \otimes \bigotimes_{i \notin S} \delta_0$$

avec γ une densité de probabilité.

On peut déjà remarquer que, en général, l'a posteriori n'a plus la forme d'un produit, ce qui avait l'avantage de simplifier les démonstrations dans la première partie.
Dans la suite, on dira que π est à décroissance exponentielle s'il existe $C > 0$ et $D < 1$ tels que $\pi(s) \leq D\pi(s-1)$ pour $s > Cp_n$. Si cette condition est satisfaite pour $C = 0$ alors on dira que π est à décroissance exponentielle stricte.

On peut dès lors se demander ce que l'on pourrait choisir pour le *prior* π sur la dimension, en voici un exemple.

Prior binomial

Si π est la loi binomiale $B(n, \alpha)$, le *prior* sur θ est exactement celui de la première partie : on retombe sur l'a priori et l'a posteriori sous forme de produit. Ce *prior* π est à décroissance exponentielle pour $\alpha \lesssim \frac{p_n}{n}$.Le choix $\alpha = \frac{p_n}{n}$ correspond à une valeur oracle, permettant d'atteindre la vitesse minimax. Si on choisit $\alpha = \frac{1}{n}$, on atteint la vitesse minimax si $p_n = n^a$, avec $a < 1$; mais si p_n est d'un ordre plus grand, on pourrait rater la vitesse minimax à un facteur logarithmique près.

2.3.1 Quelques résultats

Dans [2], Ghosal, Ghosh et Van der Vaart démontrent le théorème général suivant :

Théorème 3. *[Théorème général]*
Soit $\mathcal{P}$ un ensemble de mesures de probabilité sur $\mathbb{R}$. Soit d une métrique sur P. Soit $\Pi = (\Pi_n)_{n\in\mathbb{N}}$ une suite de priors sur $\mathcal{P}$ et supposons que les observations X soient des variables réelles i.i.d. de densité p_0, dont on note P_0 la loi associée. Soit $(\varepsilon_n)_{n\in\mathbb{N}}$ une suite de réels strictement positifs telle que $\varepsilon_n \to 0$ et $\sqrt{n}\varepsilon_n \to \infty$ quand $n \to \infty$.
On suppose qu'il existe des constantes C et L telles que :

$$\Pi(p \in \mathcal{P}; -E_{p_0}[\log(\frac{p}{p_0}(X))] \leq \varepsilon_n^2, E_{p_0}[\log(\frac{p}{p_0}(X))^2] \leq \varepsilon_n^2) \geq e^{-Cn\varepsilon_n^2}$$

et

$$\Pi(\mathcal{P} \setminus \mathcal{P}_n) \leq Le^{-(C+4)n\varepsilon_n^2}$$

pour une suite $\mathcal{P}_n \subset \mathcal{P}$ telle qu'on puisse trouver des tests $\psi_n = \psi(X_1, \ldots, X_n)$ tels que pour tous les $n \in \mathbb{N}$ et $M > 0$ assez grands :

$$E_{p_0}[\psi_n] \to 0 \text{ et } \sup_{d(p,p_0)\geq M\varepsilon_n} E_p[1-\psi_n] \leq Le^{-(C+4)n\varepsilon_n^2}.$$

Alors $\Pi(d(p,p_0) > M\varepsilon_n|X) \to 0$ quand $n \to \infty$ en $P_0^{\mathbb{N}}$-probabilité.

Il ne s'applique malheureusement pas dans notre cadre puisque nos observations X_i ne sont pas i.i.d.. Il existe cependant une version non i.i.d. de ce théorème, que l'on ne peut cependant utiliser telle quelle dans notre cas du fait d'une difficulté technique liée au fait que la classe $\ell_0[p_n]$ dans laquelle vit notre paramètre θ_0 n'est pas bornée.

Pour obtenir leurs résultats, Ismaël Castillo et Aad Van der Vaart démontrent dans [3] au préalable un résultat sur la dimension :

Théorème 4. *[Dimension]*
Si π est à décroissance exponentielle et si γ est centrée avec un moment d'ordre 2 fini, alors il existe $M > 0$ tel que quand $n \to \infty$:

$$\sup_{\theta_0\in\ell_0[p_n]} E_{\theta_0}[\Pi(|S_\theta| > Mp_n|X)] \to 0$$

Ils obtiennent alors le théorème suivant.

Théorème 5. *Si π est à décroissance exponentielle, si γ est centrée, de moment d'ordre 2 fini et de la forme e^h avec h telle que pour tous x et y réels, il existe une constante C réelle strictement positive telle que $|h(x) - h(y)| \leq C(1 + |x-y|)$, alors, avec r_n tel que*

$$r_n^2 \geq \max(p_n \log(\frac{n}{p_n}), \log(\frac{1}{\pi(p_n)}))$$

et $M > 0$ assez grand, quand $n \to \infty$:

$$\sup_{\theta_0\in\ell_0[p_n]} E_{\theta_0}[\Pi(\|\theta-\theta_0\|^2 > Mr_n^2|X)] \to 0.$$

Perspective : Cette approche adaptée de démonstration du théorème de Ghosal, Ghosh et Van der Vaart est utilisée dans [5] pour trouver des résultats de convergence pour la loi a posteriori mais par une approche qui part du maximum de vraisemblance (comme pour Johnstone et Silverman). Ne pourrait-on pas alors étendre les résultats de Johnstone et Silverman sur le *prior* Spike and Slab à la mesure a posteriori entière ? On ne peut cependant pas utiliser tels quels les résultats de Rousseau et Szabo que l'on doit adapter à notre cas spécifique où l'a priori contient une partie atomique. L'a priori proposé par Rockova semble, lui, rentrer dans le cadre des travaux de Rousseau et Szabo.

2.4 Des classes plus générales

Au lieu de travailler uniquement sur $\ell_0[p_n]$, tous ces auteurs ont également travaillé sur des classes plus générales :

$$\ell_q[p_n] = \{\theta; \sum_{i=1}^{n} |\theta_i|^q \leq n(p_n/n)^q\}$$

Les auteurs ont prouvé des résultats de convergence minimax sur ces classes pour leurs estimateurs, et adaptatifs au sens où, sans connaître la valeur de q de la classe dans laquelle notre paramètre θ à estimer vit réellement, nos estimateurs convergent à une vitesse minimax dépendant de q.

2.5 Une dernière perspective

On pourra ensuite étudier le cas d'observations $X = M\theta + \varepsilon$, où M est une matrice de design.

Bibliographie

[1] I. M. JOHNSTONE et B. W. SILVERMAN, *Needles and straw in haystacks : empirical Bayes estimates of possibly sparse sequences*, (2004). Ann. Statist. **32** 1594-1649.

[2] S. GHOSAL, J. K. GHOSH et A. W. VAN DER VAART, *Convergence rates of posterior distributions*, (2000). Ann. Statist. **28(2)** 500-531.

[3] I. CASTILLO et A. W. VAN DER VAART, *Needles and straw in haystacks : posterior concentration for possibly sparse sequences* , (2012). Ann. Statist. **40(4)** 2069-2101.

[4] I. CASTILLO et A. W. VAN DER VAART, *Supplement to 'Needles and straw in haystacks : posterior concentration for possibly sparse sequences'*, (2012). Ann. Statist. **40(4)**.

[5] J. ROUSSEAU et B. SZABO, *Asymptotic behaviour of the empirical bayes posteriors associated to maximum marginal likelihood estimator* , (2015) Preprint

[6] D. L. DONOHO, I. M. JOHNSTONE, J.C. HOCH et A.S. STERN, *Maximum entropy and the nearly black object* , (1992)J.R.Stat.Soc.Ser.B Stat. Methodol. **54** 41-81.

[7] V. ROCKOVA, *Bayesian estimayion of sparse signals with a Continuous Spike and Slab Prior*, (2015). Ann. Statist.

Annexe A

Rapport de stage de L3

Rapport de stage

Introduction

Mon stage s'est déroulé au Laboratoire Environnement et Minéralurgie de Vandoeuvre les Nancy sous la direction de Mr Alain Mailfert. Il a constitué tout d'abord en la lecture avant mon stage de deux thèses sur mon sujet, en une vérification de calculs déjà effectués puis en un travail de recherche sur mon sujet.

Remerciements

Je remercie Mr Mailfert pour son aide précieuse, sa gentillesse et sa générosité tout au long de mon stage. Je remercie aussi Delphine Martin sans qui je n'aurai jamais pu faire ce stage, et enfin tout le personnel du LEM qui m'a permis de travailler dans l'ambiance la plus agréable qu'il soit.

Plan :

I)Présentation du LEM et de ses activités

Le Laboratoire Environnement et Minéralurgie est un centre de recherche rattaché à la fois au CNRS (centre national de la recherche scientifique) et à l'INPL (institut national polytechnique de Lorraine) qui a pour principales missions de valoriser les ressources en minerais et les minéraux industriels, de rechercher de nouvelles techniques de valorisation et de traitement des minéraux et de veiller à développer des procédés en génie minéral qui soient à la fois propres et sûrs. Les différents chercheurs du LEM disposent d'un dispositif industriel unique en milieu universitaire, la station STEVAL (STation Expérimentale de VALorisation des matières premières et des substances résiduaires) où sont mis en oeuvres de nombreux procédés de séparation des minéraux (séparation magnétique, électrostatique, hydraulique,...). Le LEM est ainsi en partenariat avec de nombreux industriels notamment des opérateurs miniers, des traiteurs d'eau, ou encore des producteurs de rejets.

II)Problématique

On parle de micropesanteur lorsque la résultante des forces gravitationnelles et inertielles auxquelles est soumis un corps est très faible, ce qui ne peut être obtenu qu'en certains points de l'espace. Recréer sur Terre des conditions d'apesanteur permet notamment d'étudier le comportement des fluides dans l'espace, différent du comportement observé sur Terre puisqu'il n'y a plus de convexion thermique. De même, la combustion s'effectue différemment dans l'espace, la forme de la flamme d'une bougie devient par exemple quasi-sphérique en microgravité.
Il existe plusieurs moyens de compenser la pesanteur sur Terre, le premier auquel on peut penser est d'utiliser une accélération suffisante, il existe pour cela des tours à chute libre, qui n'offrent pas un temps d'expérimentation suffisant, des fusées sondes qui offrent des temps d'expérimentation d'une quinzaine de minutes environ. On peut aussi réaliser des vols paraboliques en avion, avec des durées de microgravité de 25 secondes environ.
On peut aussi s'aider de champs magnétiques pour compenser parfaitement la pesanteur. On souhaiterait pour cela trouver des distributions de courant qui génèreraient une force magnétique constante sur une zone utile qu'on veut la plus large possible. Cette utilisation des champs magnétiques s'avèrerait aussi utile pour la séparation magnétique des minéraux dans le cadre des activités du LEM. L'objet de mon stage est précisément l'étude de la force magnétique que l'on appellera par la suite $\vec{G}$. Un champ magnétique $\vec{B}$ exerce sur un matériau de suceptibilité magnétique χ une force :

$$\frac{\overrightarrow{df}}{dV} = \frac{\chi}{\mu_0}\overrightarrow{grad}(\| \vec{B} \|^2)$$

Avec μ_0 la perméabilité du vide. On aura ainsi compensation de la gravité si on a $\vec{G} = \frac{-2\mu_0\rho}{\chi}\vec{g}$, où ρ est la masse volumique du matériau à faire léviter et $\vec{g}$ est la pesanteur. On s'intéressera donc par la suite à la force $\vec{G} = \overrightarrow{grad}(\| \vec{B} \|^2)$, le reste étant des constantes. Il a déjà été prouvé par André Colteu et Alain Mailfert à l'aide des équations de Maxwell vérifiées par le champ magnétique dans le vide

$$\begin{cases} div\vec{B} = 0 \\ \overrightarrow{rot}\vec{B} = \vec{0} \end{cases}$$

qu'il était impossible d'avoir une force magnétique uniforme en tout point d'un espace en trois dimensions
Deux problèmes ont été soulevés : le premier étant de trouver une distribution de courants permettant d'obtenir un $\vec{G}$ constant sur toute une droite verticale de l'espace (cas vertical), le second étant d'obtenir un $\vec{G}$ constant sur tout un plan horizontal de l'espace (cas horizontal). Mon travail durant le stage a consisté tout d'abord à vérifier les résultats trouvés pour le cas vertical par Clément Lorin puis à étudier le cas horizontal. Dans tout ce

qui suit, on travaille dans un repère orthonormé $(\vec{e_x}, \vec{e_y}, \vec{e_z})$ en coordonnées sphériques (r, θ, ϕ).

III)Travaux effectués

1)Vérification des calculs en cas vertical

On travaille ici en géométrie axisymétrique (invariance par rotation autour de l'axe Oz), de ce fait le champ ne dépend pas de ϕ. Afin d'obtenir un $\vec{G}$ constant sur l'axe Oz (ou encore l'axe $\theta = 0$), soit $\vec{G}(r,0) = G_1\vec{e}_z$, les dérivées successives de $\vec{G}(r,0)$ doivent être nulles, ainsi $\forall n \in N$, on doit avoir :

$$\frac{\delta^n Gr(r,0)}{\delta r^n} = 0$$

On utilise alors la décomposition en harmoniques sphériques du champ : dans une sphère de rayon R_0, le champ est défini par une fonction des polynômes de Legendre P_n et P_n^1 :

$$\vec{H} = \sum_{n=1}^{\infty} -nC_n r^{n-1} P_n(cos\theta)\vec{e}_r + \sum_{n=1}^{\infty} -C_n r^{n-1} P_n^1(cos\theta)\vec{e}_\theta$$

On a $\vec{B} = \mu_0\vec{H}$, $\vec{G} = \overrightarrow{grad}(\parallel \vec{B} \parallel^2)$, d'où :

$$\vec{G} = 2\mu_0^2\Big(\sum_{n=1}^{\infty} nC_n r^{n-1} P_n(cos\theta) \sum_{n=1}^{\infty} nC_n(n-1)r^{n-2}P_n(cos\theta)$$

$$+\sum_{n=1}^{\infty} C_n r^{n-1} P_n^1(cos\theta) \sum_{n=1}^{\infty} (n-1)C_n r^{n-2} P_n^1(cos\theta)\Big)\vec{e}_r$$

On sait de plus que $P_n(1) = 1$ et $P_n^1(1) = 0$ $\forall n \in N$, donc :

$$Gr(r,0) = 2\mu_0^2 \sum_{k=3}^{\infty} \Big(\sum_{n=1}^{k-2} n(k-n)(k-n-1)C_n C_{k-n}\Big) r^{k-3}$$

Notre condition de départ nous donne alors, $\forall k > 3$:

$$\sum_{n=1}^{k-2} n(k-n)(k-n-1)C_n C_{k-n} = 0$$

Si on appelle B_1 la valeur du champ à l'origine, on a tout de suite $C_1 = \frac{B_1}{\mu_0}$ et $C_2 = \frac{G_1}{4\mu_0 B_1}$.

On peut alors démontrer par une (plutôt longue) récurrence que, $\forall n \in N$, $C_n = \frac{2(-1)^n(2n-5)!!G_1^{n-1}}{(2n)!!\mu_0 B_1^{2n-3}}$

On a alors l'expression de $\vec{H}$ sur l'axe : $\vec{H}_{axe} = \sum_{n=1}^{\infty} -nC_n r^{n-1}\vec{e}_r = \sum_{n=1}^{\infty} -n\frac{2(-1)^n(2n-5)!!G_1^{n-1}}{(2n)!!\mu_0 B_1^{2n-3}} r^{n-1}\vec{e}_r$

Il faut alors remarquer que cette somme est exactement le développement en série entière de la fonction $x->(1+x)^{\frac{1}{2}}$:

$$(1+x)^{\frac{1}{2}} = \sum_{n=0}^{\infty} \frac{(-1)^{n-1}(2n-3)!!}{(2n)!!}x^n = \sum_{n=1}^{\infty} \frac{(-1)^n(2n-5)!!}{(2n-2)!!}x^{n-1} = \sum_{n=1}^{\infty} \frac{2n(-1)^n(2n-5)!!}{(2n)!!}x^{n-1}$$

D'où

$$\vec{H}_{axe} = -\frac{B_1}{\mu_0}(1+\frac{G_1 r}{B_1^2})^{\frac{1}{2}}\vec{e}_r$$

D'où une force constante le long de l'axe Oz dans une sphère de rayon $R_0 = \frac{B_1^2}{G_1}$. Ce résultat est fort, dans la mesure où c'est la première fois qu'on réussit à montrer qu'une force magnétique peut être constante en un nombre infini de points de l'espace.

âge

2)Etude du cas horizontal

En géométrie cylindrique (invariance par translation), la méthode de la décomposition du champ en harmoniques a permis de résoudre le cas vertical et horizontal. On sait ainsi qu'avec une distribution de courants sinusoïdale , on peut réussir à générer une force magnétique constante sur tout le plan xOy. Il a d'ailleurs été prouvé que dans ce cas $\vec{G}$ varie exponentiellement suivant z. On cherche désormais à obtenir une force constante sur le plan xOy dans le cas axisymétrique, c'est-à-dire lorsqu'on a invariance par rotation autour de l'axe Oz. La méthode des harmoniques devient alors beaucoup plus compliquée, mais j'ai quand même souhaité chercher en utilisant cette méthode, sans résultat jusqu'à présent.

i)Méthodes des harmoniques

Cette fois, la condition sur les dérivées successives de $\vec{G}$ s'exprime comme suit, et ce $\forall n \in N$:

$$\frac{\delta^n G_r(r,\frac{\pi}{2})}{\delta r^n} = \frac{\delta^n G_\theta(r,\frac{\pi}{2})}{\delta r^n} = 0$$

Les choses se compliquent puisque cette fois G_θ n'est pas nul, et comme on est placé en $\theta = \frac{\pi}{2}$, les $P_n^1(cos\theta)$ ne sont pas nuls non plus. En effet, la valeur des P_n^p en 0 dépend de la parité de n :

$$\begin{cases} P_{2n+1}(0) = 0 \\ P_{2n}(0) = \frac{(-1)^n(2n)!}{2^{2n}(n!)^2} \\ P_{2n}^1(0) = 0 \\ P_{2n+1}^1(0) = \frac{(-1)^n(2n+1)!}{2^{2n+1}n!(n+1)!} \end{cases}$$

La condition de départ nous fournit après de longs calculs 3 conditions. On commence tout d'abord par calculer les 2 composantes du vecteur $\vec{G}$, on trouve :

$$G_r(r,\theta) = 2\mu_0^2\Big(\sum_{n=3}^{\infty}\sum_{k=1}^{n-2} kC_kP_k(cos\theta)(n-k)(n-k-1)C_{n-k}P_{n-k}(cos\theta)r^{n-3}$$

$$+\sum_{n=3}^{\infty}\sum_{k=1}^{n-2} kC_kP_k^1(cos\theta)(n-k-1)C_{n-k}P_{n-k}^1(cos\theta)r^{n-3}\Big)$$

et

$$G_\theta(r,\theta) = 2\mu_0^2\Big(\sum_{n=3}^{\infty}\sum_{k=1}^{n-2} kC_kP_k(cos\theta)(n-k)C_{n-k}(P_{n-k}(cos\theta))'r^{n-3}$$

$$+\sum_{n=3}^{\infty}\sum_{k=1}^{n-2} C_kP_k^1(cos\theta)C_{n-k}(P_{n-k}^1(cos\theta))'r^{n-3}\Big)$$

On remplace alors θ par $\frac{\pi}{2}$:

$$G_r(r,\frac{\pi}{2}) = 2\mu_0^2\Big(\sum_{n=3}^{\infty}\sum_{k=1}^{n-2} kC_kP_k(0)(n-k)(n-k-1)C_{n-k}P_{n-k}(0)r^{n-3}$$

$$+\sum_{n=3}^{\infty}\sum_{k=1}^{n-2} kC_kP_k^1(0)(n-k-1)C_{n-k}P_{n-k}^1(0)r^{n-3}\Big)$$

On sépare en pairs et impairs (on note E : $R->N$ la fonction partie entière inférieure) :

$$G_r(r,\frac{\pi}{2}) = 2\mu_0^2\Big(\sum_{n=3}^{\infty}\sum_{k=1}^{E(\frac{n-2}{2})} 2kC_{2k}P_{2k}(0)(n-2k)(n-2k-1)C_{n-2k}P_{n-2k}(0)r^{n-3}$$

$$+\sum_{n=3}^{\infty}\sum_{k=1}^{E(\frac{n-3}{2})} (2k+1)C_{2k+1}P_{2k+1}^1(0)(n-2k-2)C_{n-2k-1}P_{n-2k-1}^1(0)r^{n-3}\Big)$$

$$G_r(r,\frac{\pi}{2}) = 2\mu_0^2\Big(\sum_{n=2}^{\infty}\sum_{k=1}^{n-1} 2kC_{2k}P_{2k}(0)(2n-2k)(2n-2k-1)C_{2n-2k}P_{2n-2k}(0)r^{2n-3}$$

$$+\sum_{n=2}^{\infty}\sum_{k=0}^{n-2} (2k+1)C_{2k+1}P_{2k+1}^1(0)(2n-2k-2)C_{2n-2k-1}P_{2n-2k-1}^1(0)r^{2n-3}\Big)$$

$$G_r(r,\frac{\pi}{2}) = 2\mu_0^2\Big(\sum_{n=2}^{\infty}\sum_{k=1}^{n-1} C_{2k}C_{2n-2k}\frac{2k(2n-2k)(2n-2k-1)(-1)^n(2k)!(2n-2k)!}{2^{2n}(k!)^2((n-k)!)^2}r^{2n-3}$$

$$+\sum_{n=2}^{\infty}\sum_{k=0}^{n-2} C_{2k+1}C_{2n-2k-1}\frac{(-1)^{n-1}(2k+1)!(2n-2k-1)!(2n-2k-2)}{2^{2n}k!(n-k)!(k+1)!(n-k-1)!}r^{2n-3}\Big)$$

On obtient alors comme condition que, $\forall n > 2$,

$$\sum_{k=1}^{n-1} C_{2k}C_{2n-2k}\frac{2k(2n-2k)(2n-2k-1)(-1)^n(2k)!(2n-2k)!}{(k!)^2((n-k)!)^2}$$

$$+\sum_{k=0}^{n-2} C_{2k+1}C_{2n-2k-1}\frac{(-1)^{n-1}(2k+1)!(2n-2k-1)!(2n-2k-2)}{k!(n-k)!(k+1)!(n-k-1)!} = 0$$

En faisant de même pour G_θ, on obtient les deux conditions que, $\forall n > 2$:

$$\sum_{k=0}^{n-1} C_{2k+1}C_{2n-2k}\frac{(2k+1)!(2n-2k+2)!}{k!(n-k+1)!(k+1)!(n-k-1)!} = 0$$

et :

$$\sum_{k=0}^{n-2} C_{2k+1}C_{2n-2k-1}\frac{(2k+1)!(2n-2k)!}{k!(n-k)!(k+1)!(n-k-1)!} = 0$$

J'ai alors tenté, à l'aide des premiers termes de la suite des C_n, d'établir une formule de récurrence, mais en vain. On trouve pour les 7 premières harmoniques :

$$\begin{cases} \mathrm{C}_1 = \frac{B_1}{\mu_0} \\ \mathrm{C}_2 = \frac{G_1}{4\mu_0 B_1} \\ \mathrm{C}_3 = \frac{G_1^2}{48\mu_0 B_1^3} \\ \mathrm{C}_4 = 0 \\ \mathrm{C}_5 = \frac{-G_1^4}{3840\mu_0 B_1^7} \\ \mathrm{C}_6 = \frac{-G_1^5}{46080\mu_0 B_1^9} \\ \mathrm{C}_7 = \frac{G_1^6}{172032\mu_0 B_1^1 1} \end{cases}$$

J'ai ensuite cherché à tester des formes particulières de variation de $\vec{G}$ en fonction de z, entre autres des fonctions périodiques, les fonctions puissance et une variation en exponentielle comme le suggéraient les résultats trouvés en géométrie cylindrique, mais aucune ne vérifiait les équations de Maxwell.

ii)Démonstration de l'impossibilité pour la force magnétique de varier exponentiellement selon z

J'ai démontré qu'il est impossible, en géométrie axisymétrique, de trouver un champ $\vec{B}$ tel que $\overrightarrow{grad}(\parallel \vec{B} \parallel^2)$ varie exponentiellement suivant z et soit dirigé verticalement sur l'axe vertical, tout en étant constant et dirigé verticalement sur le plan $z = 0$:

On travaille ici en coordonnées cylindriques (r, θ, z), en considérant la sphère $B(0, R)$ de rayon R et centrée en l'origine, qui constitue la zone intérieure, tandis que $R^3 \setminus B(0, R)$ constitue la zone extérieure.

On veut sur l'axe $\overrightarrow{grad}(\parallel \vec{B} \parallel^2) = \alpha \exp(2\beta z)\vec{e}_z$, $(\alpha, \beta) \in R^{*2}$

Sur l'axe (Oz), on a $\vec{B} = B_z \vec{e}_z$

$\parallel \vec{B} \parallel^2 = B_z^2$

$\overrightarrow{grad}(\parallel \vec{B} \parallel^2) = \frac{\delta B_z^2}{\delta z}\vec{e}_z + \frac{\delta B_z^2}{\delta r}\vec{e}_r$

Donc $\frac{\delta B_z^2}{\delta z} = \alpha \exp(2\beta z)$, soit $\parallel \vec{B} \parallel^2 = \frac{\alpha}{2\beta}\exp(2\beta z) + C$, $C \in R$

D'où, sur l'axe, $\vec{B} = (\lambda \exp(\beta z) + C)\vec{e}_z$, $\lambda \in R^*$, $C \in R$

$\vec{B}$ vérifie les équations de Maxwell :

$$\begin{cases} div\vec{B} = 0 \\ \overrightarrow{rot}\vec{B} = \vec{0} \end{cases}$$

et dérive donc d'un potentiel magnétique V_m :
$\vec{B} = -\mu_0 \overrightarrow{grad} V_m$

donc $-\frac{\delta V_m}{\delta r}\vec{e}_r - \frac{\delta V_m}{\delta z}\vec{e}_z = (\frac{\lambda}{\mu_0}\exp(\beta z) + C)\vec{e}_z$ sur l'axe.

Donc $V_m = -\frac{\lambda}{\mu_0 \beta}\exp(\beta z) - Cz + C'$ sur l'axe.

On cherche alors $V_m(r,z)$ sous la forme $f(z)g(r) = (-\frac{\lambda}{\mu_0\beta}\exp(\beta z) - Cz + C')g(r)$

V_m doit alors vérifier $\Delta V_m = 0$, ce qui revient à :

$$\frac{\delta^2 V_m}{\delta r^2} + \frac{1}{r}\frac{\delta V_m}{\delta r} + \frac{\delta^2 V_m}{\delta z^2} = 0$$
$$\Leftrightarrow \frac{1}{g(r)}(\frac{d^2 g}{dr^2} + \frac{1}{r}\frac{dg}{dr}) + \frac{1}{f(z)}\frac{d^2 f}{dz^2} = 0$$

donc $\exists K \in R; \frac{1}{f(z)}\frac{d^2 f}{dz^2} = K$

soit $\frac{d^2 f}{dz^2} = Kf(z)$

ce qui implique que $-\frac{\lambda}{\mu_0\beta}\exp(\beta z)(\beta^2 - K) = C' - Cz$, $\forall z$ non nul

ce qui impose $K = \beta^2$ et $C = C' = 0$

On est finalement ramené à :

$$\frac{d^2 g}{dr^2} + \frac{1}{r}\frac{dg}{dr} + \beta^2 g(r) = 0$$

qui est une équation de Bessel :

$$\frac{d^2 y}{dx^2} + \frac{1}{x}\frac{dy}{dx} + (1 - \frac{\nu^2}{x^2})y(x) = 0$$

de solution $y(x) = \gamma J_\nu(x) + \delta Y_\nu(x)$, $(\gamma, \delta) \in R^2$

Ici, $\nu = 0$ et $g(r) = \gamma J_0(\beta r) + \delta Y_0(\beta r)$, $(\gamma, \delta) \in R^2$

Le fait que Y_0 tende vers moins l'infini en 0 nous impose premièrement que $\delta = 0$

On a alors $V_m(r,z) = -\frac{\lambda\gamma}{\mu_0\beta}\exp(\beta z)J_0(\beta r)$

Calculons dès à présent le champ $\vec{B} = -\mu_0\overrightarrow{grad}V_m$, en sachant que $J_0'(r) = -J_1(r)$:

$$\vec{B} = \lambda\gamma\exp(\beta z)(J_0(\beta r)\vec{e}_z - J_1(\beta r)\vec{e}_r)$$

donc $\parallel \vec{B} \parallel^2 = \frac{\alpha\gamma^2}{2\beta}\exp(2\beta z)(J_0^2(\beta r) + J_1^2(\beta r))$

donc $\vec{G} = \alpha\gamma^2\exp(2\beta z)\Big((J_0^2(\beta r)+J_1^2(\beta r))\vec{e}_z + J_1(\beta r)(-J_0(\beta r)+J_1'(\beta r))\vec{e}_z\Big)$

Pour que $\vec{G}$ soit constant et dirigé verticalement sur $\{z = 0\} \cap B(0,R)$, il faut que, $\forall r \in [0;R]$:

$$\left(\frac{1}{2\beta}\frac{d}{dr}(J_0^2(\beta r) + J_1^2(\beta r)) = \right)J_1(\beta r)(-J_0(\beta r) + J_1'(\beta r)) = 0$$

Par la formule $2J_\nu'(x) = J_{\nu-1} - J_{\nu+1}$, on a $J_1'(\beta r) = \frac{\beta}{2}(J_0(\beta r) - J_2(\beta r))$,

et l'équation $J_1(\beta r)(-J_0(\beta r) + J_1'(\beta r)) = 0$ équivaut à

ou bien $J_1(\beta r) = 0$, ce qui n'est pas vérifié en tout point de $[0;R]$,

ou bien $J_0(\beta r) = \frac{\beta}{\beta-2}J_2(\beta r)$, $\forall r \in [0,R]$

Mais au vu des deux expressions :

$J_0(\beta r) = \sum_{k=0}^{\infty}\frac{(-1)^k}{(k!)^2}(\frac{\beta r}{2})^{2k}$ et $J_2(\beta r) = (\frac{\beta r}{2})^2\sum_{k=0}^{\infty}\frac{(-1)^k}{k!(k+2)!}(\frac{\beta r}{2})^{2k}$
on remarque aisément que ceci n'est pas vérifié non plus

On a ici démontré qu'il est impossible, en géométrie axisymétrique, de trouver une distribution de courant telle que la force magnétique qui en résulte soit constante en tout point du plan xOy tout en variant exponentiellement selon z, et donc que ce n'est pas la peine de chercher $\vec{G}$ en exponentielle comme c'était le cas dans le cadre d'une géométrie cylindrique.

iii) Dernière tentative à l'aide des équations de Maxwell

On se place encore une fois en coordonnées cylindriques.
La dernière tentative a consisté à écrire un système d'équations d'inconnues les deux composantes du champ B_r et B_z. Tout d'abord, on veut $\vec{G} = \overrightarrow{grad}(\parallel \vec{B} \parallel^2)$ en tant que fonction ne dépendant que de z, $\vec{G} = G_0(z)\vec{e}_z$. On a $\parallel \vec{B} \parallel^2 = B_z^2 + B_r^2$. Donc

$$\overrightarrow{grad}(\parallel \vec{B} \parallel^2) = 2\left(B_z \frac{\delta B_z}{\delta z} + B_r \frac{\delta B_r}{\delta z}\right)\vec{e}_z + 2\left(B_z \frac{\delta B_z}{\delta r} + B_r \frac{\delta B_r}{\delta r}\right)\vec{e}_r$$

Le champ vérifie de plus les équations de Maxwell :

$$\begin{cases} div\vec{B} = 0 \\ \overrightarrow{rot}\vec{B} = \vec{0} \end{cases}$$

On obtient ainsi le système :

$$\begin{cases} B_z \frac{\delta B_z}{\delta z} + B_r \frac{\delta B_r}{\delta z} = G_0(z) \\ B_z \frac{\delta B_z}{\delta r} + B_r \frac{\delta B_r}{\delta r} = 0 \\ \frac{1}{r} B_r + \frac{\delta B_r}{\delta r} + \frac{\delta B_z}{\delta z} = 0 \\ \frac{\delta B_r}{\delta z} = \frac{\delta B_z}{\delta r} \end{cases}$$

d'inconnues $B_z(r, z)$ et $B_r(r, z)$, que je n'ai pas réussi à résoudre..

IV)Notions mathématiques utilisées

1)Polynômes de Legendre

Les polynômes de Legendre sont solutions de l'équation de Legendre :

$$\frac{d}{dx}\left((1-x^2)\frac{dy}{dx}\right)+n(n+1)y=0$$

On définit donc les polynômes de Legendre :

$$\begin{cases} \frac{d}{dx}\left((1-x^2)\frac{dP_n}{dx}\right)+n(n+1)P_n(x)=0 \\ P_n(1)=1 \end{cases}$$

On utilise aussi 2 autres formules pour définir P_n :

$$\begin{cases} (n+1)P_{n+1}(x)=(2n+1)xP_n(x)-nP_{n-1}(x) \\ P_0(x)=1 \\ P_1(x)=x \end{cases}$$

et :

$$P_n(x)=\frac{1}{2^n n!}\frac{d^n}{dx^n}\left((x^2-1)^n\right)$$

On peut aussi définir les polynômes de Legendre sous forme de somme :

$$P_n(x)=\sum_{k=0}^{E(\frac{n}{2})}(-1)^k C_n^k C_{2n-2k}^n x^{n-2k}$$

(d'où $P_{2n}(0)=\frac{(-1)^n(2n)!}{2^{2n}(n!)^2}$)

On a de plus la formule, $\forall(n,m)\in N^2$:

$$P_n^m(x)=(1-x^2)^{\frac{m}{2}}\frac{d^m P_n(x)}{dx^m}$$

On dispose aussi d'une formule pour trouver la parité des P_n^m :

$$P_n^m(-x)=(-1)^{n+m}P_n^m(x)$$

2)Double factorielle et fonction factorielle

On définit la double factorielle de façon récurrente : $\begin{cases} 0!! = 1 \\ 1!! = 1 \\ \text{n}\,!\,!\text{=n((n-2)}\,!\,!\text{)}\ \forall n > 1 \end{cases}$

On a immédiatement les formules :

$$(2n)!! = 2^n n!$$

et

$$(2n+1)!! = \frac{(2n+1)!}{(2n)!!} = \frac{(2n+1)!}{2^n n!}$$

La fonction factorielle peut être prolongée sur l'ensemble des complexes grçace à la fonction gamma d'Euler :

$$z! = \Gamma(z+1) = \int_0^\infty x^z e^{-x} dx$$

On définit la fonction gamma d'Euler comme suit :

$$\frac{1}{\Gamma(z)} = z e^{\gamma z} \prod_{m=1}^{\infty} (1 + \frac{z}{m}) e^{\frac{-z}{m}}$$

avec γ la constante d'Euler.On peut alors prouver un certain nombre de propriétés :

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$$

$$\Gamma(z+1) = z\Gamma(z)$$

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{sin(\pi z))}$$

$$\Gamma(1-z) = -z\Gamma(-z)$$

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}$$

$$\Gamma(n+\frac{1}{2}) = \sqrt{\pi}\frac{(2n-1)!!}{2^n}$$

3)Equations de Bessel [4]

Elles prennent la forme suivante :

$$\frac{d^2y}{dx^2} + \frac{1}{x}\frac{dy}{dx} + (1 - \frac{\nu^2}{x^2})y(x) = 0$$

Ces équations ont pour solution $y(x) = \delta J_\nu(x) + \gamma Y_\nu(x)$, $(\delta, \gamma) \in R^2$

Avec $J_\nu(z) = (\frac{z}{2})^\nu \sum_{r=0}^{\infty} \frac{(-1)^r}{r!\Gamma(\nu+r+1)}(\frac{z}{2})^{2r}$

Soit, si ν est un entier n : $J_n(z) = (\frac{z}{2})^n \sum_{r=0}^{\infty} \frac{(-1)^r}{r!(r+n)!}(\frac{z}{2})^{2r}$

En particulier, $J_0(z) = \sum_{r=0}^{\infty} (-1)^r (r!)^2 (\frac{z}{2})^{2r}$

On a $Y\nu(z) = \frac{\cos \pi\nu J\nu(z) - J_{-\nu}(z)}{\sin \pi\nu}$

Calculons maintenant :

$$zJ'_\nu(z) = \sum_{r=0}^{\infty} \frac{(-1)^r}{r!\Gamma(\nu+r+1)}(2r+\nu)(\frac{z}{2})^{2r+\nu}$$

$$zJ'_\nu(z) = \nu J_\nu(z) + \sum_{r=0}^{\infty} \frac{(-1)^r}{r!\Gamma(\nu+r+1)} 2r(\frac{z}{2})^{2r+\nu}$$

$$zJ'_\nu(z) = \nu J_\nu(z) - z\sum_{r=0}^{\infty} \frac{(-1)^{r-1}}{(r-1)!\Gamma(r-1+\nu+1+1)}(\frac{z}{2})^{2(r-1)+\nu+1}$$

D'où la formule :

$$zJ'_\nu(z) = \nu J_\nu(z) - zJ_{\nu+1}(z)$$

On a de même :

$$zJ'_\nu(z) = -\nu J_\nu(z) + zJ_{\nu-1}(z)$$

Ce qui nous donne, en additionnant les deux formules :

$$2J'_\nu(z) = J_{\nu-1}(z) - J_{\nu+1}(z)$$

En particulier, $J'_0(z) = -J_1(z)$.

V)Conclusion

Mon stage aura été un travail de découverte et d'approfondissement d'un domaine physique qui me plaisait, et qui m'a permis de découvrir un peu ce à quoi ressemblait le métier de chercheur. J'ai énormément apprécié ce stage au LEM, d'une part grâce à la sympathie de mon entourage pendant mon stage, et d'autre part pour les nouvelles connaissances que j'ai pu acquérir, à la fois dans le domaine physique et mathématique.

Bibliographie

[**1**]L.Quettier et European Physical Journal, Ap.Phys.,32,167-175 (2005)

[**2**]C.Lorin, thèse de doctorat, INPL (2008)

[**3**]C.Lorin, A.Mailfert, publication en cours

[**4**]A.Angot,Compléments de mathématiques à l'usage des ingénieurs de l'électrotechnique et des télécommunications, 6ème édition, ISBN 2-225-34254-7

[**5**]G.B.Arfken et H.J.Weber, Mathematical methods for physicists, 5th edition, ISBN 0-12-059825-6

Annexe B

TER de M1

Equation $\overline{\partial}$ sur les ouverts de C

Romain Mismer

13 octobre 2015

Introduction

Ce rapport a pour but d'étudier l'équation $\frac{\partial u}{\partial \bar{z}} = f$, en traiter l'existence de solutions et de solutions vérifiant certaines propriétés. Pour ce faire, on démontrera des résultats classiques de la théorie des fonctions holomorphes comme la formule intégrale de Cauchy ou le principe du maximum, puis des théorèmes essentiels comme le théorème d'approximation de Runge, mais aussi des résultats qui peuvent paraître incroyables, comme celui de Weierstrass, qui énonce qu'on peut fabriquer une fonction méromorphe dont les pôles et les zéros ont été entièrement prescrits à l'avance. On commencera donc par rappeler dans un chapitre préliminaire des résultats de base sur les fonctions holomorphes, mais qui vont s'avérer très utiles pour démontrer des résultats essentiels comme le théorème de Runge, théorème qui servira à démontrer le théorème de Mittag-Leffler et l'existence de solutions de notre équation de départ. Ces résultats nous permettrons alors de résoudre le problème de Cousin (additif) et de démontrer le théorème de Weierstrass. Nous verrons ensuite quelques propriétés des fonctions sous-harmoniques, puis reviendrons à notre équation initiale dont on prouvera l'existence de solutions vérifiant certaines propriétés grâce aux fonctions sous-harmoniques.

Plan

Chapitre 1

Préliminaires : formule de Stokes et formule intégrale de Cauchy

1.1 Première définition : fonction holomorphe

Dans tout ce document, on notera Ω un ouvert du plan complexe C, qu'on identifiera à R^2. on note x et y les coordonnées réelles, et $z = x+iy$, $\overline{z} = x-iy$. Pour u une fonction $C^1(\Omega)$ à valeurs complexes, on note $\frac{\partial u}{\partial z} = \frac{1}{2}\left(\frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}\right)$ et $\frac{\partial u}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial u}{\partial x} + i\frac{\partial u}{\partial y}\right)$. Ainsi, la différentielle de u s'écrit comme une combinaison linéaire de dz et de $d\overline{z}$:

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy = \frac{\partial u}{\partial z}dz + \frac{\partial u}{\partial \overline{z}}d\overline{z}$$

Définition 1 (Fonction holomorphe)**.** *Soit $u \in C^1(\Omega)$. u est dite holomorphe (ou analytique) sur Ω si $\frac{\partial u}{\partial \overline{z}} = 0$ dans Ω (équation de Cauchy-Riemann). On note alors $u \in H(\Omega)$. Ceci est équivalent à du proportionnelle à dz. On notera aussi u' au lieu de $\frac{\partial u}{\partial z}$, de telle sorte que u holomorphe vérifie $du = u'dz$.*

Exemple (1)$\forall n \in N$, $d(z^n) = nz^{n-1}$, et donc tout polynôme $P(z) = \sum_{k=0}^{n} a_k z^k$ est holomorphe et $P'(z) = \sum_{k=1}^{n} k a_k z^{k-1}$.
(2)En posant $e^z = e^x(\cos y + i \sin y)$, l'exponentielle complexe est holomorphe, et $d(e^z) = e^z dz$.

Comme $\overline{\partial} = \frac{\partial}{\partial \overline{z}}$ est un opérateur linéaire, il est clair que toute combinaison linéaire complexe de fonctions holomorphes est holomorphe. Il en est de même pour un produit de fonctions holomorphes. Si v est une fonction holomorphe sur un ouvert contenant l'image de u, alors $v \circ u$ est holomorphe sur Ω. Au voisinage d'un point z_0 où la dérivée de u est non nulle, u est une transformation conforme. Sa différentielle en z_0 est $h \in C \to u'(z_0).h$, c'est une similitude directe du plan pour $u'(z_0) \neq 0$. Par le théorème des fonctions implicites, u^{-1} est C^1 sur un voisinage de $u'(z_0)$ et $\frac{\partial u^{-1}}{\partial z} = \frac{1}{u'(u^{-1}(z))}$.

1.2 Formule de Stokes

Définition 2 (1-forme différentielle)**.** *Soit U un ouvert de R^n, une 1-forme différentielle est une application f de U dans $(R^n)^*$.*

$$\forall a \in R^n, \forall x \in U, f(x)(a) =< f(x), a >= \sum_{i=1}^{n} f_i(x)a_i$$

$(dx_i)_{i\in[|1;n|]}$ base de $(R^n)^*$ ($\forall a \in R^n$, $dx_i(a) = a_i$) : $\forall x \in U$, $f(x) = \sum_{i=1}^{n} f_i(x)dx_i$.

Théorème 1 (Formule de Stokes)**.** *Soit K un compact à bord régulier. Si w est une 1-forme différentielle de classe C^1 au voisinage de K, alors*

$$\int_{\partial K} w = \int_K dw$$

Preuve On va simplement ici montrer la formule de Stokes pour $K = R = [a;b] \times [c;d]$ un rectangle du plan. Si $w = Pdx + Qdy$, alors $dw = (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dx \wedge dy$. La formule à prouver s'écrit donc

$$\int_{\partial R} Pdx + Qdy = \int_R (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy$$

On a

$$\int_{\partial R} Pdx + Qdy = \int_a^b P(x,c)dx + \int_c^d Q(b,y)dy - \int_a^b P(x,d)dx - \int_c^d Q(a,y)dy$$

Donc

$$\int_{\partial R} Pdx + Qdy = \int_a^b (P(x,c) - P(x,d))dx + \int_c^d (Q(b,y) - Q(a,y))dy$$

D'où

$$\int_{\partial R} Pdx + Qdy = \int_a^b (-\int_c^d \frac{\partial P(x,y)}{\partial y}dy)dx + \int_c^d (\int_a^b \frac{\partial Q(x,y)}{\partial x}dx)dy$$

Et finalement on a bien :

$$\int_{\partial R} Pdx + Qdy = \int_R (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy$$

1.3 Formule intégrale de Cauchy

La formule de Stokes s'écrit, si w est un ouvert borné de C et si u est une fonction $C^1(\overline{w})$:

$$\int_{\partial w} u dz = \int\int_w du \wedge dz$$

On a $du \wedge dz = (\frac{\partial u}{\partial z}dz + \frac{\partial u}{\partial \overline{z}}d\overline{z}) \wedge dz = \frac{\partial u}{\partial \overline{z}}d\overline{z} \wedge dz$ et $d\overline{z} \wedge dz = (dx - idy) \wedge (dx + idy) = idx \wedge dy - idy \wedge dx = 2idx \wedge dy$, d'où $du \wedge dz = 2i\frac{\partial u}{\partial \overline{z}}dx \wedge dy$. D'où la formule :

$$\int_{\partial w} u dz = 2i \int\int_w \frac{\partial u}{\partial \overline{z}} dx \wedge dy = \int\int_w \frac{\partial u}{\partial \overline{z}} d\overline{z} \wedge dz (*)$$

On a immédiatement en conséquence :

Proposition 1. *Soit w un ouvert borné de C et $u \in C^1(\overline{w})$ holomorphe, alors*

$$\int_{\partial w} u dz = 0$$

On obtient de plus la formule intégrale de Cauchy :

Théorème 2 (Formule intégrale de Cauchy). *Soit w un ouvert borné de C et $u \in C^1(\overline{w})$. $\forall \zeta \in w$,*

$$u(\zeta) = \frac{1}{2i\pi}\left(\int_{\partial w} \frac{u(z)}{z - \zeta} dz + \int\int_w \frac{\partial u}{\partial \overline{z}} \frac{1}{z - \zeta} dz \wedge d\overline{z}\right)$$

Preuve Soit $\zeta \in w$. Soit $\epsilon \in]0; d(\zeta, w^c)[$, on pose $w_\epsilon = \{z \in w; |z - \zeta| > \epsilon\}$ et $\phi : z \rightarrow \frac{u(z)}{z-\zeta}$.
On a $\phi \in C^1(\overline{w_\epsilon})$ et $\frac{\partial \phi(z)}{\partial \overline{z}} = \frac{1}{z-\zeta}\frac{\partial u(z)}{\partial \overline{z}}$. Donc $(*)$ appliquée à ϕ donne :

$$\int\int_{w_\epsilon} d\phi \wedge dz = \int_{\partial w_\epsilon \phi dz}$$

Soit

$$\int\int_{w_\epsilon} \frac{1}{z - \zeta} \frac{\partial u(z)}{\partial \overline{z}} d\overline{z} \wedge dz = \int_{\partial w_\epsilon} \frac{u(z)}{z - \zeta} dz$$

Donc $\int\int_{w_\epsilon} \frac{1}{z-\zeta}\frac{\partial u(z)}{\partial \overline{z}} d\overline{z} \wedge dz = \int_{\partial w} \frac{u(z)}{z-\zeta}dz - \int_{S(\zeta,\epsilon)} \frac{u(z)}{z-\zeta}dz$, où $S(\zeta, \epsilon)$ désigne le cercle de centre ζ et de rayon ϵ.
Or $\int_{S(\zeta,\epsilon)} \frac{u(z)}{z-\zeta}dz = \int_0^{2\pi} \frac{u(\gamma(\theta))}{\gamma(\theta)-\zeta}\gamma'(\theta)d\theta$, avec γ :

$$\begin{cases} [0; 2\pi] \rightarrow C \\ \theta \rightarrow \zeta + \epsilon e^{i\theta} \end{cases}$$

On a $\gamma'(\theta) = i\epsilon e^{i\theta}$ et donc :

$$\int_{S(\zeta,\epsilon)} \frac{u(z)}{z - \zeta} dz = \int_0^{2\pi} \frac{u(\zeta + \epsilon e^{i\theta}) i\epsilon e^{i\theta}}{\zeta + \epsilon e^{i\theta} - \zeta} d\theta = i \int_0^{2\pi} u(\zeta + \epsilon e^{i\theta}) d\theta$$

donc en faisant tendre ϵ vers 0, on a bien :

$$-\int\int_w \frac{1}{z-\zeta}\frac{\partial u(z)}{\partial \overline{z}}dz\wedge d\overline{z} = \int_{\partial w}\frac{u(z)}{z-\zeta}dz - i\int_0^{2\pi}u(\zeta)d\theta = \int_{\partial w}\frac{u(z)}{z-\zeta}dz - 2i\pi u(\zeta)$$

D'où le résultat.

1.4 Applications de la formule de Cauchy

Théorème 3. *Soit μ une mesure à support K compact dans C. $u :\begin{cases} C \to C \\ \zeta \to \int \frac{1}{z-\zeta}d\mu(z)\end{cases}$ définit une fonction holomorphe C^∞ en dehors de K. Sur tout ouvert Ω où $d\mu = \frac{1}{2i\pi}\phi dz\wedge d\overline{z}$ avec $\phi \in C^k(\Omega)$, $u \in C^k(\Omega)$ et $\frac{\partial u}{\partial \overline{z}} = \phi$ si $k \geq 1$.*

Preuve $*$ On pose $\psi : \begin{cases} K\times K^c \to C \\ (z,\zeta) \to \frac{1}{z-\zeta}\end{cases}$. On a $\psi \in C^\infty$ et $\frac{\partial \psi}{\partial \overline{\zeta}} = 0$, donc par dérivation sous l'intégrale, $u \in C^\infty$ et est holomorphe en dehors de K.
$*$ Supposons d'abord $\Omega = R^2$

$$u(\zeta) = \frac{1}{2i\pi}\int\int \frac{1}{z-\zeta}\phi(z)dz\wedge d\overline{z}$$

$$u(\zeta) = \frac{-1}{2i\pi}\int\int \frac{1}{z}\phi(\zeta - z)dz\wedge d\overline{z}$$

Comme $z \to \frac{1}{z}$ est intégrable sur tout compact, par dérivation sous l'intégrale k fois, on a $u \in C^k$ et

$$\frac{\partial u(\zeta)}{\partial\overline{\zeta}} = \frac{-1}{2i\pi}\int\int\frac{1}{z}\frac{\partial\phi(\zeta - z)}{\partial\overline{\zeta}}dz\wedge d\overline{z}$$

$$\frac{\partial u(\zeta)}{\partial\overline{\zeta}} = \frac{1}{2i\pi}\int\int\frac{1}{z-\zeta}\frac{\partial\phi(z)}{\partial\overline{\zeta}}dz\wedge d\overline{z}$$

et la formule de Cauchy appliquée à ϕ sur un disque contenant le support de ϕ donne $\frac{\partial u}{\partial\overline{\zeta}} = \phi$
$*$ Si maintenant Ω est un ouvert quelconque, pour $z_0 \in \Omega$, $\exists \psi \in C_0^k(\Omega)$ qui vaut 1 dans un voisinage V de z_0.
On pose $\mu_1 = \psi\mu$ et $\mu_2 = (1-\psi)\mu$, on a alors $u = u_1 + u_2$, avec $u_i(\zeta) = \int\frac{1}{z-\zeta}d\mu_i(z)$
On a $\mu_1 = \frac{1}{2i\pi}\psi\phi dz\wedge d\overline{z}$ avec $\psi\phi \in C_0^k(R^2)$ donc $u_1 \in C^k$ et $\frac{\partial u_1}{\partial\overline{\zeta}} = \psi\phi$. De plus μ_2 est nulle sur V et $u \in C^k(V)$, et $\frac{\partial u}{\partial\overline{\zeta}} = \phi$ au voisinage de z_0.

Corollaire 1. $H(\Omega) = \{$*fonctions holomorphes sur* $\Omega\} \subset C^\infty(\Omega)$
Si $u \in H(\Omega)$ *alors* $u' \in H(\Omega)$

Preuve Soit w un disque d'adhérence incluse dans Ω, et soit $\zeta \in w$.

$$u(\zeta) = \frac{1}{2i\pi} \int_{\partial w} \frac{u(z)}{z-\zeta} dz$$

Donc $u(\zeta) = \int_{\partial w} \frac{1}{z-\zeta} d\mu(z)$ avec $d\mu(z) = \frac{1}{2i\pi} u(z) dz$

$$u'(\zeta) = \int_{\partial w} \frac{1}{(z-\zeta)^2} d\mu(z)$$

donc $\frac{\partial u'}{\partial \bar{\zeta}}(\zeta) = 0$

Théorème 4. *$\forall K \subset\subset \Omega$ et $\forall$ voisinage ouvert $w \subset \Omega$ de K, $\exists C_j$, $j = 0, 1, ..$ tels que $\sup_{z \in K} |u^{(j)}(z)| \leq C_j \|u\|_{L^1(w)} \forall u \in H(\Omega)$*

Preuve On prend $\psi \in C_0^\infty(w)$ telle que $\psi = 1$ au voisinage de K. Soit $u \in H(\Omega)$, $\frac{\partial(\psi u)}{\partial \bar{z}} = u \frac{\partial \psi}{\partial \bar{z}}$.
La formule de Cauchy donne $\psi(\zeta) u(\zeta) = \frac{1}{2i\pi} \int_w \frac{\partial \psi}{\partial \bar{z}} \frac{u(z)}{z-\zeta} dz \wedge d\bar{z}$, donc, comme ψ vaut 1 au voisinage de K, $\forall \zeta \in K$:

$$u^{(j)}(\zeta) = \frac{1}{2i\pi} \int_w \frac{\partial \psi}{\partial \bar{z}} \frac{u(z) j!}{(z-\zeta)^{j+1}} dz \wedge d\bar{z}$$

d'où le résultat.

Corollaire 2. *Si $u_n \in H(\Omega)$ et u_n converge vers u quand n tend vers l'infini uniformément sur tout compact de Ω, alors $u \in H(\Omega)$*

Preuve En appliquant le théorème 4 à $u_n - u_m$, on obtient que $\left(\frac{\partial u_n}{\partial z}\right)_n$ converge uniformément. $\forall n$, $\frac{\partial u_n}{\partial \bar{z}} = 0$ donc $\left(\frac{\partial u_n}{\partial x}\right)_n$ et $\left(\frac{\partial u_n}{\partial y}\right)_n$ convergent uniformément sur tout compact. Donc $u \in C^1$ et $\frac{\partial u}{\partial \bar{z}} = \lim \frac{\partial u_n}{\partial \bar{z}} = 0$

Corollaire 3 (Stieljes-Vitali)**.** *Si $u_n \in H(\Omega)$ et $|u_n|$ est bornée uniformément sur tout compact de Ω, $\exists (u_{n_j})_j$ qui converge uniformément sur tout compact de Ω vers une limite $u \in H(\Omega)$*

Preuve Soit $z \in \Omega$ et $\epsilon > 0$. $\forall z' \in \Omega$, $|u_n(z) - u_n(z')| \leq |z - z'| \sup_K |u_n'| \leq |z - z'| C_K$ pour tout $K \subset\subset \Omega$. Donc si $|z - z'| < \delta = \frac{\epsilon}{C_K}$, $|u_n(z) - u_n(z')| < \epsilon$, donc cette suite est équicontinue. On a de plus $\{u_n(z), n \geq 0\}$ relativement compacte pour tout z, donc par le théorème d'Arzéla-Ascoli, il existe bien une sous-suite qui converge uniformément sur tout compact de Ω, vers une limite u qui est holomorphe par le corollaire précédent.

Corollaire 4. *La série $\sum_{n=0} a_n z^n$ converge uniformément sur tout disque inclus dans son disque de convergence, $u(z) = \sum_{n=0}^\infty a_n z^n$ est donc holomorphe sur son disque de convergence.*

Théorème 5. *Soit $u \in H(\Omega)$, $\Omega = D(0,r)$. On a $u(z) = \sum_{n=0}^{\infty} \frac{u^{(n)}(0)}{n!} z^n$, avec convergence uniforme sur tout compact de Ω*

Preuve Si $r_1 < r_2 < r$, on a, $\forall |\zeta| \leq r_1$, $u(z) = \frac{1}{2i\pi} \int_{|\zeta|=r_2} \frac{u(\zeta)}{\zeta - z} d\zeta$, donc

$$u^{(n)}(z) = \frac{1}{2i\pi} \int_{|\zeta|=r_2} \frac{u(\zeta)n!}{(\zeta - z)^{n+1}} d\zeta$$

Pour $|\zeta| \leq r_1$ et $|\zeta| = r_2$, on a $\frac{1}{\zeta - z} = \sum_{n=0}^{\infty} \frac{1}{\zeta} \left(\frac{z}{\zeta}\right)^n$. Comme la série est absolument et uniformément convergente, on obtient le théorème en intégrant terme à terme :

$$u(z) = \sum_{n=0}^{\infty} z^n \frac{1}{2i\pi} \int_{|\zeta|=r_2} \frac{u(\zeta)}{\zeta^{n+1}} d\zeta$$

d'où $u(z) = \sum_{n=0}^{\infty} \frac{u^{(n)}(0)}{n!} z^n$.

Corollaire 5 (Unicité du prolongement analytique)**.** *Si $u \in H(\Omega)$ et $\exists z \in \Omega$ tel que, $\forall k \geq 0$, $u^{(k)}(z) = 0$, alors $u = 0$ sur Ω si Ω est connexe. Donc si Ω est connexe, et que f et g sont 2 fonctions holomorphes sur Ω qui coïncident sur un ouvert non vide de Ω, alors elles sont égales sur Ω tout entier.*

Preuve On pose $A = \{z \in \Omega; \forall k \geq 0, u^{(k)}(z) = 0\}$. A est un fermé de Ω car tous les $u^{(k)}$ sont continus, mais par le théorème qui précède, si $z_0 \in A$ et si r est tel que $D = D(z_0, r)$ soit contenu dans Ω, alors u est identiquement nul sur D puisque $u(z) = \sum_{n=0}^{\infty} u^{(n)}(z_0) \frac{(z-z_0)^n}{n!}$ sur ce disque. Ses dérivées successives le sont aussi et donc D est inclus dans A. Ceci prouve que A est aussi ouvert donc comme Ω est connexe, A est soit vide soit Ω tout entier, or il n'est pas vide par hypothèse, d'où le résultat.

Corollaire 6. *Soit $u \in H(D(0,r))$. Si u n'est pas identiquement nulle, il existe un unique $v \in H(D(0,r))$ et un unique $n \in N$ tels que $v(0) \neq 0$ et $u(z) = z^n v(z)$*

En effet, on pose $n = \min\{k; u^{(k)}(0) \neq 0\}$, on a $u(z) = \sum_{k=n}^{\infty} u^{(k)}(0) \frac{z^k}{k!} = z^n \sum_{k=n}^{\infty} u^{(k)}(0) \frac{z^{k-n}}{k!} = z^n v(z)$

Théorème 6. *Soit $u \in H(D(z_0, r))$ tel que $|u(z)| \leq |u(z_0)|$ pour tout $z \in D(z_0, r)$, alors u est constante sur $D(z_0, r)$*

Preuve Supposons $u(z_0) \neq 0$. Par la formule de Cauchy, on obtient :

$$u(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + \rho e^{i\theta}) d\theta, 0 < \rho < r$$

Donc $\int_0^{2\pi} (1 - \frac{u(z_0 + \rho e^{i\theta})}{u(z_0)}) d\theta = 0$

On a $Re(1 - \frac{u(z_0+\rho e^{i\theta})}{u(z_0)}) \geq 0$ et $= 0$ si et seulement si $u(z_0) = u(z_0 + \rho e^{i\theta})$.

D'où $\forall 0 < \rho < r$ et $\theta \in [0; 2\pi]$,$u(z_0) = u(z_0 + \rho e^{i\theta})$.

Corollaire 7 (Principe du maximum)**.** *Soit Ω un ouvert borné et $u \in C(\overline{\Omega})$ holomorphe sur Ω. Alors le maximum de $|u|$ sur $\overline{\Omega}$ est atteint en un point du bord.*

Preuve S'il est atteint en un point intérieur z_0, u est constant sur tout disque de centre z_0 contenu dans Ω, donc est constant sur la composante connexe de Ω contenant z_0 par unicité du prolongement analytique. Donc le maximum de $|u|$ est atteint en un point du bord.

Chapitre 2

Le théorème d'approximation de Runge et ses applications

2.1 Enoncé

On a vu qu'une fonction holomorphe sur un disque peut être approchée uniformément par des fonctions polynomiales sur tout disque plus petit, par exemple en prenant les sommes partielles de son développement en série entière. Le théorème de Runge fournit un résultat d'approximation plus général :

Théorème 7 (Runge)**.** *Soit Ω un ouvert de C et K un compact inclus dans Ω. Les assertions suivantes sont équivalentes :*
(a)Toute fonction holomorphe sur un voisinage $V(K)$ de K peut être approchée uniformément sur K par des fonctions de $H(\Omega)$
(b)$\Omega \setminus K$ n'a pas de composante relativement compacte dans Ω
(c)$\forall z \in \Omega \setminus K$, $\exists f \in H(\Omega)$ telle que $|f(z)| > \sup_K |f|$

Corollaire 8. *Toute fonction holomorphe sur un voisinage de K compact peut être approchée uniformément par des polynômes sur K si et seulement si K^c est connexe (si et seulement si pour tout $z \in K^c$,$\exists f$ un polynôme tel que $|f(z)| > \sup_K |f|$)*

Preuve *(c) implique (b)
Supposons que $\Omega \setminus K$ a une composante O telle que $\overline{O}$ est un compact de Ω. Le principe du maximum nous dit que $\forall f \in H(\Omega)$ le maximum de $|f|$ sur $\overline{O}$ est atteint sur $\partial O \subset K$, donc $\forall f \in H(\Omega)$, $\sup_{\overline{O}} |f| \leq \sup_K |f|$. Or d'après (c), $\forall z \in \Omega \setminus K \exists f \in H(\Omega)$ telle que $|f(z)| > \sup_K |f|$, or pour $z \in O$, $|f(z)| \leq \sup_{\overline{O}} |f|$, et on obtient finalement $\sup_K |f| > \sup_K |f|$, ce qui est absurde.

(a) implique (b)

On suppose toujours que (b) n'est pas vérifié, (a) nous donne : $\forall f \in H(V(K))$, $\exists f_n \in H(\Omega)$ qui converge uniformément vers f sur K.

$$\sup_{\overline{O}} |f_n - f_m| \leq \sup_K |f_n - f_m|$$

Donc $(f_n)_n$ converge uniformément sur $\overline{O}$ vers une fonction F. Sur ∂O on a $F = f$, et $F \in H(O)$ et est continue sur $\overline{O}$.
On pose $f(z) = \frac{1}{z-\zeta}$, $\zeta \in O. \forall z \in \partial O$, $(z-\zeta)F(z) = 1$, donc $\forall z \in O$,$(z-\zeta)F(z) = 1$, ce qui est absurde en prenant $z = \zeta$.

(b) implique (a)
On veut prouver que $G \subset F$, F dénotant l'adhérence des restrictions à K des fonctions holomorphes au voisinage de K et G les restrictions à K des fonctions holomorphes sur Ω. Cela revient à prouver que toute forme linéaire continue de $C(K) = \{\text{formes linéaires continues de } K\}$ s'annulant sur G s'annule sur F. Or par le théorème de représentation de Riesz, pour toute forme linéaire L de $C(K)$, $\exists! \mu$ mesure borélienne complexe telle que pour toute $f \in C(K)$, $L(f) = \int f d\mu$. Nous allons donc prouver que toute mesure μ sur K qui est orthogonale à $H(\Omega)$ est aussi orthogonale à toute fonction holomorphe sur $V(K)$.
Posons, pour $\zeta \in K^c$, $\phi(\zeta) = \int \frac{1}{z-\zeta} d\mu(z)$. ϕ est holomorphe sur K^c et $\forall \zeta \in K^c \cap \Omega^c$, $\phi^{(k)}(\zeta) = k! \int \frac{1}{(z-\zeta)^{k+1}} d\mu(z) = 0$ car $z \to \frac{1}{(z-\zeta)^{k+1}}$ est holomorphe sur Ω si $\zeta \in \Omega^c$. Donc $\phi = 0$ sur toute composante connexe de K^c qui rencontre Ω^c. De plus $\int z^n d\mu(z) = 0$ pour tout n et $z \to \frac{1}{z-\zeta}$ est développable en une série entière qui converge uniformément sur K dès que $|\zeta| > \sup_K |z|$, donc $\phi = 0$ sur la composante non bornée de K. Finalement, par (b), $\Omega \setminus K$ n'a pas de composante relativement compacte dans Ω, d'où $\phi = 0$ sur K^c.
Soient $f \in H(V(K))$ et $\psi \in C_0^\infty(V(K))$ telle que $\psi = 1$ sur $K. \forall z \in K$,

$$f(z) = \psi(z)f(z) = \frac{1}{2i\pi} \int\int \frac{f(\zeta)}{\zeta - z} \frac{\partial \psi}{\partial \overline{\zeta}}(\zeta) d\zeta \wedge d\overline{\zeta}$$

Donc

$$f(z) = \frac{-1}{2i\pi} \int\int f(\zeta)\phi(\zeta) \frac{\partial \psi}{\partial \overline{\zeta}}(\zeta) d\zeta \wedge d\overline{\zeta}$$

Donc, puisque $\phi = 0$ sur K^c et que $\frac{\partial \psi}{\partial \overline{\zeta}} = 0$ sur K, $\int_K f(z) d\mu(z) = 0$, d'où le résultat.
*(b) implique (c)
Soit $z \in \Omega \setminus K$.
On pose $L = \overline{D}(z, \delta)$, avec δ choisi tel que $L \subset \Omega \setminus K$. $\Omega \setminus (K \cup L)$ n'a pas de composante relativement compacte dans Ω donc on sait que par (a), on peut approcher la fonction qui vaut 0 sur $V(K)$ et 1 sur $V(L)$: $\exists f \in H(\Omega)$ telle que $|f| < \frac{1}{2}$ sur K et $|f - 1| < \frac{1}{2}$ sur L. Donc $\frac{-1}{2} < f(z) - 1 < \frac{1}{2}$, donc $\frac{1}{2} < f(z) < \frac{3}{2}$, d'où $|f(z)| > \frac{1}{2} > \sup_K |f|$

2.2 Enveloppe d'holomorphie

Une première application de ce théorème est l'enveloppe d'holomorphie d'un compact :

Définition 3 (Enveloppe d'holomorphie). *Soit $K \subset\subset \Omega$. L'enveloppe d'holomorphie de K est définie par $\widehat{K} = \widehat{K}_\Omega = \{z; z \in \Omega; \forall f \in H(\Omega), |f(z)| \leq \sup_K |f|\}$.*

On peut d'abord remarquer que $K \subset \widehat{K}$ et $\widehat{\widehat{K}} = \widehat{K}$. On a aussi les propriétés suivantes :

Proposition 2 (Quelques propriété de l'enveloppe d'holomorphie). *(i) $d(K, \Omega^c) = d(\widehat{K}, \Omega^c)$*
(ii)$\widehat{K}$ est inclus dans l'enveloppe convexe $conv(K)$ de K.

Preuve (i)Soit $\zeta \in \Omega^c$, $z \to \frac{1}{z-\zeta}$ est holomorphe sur Ω. Soit $z_0 \in \widehat{K}$, $|\frac{1}{z_0-\zeta}| \leq \sup_K |\frac{1}{z-\zeta}| = \frac{1}{d(K,\Omega^c)}$, donc $d(K, \Omega^c) \leq d(\widehat{K}, \Omega^c)$. On a bien sûr $d(K, \Omega^c) \geq d(\widehat{K}, \Omega^c)$, d'où l'égalité.

(ii) On a $conv(K) = \{z \in C; \forall a \in C, < z, a >\leq \sup_{\zeta \in K} < \zeta, a >\}$. Soit $z_0 \in \widehat{K}$, $a \in C$, $z \to e^{az}$ est holomorphe sur $\Omega.|e^{az_0}| \leq \sup_K |e^{az}| = \sup_K e^{Re(<a,z>)}$, d'où $Re(< a, z_0 >) \leq \sup_K Re(< a, z >)$

L'enveloppe d'holomorphie de K apparaît donc comme le compact contenant K remplissant les conditions du théorème d'approximation de Runge. On peut la caractériser d'avantage grâce à la condition (b) du théorème :

Théorème 8. *$\widehat{K}$ est la réunion de K et des composantes connexes de $\Omega \setminus K$ qui sont relativement compactes dans Ω.*

Preuve On note U la réunion de K et des composantes connexes de $\Omega \setminus K$ qui sont relativement compactes dans Ω. Soit O une telle composante, on a prouvé que $\sup_{\overline{O}} |f| \leq \sup_K |f|$, donc $O \subset \widehat{K}$. Ceci prouve que $U \subset \widehat{K}$.
$\Omega \setminus U$ est ouvert en tant qu'union de composantes ouvertes de $\Omega \setminus K$. Donc U est compact et vérifie de plus la condition (b) du théorème de Runge : $\Omega \setminus U$ n'a pas de composante relativement compacte dans Ω. Donc la condition (c) du théorème de Runge nous indique que $U = \widehat{U}$. Donc $\widehat{K} \subset U = \widehat{U}$, d'où $\widehat{K} = U$

Exemple On prend $\Omega = \{z \in C; \frac{1}{2} < |z| < 2\}$, $K_1 = \{z \in \Omega; |z - 1| = \frac{1}{4}\}$ et $K_2 = \{z \in \Omega; |z| = 1\}$. Alors $\widehat{K_1} = \overline{D}(1, \frac{1}{4})$ et $\widehat{K_2} = K_2$.

On a désormais une variante du théorème de Runge pour 2 ouverts :

Théorème 9. *Soient $\Omega_1 \subset \Omega_2$ deux ouverts de C. Les assertions suivantes sont équivalentes :*
(a)Toute fonction holomorphe sur Ω_1 peut être approchée par des fonctions holomorphes sur Ω_2 uniformément sur chaque compact de Ω_1.
(b)Si $\Omega_2 \setminus \Omega_1$ est l'union disjointe d'un fermé F de Ω_2 et d'un compact K, alors K est vide.
(c_1)$\forall K \subset\subset \Omega_1$, $\widehat{K}_{\Omega_2} = \widehat{K}_{\Omega_1}$.
(c_2)$\forall K \subset\subset \Omega_1$, $\widehat{K}_{\Omega_2} \cap \Omega_1 = \widehat{K}_{\Omega_1}$.
(c_3)$\forall K \subset\subset \Omega_1$, $\widehat{K}_{\Omega_2} \cap \Omega_1$ est compact.

Preuve On a clairement (a) implique (c_2) et (c_2) implique (c_3).
On pose $K' = \widehat{K}_{\Omega_2} \cap \Omega_1$ et $K'' = \widehat{K}_{\Omega_2} \cap \Omega_1^c$, si ($c_3$) est vérifié, K' et K'' sont des compacts. Soit $f \in H(\Omega_1)$, par le théorème de Runge, la fonction qui vaut f sur K' et 1 sur K'' peut être approchée uniformément sur $K' \cup K''$ par des fonctions holomorphes sur Ω_2. Donc (c_3) implique (a). Si on choisit de plus $f = 0$, on obtient que K'' est vide, donc $\widehat{K}_{\Omega_2} = \widehat{K}_{\Omega_2} \cap \Omega_1 = \widehat{K}_{\Omega_1}$, on obtient donc l'équivalence de (a), (c_1), (c_2) et (c_3).

(c_1) implique (b)
Si $\Omega_2 \setminus \Omega_1 = F \sqcup K$, soit un ouvert $\omega \supset K$ tel que $\overline{\omega} \cap F$ est vide et tel que ω est relativement compact dans Ω_2. Comme $\partial\omega \cap \Omega_2 \setminus \Omega_1$ est vide, et comme $\partial\omega \subset \Omega_2$, on a $\partial\omega \subset \Omega_1$. Donc par principe du maximum l'enveloppe d'holomorphie relativement à Ω_2 de $\partial\omega$ contient ω et donc contient K. Donc $K \subset \Omega_1$ par (c_1), d'où $K = \emptyset$.

*(b) implique (c_1)
Soit $K \subset\subset \Omega_1$, soit O une composante de $\Omega_2 \setminus K$ relativement compacte dans Ω_2. On a $\partial O \subset K \subset \Omega_1$ donc l'ensemble $K' = \overline{O} \cap (\Omega_2 \setminus \Omega_1) = O \cap (\Omega_2 \setminus \Omega_1)$ est un ensemble compact de O. Et comme $O^c \cap (\Omega_2 \setminus \Omega_1)$ est un fermé de Ω_2, on a $K' = \emptyset$ par (b). D'où $O \subset \Omega_1$, donc par le théorème qui précède, on a $\widehat{K}_{\Omega_2} \subset \widehat{K}_{\Omega_1}$, l'autre inclusion est évidente.

Chapitre 3

Le théorème de Mittag-Leffler

3.1 Fonctions méromorphes

Pour tout $z \in C$, on note A_z l'ensemble des classes d'équivalence des fonctions f qui sont holomorphes au voisinage de z pour la relation d'équivalence $f \sim g \Leftrightarrow f = g$ au voisinage de z. Si f est holomorphe au voisinage de z, on note f_z la classe de f. A_z est un anneau intègre donc on peut considérer son corps des fractions noté M_z.

Définition 4 (Fonction méromorphe). *Une fonction méromorphe ϕ sur Ω est une application de $\Omega \to \bigcup_{z\in\Omega} M_z$ telle que, $\forall z \in \Omega$, $\phi(z) \in M_z$ et $\exists$ un voisinage ω de z et des fonctions f et g holomorphes sur ω telles que, $\forall z \in \omega$, $\phi(z) = \frac{f_z}{g_z}$. L'ensemble des fonctions méromorphes sur Ω est noté $M(\Omega)$.*

En particulier, si $f \in H(\Omega)$, l'application $z \to f_z$ est une fonction méromorphe. Comme des fonctions holomorphes différentes définissent des fonctions méromorphes différentes, on identifie $H(\Omega)$ comme un sous-ensemble de $M(\Omega)$.
Les fonctions méromorphes forment un anneau dans lequel tout élément qui ne s'annule pas identiquement sur une composante connexe de Ω est inversible.
Pour tout $q \in M_\zeta$, on peut assigner à ζ une valeur $q(\zeta)$. Pour cela on prend f et g holomorphes au voisinage de ζ telles que $q = \frac{f_\zeta}{g_\zeta}$, $g_\zeta \neq 0$. Si $q = 0$, on pose $q(\zeta) = 0$. Sinon, on sait qu'il existe n,m des entiers et f_1 et g_1 holomorphes au voisinage de ζ tels que $f_1(\zeta)g_1(\zeta) \neq 0$ et $f(z) = (z-\zeta)^n f_1(z)$ et $g(z) = (z-\zeta)^m g_1(z)$. Il est clair que $n - m$ et $\frac{f_1(\zeta)}{g_1(\zeta)}$ ne dépendent que de q et non du choix de f et g. On définit ainsi

$$q(\zeta) = \begin{cases} \infty, n < m \\ \frac{f_1(\zeta)}{g_1(\zeta)}, n = m \\ 0, n > m \end{cases}$$

Si $\phi \in M(\Omega)$, on obtient une application $z \to \phi_z(z) = F(z) \in C \cup \{\infty\}$ de telle sorte que F est holomorphe sur le complémentaire d'un ensemble discret $D \subset \Omega$ et $\frac{1}{F}$ est holomorphe au voisinage de D (avec la convention $\frac{1}{\infty} = 0$). Réciproquement, si on a une fonction F avec ces propriétés, une fonction méromorphe ϕ est définie par $\phi_z = F_z$ si $z \notin D$ et $\phi_z = \frac{1}{(\frac{1}{F})_z}$ si $z \in D$ et $\phi_z(z) = F(z)$ pour tout z. La définition donnée précédemment est donc équivalente à la définition classique d'une fonction méromorphe. Les points z pour lesquels $F(z) = \infty$ sont les pôles de F.

Théorème 10. *Si F est méromorphe au voisinage de ζ, alors $\exists (A_1, .., A_n) \in C^n$, G une fonction holomorphe et ω un voisinage de ζ sur lequel*

$$F(z) = \sum_{k=1}^{n} \frac{A_k}{(z-\zeta)^k} + G(z)$$

Cette écriture est unique. Si $F_\zeta \neq 0$, F s'écrit aussi de manière unique sous la forme $F(z) = (z-\zeta)^n G(z)$, avec n un entier et $G(\zeta) \neq 0$. Si $n > 0$, ζ est un zéro d'ordre n ; si $n < 0$, ζ est un pôle d'ordre $-n$.

La preuve découle directement du corollaire 6 du chapitre 1.

3.2 Théorème de Mittag-Leffler

Théorème 11 (Mittag-Leffler)**.** *Soit $(z_j)_j$ une suite discrète de points de Ω, et $\forall j$, f_j méromorphe au voisinage de z_j. Alors il existe une fonction méromorphe f sur Ω telle que f est holomorphe en dehors des z_j et $f - f_j$ est holomorphe au voisinage de z_j pour tout j.*

Preuve Par le théorème qui précède, on peut écrire $f_j(z) = \sum_{k=1}^{n_j} \frac{A_{jk}}{(z-z_j)^k}$. On cherche des fonctions $u_j \in H(\Omega)$ telles que la série

$$f(z) = \sum_{j=1}^{\infty} (f_j(z) - u_j(z))$$

définisse une fonction avec les propriétés requises. Pour cela on choisit une suite croissante (pour l'inclusion) de compacts $K_j \subset \Omega$ avec $K_j = \widehat{K}_j$ de sorte que tout compact de Ω soit inclus dans un K_j. On peut supposer que $\forall k \geq j$, $z_k \notin K_j$ puisque la suite n'a pas de point d'accumulation dans Ω. Donc par le théorème de Runge, il existe $u_j \in H(\Omega)$ tel que, $\forall z \in K_j$,

$$|f_j(z) - u_j(z)| < 2^{-j}$$

Ainsi la série $\sum_{j=k}^{\infty} (f_j(z) - u_j(z))$ converge uniformément sur K_k pour tout k vers une fonction holomorphe à l'intérieur de K_k. La définition de f ci-dessus a donc un sens et f a les propriétés requises.

On a aussi la formulation suivante :

Théorème 12 (Mittag-Leffler bis). *Soit $\Omega = \bigcup_j \Omega_j$ avec Ω_j des ouverts de C. Si $f_j \in M(\Omega_j)$ et $f_j - f_k \in H(\Omega_j \cap \Omega_k)$ pour tous j et k, alors il existe $f \in M(\Omega)$ telle que $f - f_j \in H(\Omega_j)$ pour tout j.*

Preuve On a bien équivalence entre les deux formulations : Si ce théorème bis est vérifié, soit $(z_j)_j$ une suite discrète de points de Ω, et $\forall j$, f_j méromorphe au voisinage de z_j. On écrit alors $\Omega = \bigcup_j \Omega_j$, avec $z_j \in \Omega_j$ pour tout j, on a $f_j \in M(\Omega_j)$ et $f_j - f_k \in H(\Omega_j \cap \Omega_k)$ pour tous j et k, donc il existe $f \in M(\Omega)$ telle que $f - f_j \in H(\Omega_j)$ pour tout j, donc le théorème bis implique bien le théorème de Mittag-Leffler.

Si le théorème de Mittag-Leffler est vérifié, et si $\Omega = \bigcup_j \Omega_j$ avec Ω_j des ouverts de C, soient $f_j \in M(\Omega_j)$ telles que $f_j - f_k \in H(\Omega_j \cap \Omega_k)$ pour tous j et k. On utilise alors Mittag-Leffler pour construire une fonction f holomorphe en dehors des pôles des f_j et telle que $f - f_j$ est holomorphe au voisinage des pôles de f_j pour tout j. Donc pour tout j, on a bien $f - f_j \in H(\Omega_j)$

3.3 Equation $\frac{\partial u}{\partial \bar{z}} = f$ et problème de Cousin

On peut, à l'aide d'une preuve très analogue à celle donnée pour le théorème de Mittag-Leffler, prouver le résultat suivant :

Théorème 13. *Pour toute $f \in C^\infty(\Omega)$, l'équation $\frac{\partial u}{\partial \bar{z}} = f$ a une solution $u \in C^\infty(\Omega)$*

Preuve On commence par choisir une suite croissante de compacts $K_j \subset \Omega$ avec $K_j = \widehat{K}_j$ de sorte que tout compact de Ω soit inclus dans un K_j. Soit alors $\psi_j \in C_0^\infty(\Omega)$ égale à 1 au voisinage de K_j, on pose alors $\phi_1 = \psi_1$ et, pour tout $j > 1$, $\phi_j = \psi_j - \psi_{j-1}$.
On a $\phi_j = 0$ au voisinage de K_{j-1} et $\sum_{j=1}^\infty \phi_j = 1$ sur Ω. Par le théorème 3 du chapitre 1, $\exists u_j \in C^\infty(R^2)$ telle que $\frac{\partial u_j}{\partial \bar{z}} = \phi_j f$. En particulier on a u_j holomorphe au voisinage de K_{j-1}.
Par le théorème de Runge $\exists v_j \in H(\Omega)$ telle que $|u_j - v_j| < 2^{-j}$ sur K_{j-1}. La série

$$u(z) = \sum_{j=1}^{\infty} (u_j(z) - v_j(z))$$

est uniformément convergente sur tout compact de Ω. De plus la somme qui débute au terme $k+1$ est constituée de termes qui sont holomorphes au voisinage de K_k et converge uniformément sur K_k vers une fonction holomorphe à l'intérieur de K_k. Donc $u \in C^\infty(\Omega)$ et en dérivant terme à terme on a

$$\frac{\partial u}{\partial \bar{z}} = \sum_{j=1}^{\infty} \phi_j f = f$$

Le problème de Cousin additif est le suivant :
Soit $\Omega = \bigcup_{j \in J} \Omega_j$ et soit pour tout couple $(i, j) \in J^2$ une fonction $g_{ij} \in H(\Omega_i \cap \Omega_j)$. Est-il possible de trouver des fonctions $g_j \in H(\Omega_j)$ pour tout $j \in J$ telles que $g_{ij} = g_j - g_i$ sur $\Omega_i \cap \Omega_j$ pour tout $(i, j) \in J^2$?

Si ce problème admet des solutions, il est clair que les fonctions g_{ij} doivent vérifier $g_{ij} = -g_{ji}$ sur $\Omega_i \cap \Omega_j$ et $g_{ij} + g_{jk} + g_{ki} = 0$ sur $\Omega_i \cap \Omega_j \cap \Omega_k$.
Il s'avère que ces conditions suffisent à elles seules à assurer l'existences des g_j :

Théorème 14. *Si ces conditions sont vérifiées, il existe des solutions pour le problème de Cousin additif*

On introduit d'abord la notion de partition C_0^∞ de l'unité :

Définition 5 (Partition C_0^∞ de l'unité)**.** *On appelle partition C_0^∞ de l'unité d'un espace topologique X la donnée de fonctions $(\phi_j)_{j \in J}$ qui sont C_0^∞ à valeur dans $[0; 1]$ et telles que,$\forall x \in X$,*
(i) Il existe un voisinage de x tel que les ϕ_j sont nulles sur ce voisinage à l'exception d'un nombre fini d'entre elles,
(ii)$\sum_{j \in J} \phi_j(x) = 1$

Définition 6 (Partition C_0^∞ de l'unité subordonnée à un recouvrement)**.** *On appelle partition C_0^∞ de l'unité subordonnée au recouvrement $(\Omega_j)_{j \in J}$ une partition $(\phi_j)_{j \in J}$ au sens de la première définition, indexée par le même ensemble que le recouvrement, et telle que, pour tout $j \in J$, le support de ϕ_j soit inclus dans Ω_j*

On sait alors qu'on peut toujours trouver une telle partition subordonnée au recouvrement de tout ouvert $\Omega \subset R^n$.

Preuve On choisit une partition C_0^∞ de l'unité subordonnée au recouvrement $(\Omega_j)_{j \in J}$ notée $(\phi_j)_{j \in J}$. Si $g_{ij} = g_j - g_i$ sur $\Omega_i \cap \Omega_j$ pour tout $(i, j) \in J^2$, on a $\forall i \in J$, $\phi_i g_j = \phi_i g_{ij} + \phi_i g_i$, donc

$$g_j = h_j + u$$

avec $h_j = \sum_{i \in J} \phi_i g_{ij}$ et $u = \sum_{i \in J} \phi_i g_i$.$h_j$ est bien définie par les fonctions g_{ij} et est dans $C^\infty(\Omega_j)$. On a de plus

$$h_j - h_i = \sum_{k \in J} \phi_k (g_{kj} - g_{ki}) = \sum_{k \in J} \phi_k g_{ij} = g_{ij}$$

g_{ij} est holomorphe sur $\Omega_i \cap \Omega_j$ donc $\frac{\partial h_j}{\partial \bar{z}} = \frac{\partial h_i}{\partial \bar{z}}$ sur $\Omega_i \cap \Omega_j$. Donc il existe une fonction $\psi \in C^\infty(\Omega)$ telle que $\psi = \frac{\partial h_j}{\partial \bar{z}}$ sur Ω_j pour tout $j \in J$.
On sait qu'il existe une solution u de $\frac{\partial u}{\partial \bar{z}} = -\psi$. Ainsi, on a bien $g_j = h_j + u \in H(\Omega_j)$ pour tout j et $g_{ij} = g_j - g_i$ sur $\Omega_i \cap \Omega_j$ pour tout $(i, j) \in J^2$.

Ce théorème est en fait une forme plus forte du théorème de Mittag-Leffler, montrons en effet qu'il implique la forme bis du théorème :

Si $\Omega = \bigcup_j \Omega_j$ avec Ω_j des ouverts de C, soient $f_j \in M(\Omega_j)$ telles que $f_j - f_k \in H(\Omega_j \cap \Omega_k)$ pour tous j et k. On pose $g_{ij} = f_i - f_j$. On a bien $g_{ij} = -g_{ji}$ sur $\Omega_i \cap \Omega_j$ et $g_{ij} + g_{jk} + g_{ki} = 0$ sur $\Omega_i \cap \Omega_j \cap \Omega_k$, donc $\exists g_j \in H(\Omega_j)$ telle que

$$f_i - f_j = g_{ij} = g_j - g_i$$

sur $\Omega_i \cap \Omega_j$. Donc sur $\Omega_i \cap \Omega_j$, $f_i + g_i = f_j + g_j$. On a donc $f = f_i + g_i$ sur Ω_i pour tout i et $f - f_j = g_j \in H(\Omega_i)$ pour tout i.

3.4 Théorème de Weierstrass et applications

Théorème 15 (Weierstrass). *Soit $(z_j)_j$ une suite discrète de points de Ω, et n_j des entiers. Alors il existe une fonction méromorphe f telle que f soit holomorphe et non nulle mis à part en les z_j, et $z \to f(z)(z - z_j)^{-n_j}$ est holomorphe et non nulle au voisinage de z_j pour tout $j \in J$.*

Ainsi on peut trouver une fonction méromorphe f avec des zéros et des pôles d'ordres donnés prescrits à l'avance !

Preuve On commence par choisir une suite croissante de compacts $K_j \subset \Omega$ avec $K_j = \widehat{K}_j$ de sorte que tout compact de Ω soit inclus dans un K_j.
On choisira successivement des fonctions rationnelles f_j avec les zéros et pôles désirés dans K_j et des fonctions holomorphes sur Ω telles que $|f_{j+1}f_j^{-1}e^{g_j} - 1| < \epsilon_j$ sur K_j, avec $\sum_{j\in J} \epsilon_j < \infty$.
Supposons qu'on ait déjà choisi de telles $f_1, .., f_j$ et $g_1, .., g_{j-1}$. Soit f une fonction rationnelle avec les zéros et pôles prescrits sur K_{j+1}. On peut écrire

$$\frac{f}{f_j} = c \prod_{k \in L} (z - w_k)^{m_k}$$

avec L un ensemble fini, $w_k \in K_j^c$ pour tout j et les m_k sont des entiers. Comme aucune composante connexe de $\Omega \setminus K$ n'est relativement compacte dans Ω, on peut choisir pour tout k des $\zeta_k \in K_{j+1}^c$ dans la même composante connexe de K_j^c que w_k. La fonction $f_{j+1} = f \prod_{k\in L} (z - \zeta_k)^{-m_k}$ a ainsi les bons zéros et pôles sur K_{j+1}, et

$$\log\left(\frac{f_{j+1}(z)}{f_j(z)}\right) = \log c + \sum_{k\in L} m_k \log\left(\frac{z - w_k}{z - \zeta_k}\right)$$

peut être défini comme une fonction holomorphe sur un voisinage de K_j puisque w_k et ζ_k sont dans la même composante connexe de K_j^c. On peut par Runge choisir $g_j \in H(\Omega)$ telle que $|\log(\frac{f_{j+1}}{f_j}) + g_j| < \log(1+\epsilon_j)$, d'où $|f_{j+1}f_j^{-1}e^{g_j} - 1| < \epsilon_j$ sur K_j. Et finalement on a

$$\lim_{j\to\infty} f_{j+1}\prod_{k=1}^{j} e^{g_k} = f_1 \prod_{k=1}^{\infty} \frac{f_{j+1}}{f_j} e^{g_j}$$

ce qui définit une fonction f avec les propriétés requises, le produit débutant au terme j convergeant vers une fonction holomorphe à l'intérieur de K_j.

Corollaire 9. *Toute fonction méromorphe F sur Ω peut s'écrire sous la forme $\frac{f}{g}$ avec f et g holomorphes sur Ω.*

Preuve Si F a des pôles z_j d'ordres n_j, le théorème de Weierstrass donne une fonction holomorphe g dont les zéros sont les z_j d'ordre n_j. $f = Fg$ est alors holomorphe, d'où le résultat.

Corollaire 10. *Il existe une fonction $f \in H(\Omega)$ qui ne peut être prolongée holomorphiquement à aucun ensemble plus grand, même en fonction méromorphe.*

Preuve On note les points de coordonnées rationnelles de Ω en une suite (z_j) telle que chaque point apparaisse une infinité de fois. On pose $r_j = d(z_j, \Omega^c)$.On choisit une suite croissante de compacts $K_j \subset \Omega$ avec $K_j = \widehat{K}_j$ de sorte que tout compact de Ω soit inclus dans un K_j. On prend ensuite pour chaque j un point $w_j \in K_j^c$ tel que $|w_j - r_j| < r_j$.
Comme la suite $(w_j)j$ est discrète dans Ω, le théorème de Weierstrass nous donne une fonction f holomorphe sur Ω qui a pour uniques zéros les w_j. Si $a \in \Omega$ a des coordonnées rationnelles, on note $r = d(a, \Omega^c)$. Le disque $D(a,r)$ contient une infinité de points w_j puisque $a = z_j$ pour une infinité de j.
Donc f ne peut être prolongée en une fonction méromorphe sur aucun disque contenant $\overline{D}(a,r)$, puisque les zéros d'une fonction méromorphe non identiquement nulle sont isolés.

Corollaire 11. *Soient $(z_j)_j$ une suite discrète de points de Ω, f_j holomorphe au voisinage de z_j, et n_j des entiers positifs. Alors il existe $f \in H(\Omega)$ telle que, pour tout j :*

$$f(z) - f_j(z) = O_{z\to z_j}(|z - z_j|^{n_j+1})$$

Preuve Soit $g \in H(\Omega)$ avec pour zéros d'ordre $n_j + 1$ les z_j. Par Mittag-Leffler, il existe une fonction méromorphe h dont les uniques pôles sont les z_j et telle que $h - \frac{f_j}{g}$ est holomorphe au voisinage de z_j pour tout j. Alors la fonction $f = gh$ convient.

Chapitre 4

Fonctions sous-harmoniques

4.1 Définition et propriétés

Définition 7 (Fonction harmonique)**.** *Une fonction* $h \in C^2(\Omega)$ *est harmonique si et seulement si* $\Delta h = 4\frac{\partial^2 h}{\partial z \partial \bar{z}} = 0$.

Définition 8 (Fonction sous-harmonique)**.** u *de* $\Omega \to [-\infty; +\infty)$ *est sous-harmonique si et seulement si :*
u est SCS(semi-continue par valeurs supérieures) : $\forall s \in R$,$\{z \in \Omega; u(z) < s\}$ *est un ouvert.*
$\forall K \subset\subset \Omega$ *et* $\forall h$ *continue sur* K *harmonique à l'intérieur de* K*, si* $h \geq u$ *sur* ∂K*, alors* $h \geq u$ *sur* K *entier.*

Théorème 16. *(i)u sous-harmonique et* $c \in R^{+*} \Rightarrow cu$ *sous-harmonique.*
(ii)$(u_\alpha)_{\alpha \in A}$ *famille de fonctions sous-harmoniques* $\Rightarrow u = \sup_{\alpha \in A} u_\alpha$ *est sous-harmonique si* $u < +\infty$*, et* u *est SCS.*
(iii)$(u_n)_{n \in N}$ *suite décroissante de fonctions sous-harmoniques* $\Rightarrow u = \lim_{j \to +\infty} u_j$ *sous-harmonique.*

Preuve (i)$\forall s \in R$, $\{z \in \Omega; cu(z) < s\} = \{z \in \Omega; u(z) < \frac{s}{c}\}$ est un ouvert, donc cu est SCS. Soit $K \subset\subset \Omega$ et h continue sur K et harmonique à l'intérieur de K avec $h \geq cu$ sur ∂K. $\frac{h}{c}$ est continue sur K et harmonique à l'intérieur de K avec $\frac{h}{c} \geq u$ sur ∂K donc $\frac{h}{c} \geq u$ sur K entier, d'où cu sous-harmonique.

(ii)$\forall s \in R$,$\{z \in \Omega; u(z) < s\} = \bigcup_{\alpha \in A}\{z \in \Omega; u_\alpha(z) < s\}$ est un ouvert en tant qu'union quelconque d'ouverts, donc u est SCS. Soit $K \subset\subset \Omega$ et h continue sur K harmonique à l'intérieur de K, avec $h \geq u$ sur ∂K. On a, $\forall \alpha \in A$, $h \geq u_\alpha$ sur ∂K donc sur K entier, donc $h \geq u$ sur K entier, donc u est sous-harmonique.

(iii)$\forall s \in R$,$\{z \in \Omega; u(z) < s\} = \bigcup_{j=1}^{\infty}\{z \in \Omega; u_j(z) < s\}$ est un ouvert en tant qu'union quelconque d'ouverts, donc u est SCS. Soit $K \subset\subset \Omega$ et h continue sur K harmonique à l'intérieur de K, avec $h \geq u$ sur ∂K. Soit $\epsilon > 0$, on pose $K_j = \{z \in \partial K; u_j(z) \geq h(z)+\epsilon\}$. C'est une suite décroissante de compacts, $\bigcap_{j=1}^{\infty} K_j = \emptyset$, donc $\exists j_0$ tel que pour tout $j \geq j_0$, $K_j = \emptyset$. Donc $u_j < h + \epsilon$ sur ∂K donc sur K pour tout $j \geq j_0$, donc $h \geq u$ sur K, donc u est sous-harmonique.

Théorème 17. *Soit $u : \Omega \to [-\infty; +\infty)$ une fonction SCS. Les assertions suivantes sont des conditions nécessaires et suffisantes pour que u soit sous-harmonique :*
(i)Pour tout disque fermé $D \subset \Omega$ et f polynôme tel que $u \leq Re(f)$ sur ∂D, $u \leq Re(f)$ sur D.
(ii)Si $\Omega_\delta = \{z; d(z, \Omega^c) > \delta\}$,on a, pour toute mesure μ positive sur $[0, \delta]$, pour tout $z \in \Omega_\delta$,

$$2\pi u(z) \int d\mu(r) \leq \int \int_0^{2\pi} u(z + re^{i\theta}) d\theta d\mu(r)$$

(iii)Pour tout $\delta > 0$ et tout $z \in \Omega_\delta$, il existe une mesure positive μ à support dans $[0, \delta]$ qui charge des points en dehors de l'origine et telle que pour tout $z \in \Omega_\delta$,

$$2\pi u(z) \int d\mu(r) \leq \int \int_0^{2\pi} u(z + re^{i\theta}) d\theta d\mu(r)$$

Preuve Il est clair que u sous-harmonique implique (i) et que (ii) implique (iii).
(i) $\Rightarrow$ (ii)
Soit $z \in \Omega_\delta$ et $0 < r \leq \delta$. On pose $D = \overline{D}(z, r) \subset \Omega$.
Si $\phi(\theta) = \sum a_k e^{ik\theta}$ est un polynôme trigonométrique tel que $u(z+re^{i\theta} \leq \phi(\theta)$ pour tout θ, le polynôme $f(\zeta) = a_0 + 2\sum_{k>0} a_k \frac{(\zeta - z)^k}{r^k}$ vérifie $Re(f) \geq u$ sur ∂D, donc sur D, et en particulier

$$u(z) \leq a_0 = \frac{1}{2\pi} \int_0^{2\pi} \phi(\theta) d\theta$$

Si maintenant ϕ est une fonction continue telle que $u(z + re^{i\theta} \leq \phi(\theta)$, pour tout $\epsilon > 0$ on peut trouver un polynôme trigonométrique ϕ_1 avec $\phi \leq \phi_1 \leq \phi + \epsilon$, donc pour toute fonction ϕ continue telle que $u(z + re^{i\theta} \leq \phi(\theta)$, on a

$$u(z) \leq \frac{1}{2\pi} \int_0^{2\pi} \phi(\theta) d\theta$$

Donc comme u est SCS,

$$u(z) \leq \frac{1}{2\pi} \int_0^{2\pi} u(z + re^{i\theta}) d\theta$$

D'où (ii) en intégrant.

(iii) $\Rightarrow u$ sous-harmonique
Soit $K \subset\subset \Omega$ et h continue sur K harmonique à l'intérieur de K, avec $h \geq u$ sur ∂K. On pose $v = u - h$ et $M = \sup_K v$. Si $M > 0$, on a v SCS donc $v = M$ sur un compact F non vide inclus dans l'intérieur de K. Soit $z_0 \in F$ tel que la distance de z_0 à K soit minimale. Si cette distance est $> \delta$, alors tout cercle centré en z_0 et de rayon $r \leq \delta$ contient des points sur lesquels $v < M$, c'est même un arc de cercle entier puisque v est semi-continue. On a donc

$$\int\int v(z_0 + re^{i\theta})d\theta d\mu(r) < 2\pi M \int d\mu(r) = 2\pi v(z_0) \int d\mu(r)$$

pour μ une mesure comme dans (iii). Mais cela contredit (iii), d'où le résultat.

Corollaire 12. *u_1 et u_2 sous-harmoniques $\Rightarrow u_1 + u_2$ sous-harmonique.*

Corollaire 13. *Une fonction u est sous-harmonique si chaque point de Ω admet un voisinage où u est sous-harmonique. Autremnt dit, la sous-harmonicité est une propriété locale.*

Corollaire 14. *Si $f \in H(\Omega)$ alors $\log |f|$est sous-harmonique sur Ω.*

Preuve Cela découle du principe du maximum en utilisant le point (i) du théorème précédent.

Ceci suggère donc que $|f|$ est sous-harmonique si f est holomorphe. En fait, on a le résultat suivant :

Théorème 18. *Soit ϕ une fonction convexe croissante sur R. On pose $\phi(-\infty) = \lim_{x\to-\infty} \phi(x)$. Si u est sous-harmonique, alors $\phi(u)$ est sous-harmonique.*

Preuve Soit $x_0 \in R$. $\forall x \in R$, $\exists k \in R$ tel que $\phi(x) \geq \phi(x_0) + k(x - x_0)$. Donc

$$\frac{1}{2\pi}\int_0^{2\pi} \phi(u(z + re^{i\theta}))d\theta \geq \phi(x_0) + k(\frac{1}{2\pi}\int_0^{2\pi} u(z + re^{i\theta})d\theta - x_0)$$

On prend alors $x_0 = \frac{1}{2\pi}\int_0^{2\pi} u(z + re^{i\theta})d\theta$, on obtient :

$$\frac{1}{2\pi}\int_0^{2\pi} \phi(u(z + re^{i\theta}))d\theta \geq \phi(\frac{1}{2\pi}\int_0^{2\pi} u(z + re^{i\theta})d\theta) \geq \phi(u(z))$$

par croissance de ϕ et sachant que u est sous-harmonique. De plus $\phi(u)$ est SCS, d'où le résultat.

Corollaire 15. *Soient $u_1, u_2 \geq 0$ tels que $\log u_1$ et $\log u_2$ soient sous-harmoniques ($\log 0 = -\infty$). Alors $\log(u_1 + u_2)$ est sous-harmonique.*

Preuve Soit D un disque de Ω et f un polynôme tel que $Re(f) \geq \log(u_1 + u_2)$ sur ∂D. On a donc $u_1 + u_2 \leq |e^f|$ sur ∂D. Puisque $\log u_i - Re(f)$ est sous-harmonique et comme exp est croissante et convexe, $u_i|e^{-f}|$ est sous-harmonique. Donc $(u_1+u_2)|e^{-f}|$ est sous-harmonique, donc $(u_1+u_2)|e^{-f}| < 1$ sur D, d'où $Re(f) \geq \log(u_1 + u_2)$ sur D, d'où le résultat.

Théorème 19. *Soit u sous-harmonique non identiquement égale à $-\infty$ sur aucune des composantes connexes de Ω. Alors $u \in L^1_{loc}(\Omega)$, ce qui implique que $u > -\infty$ presque partout.*

Preuve Soit $z \in \Omega$ tel que $u(z) > -\infty$ et D un disque fermé de centre z contenu dans Ω, u est intégrable sur D. Si $E = \{z \in \Omega; u$ est intégrable au voisinage de $z\}$, on a $u = -\infty$ au voisinage de chaque point de $\Omega \setminus E$. Donc E et $\Omega \setminus E$ sont tous deux ouverts donc $\Omega \setminus E$ est une union de composantes de Ω, mais ces composantes sont forcément vide par hypothèse. D'où $\Omega = E$.

On donne maintenant une autre description des fonctions sous-harmoniques :

Théorème 20. *Soit u sous-harmonique non identiquement égale à $-\infty$ sur aucune des composantes connexes de Ω. On a, pour tout $v \in C^2_0(\Omega)$ positif :*

$$\int u \Delta v d\lambda \geq 0$$

(λ désigne la mesure de Lebesgue)

Preuve Soit $0 < r < d(supp(v), \Omega^c)$, pour tout $z \in supp(v)$, on a

$$a\pi u(z) \leq \int_0^{2\pi} u(z + re^{i\theta}) d\theta$$

On multiplie ensuite par v et on intègre par rapport à z, on obtient (après un changement de variables) :

$$\int u(z)(\int_0^{2\pi} v(z - re^{i\theta}) d\theta - 2\pi v(z)) d\lambda(z) \geq 0$$

On obtient finalement l'inégalité désirée en multipliant par $\frac{2}{\pi r^2}$ et, après un développement de Taylor de v, en faisant tendre r vers 0.

Si $u \in C^2(\Omega)$, en intégrant cette inégalité par parties, on obtient :

Corollaire 16. *Soit $u \in C^2(\Omega)$, u est sous-harmonique si et seulement si $\Delta \geq 0$ (on dit que u est strictement sous-harmonique si $\Delta u > 0$)*

Preuve Réciproquement : si $\Delta \geq 0$: On a $\Delta u \geq 0$, donc :

$$\int (\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2})u(z + re^{i\theta})d\theta$$

Soit $M(r) = \frac{1}{2\pi}\int_0^{2\pi} u(z + re^{i\theta})d\theta$, on a donc $M''(r) + \frac{1}{r}M'(r) \geq 0$ donc $r \to rM'(r)$ est croissante, comme elle tend vers 0 quand r tend vers 0, on obtient que $M'(r) \geq 0$, donc M est croissante. On a donc $M(0) \leq M(r)$, d'où u sous-harmonique.

Théorème 21. *Soit $u \in L^1_{loc}(\Omega)$ tel que l'inégalité du théorème précédent soit vérifiée pour u. Alors $\exists! U$ fonction sous-harmonique de Ω qui vaut u presque partout. Si ϕ est une fonction de $|z|$ positive, intégrable et à support compact, alors pour tout $z \in \Omega$*

$$U(z) = \lim_{\delta \to 0} \frac{\int u(z - \delta z')\phi(z')d\lambda(z')}{\int \phi(z')d\lambda(z')}$$

Preuve Si U est sous-harmonique, on a pour δ petit :

$$U(z) \leq \frac{\int U(z - \delta z')\phi(z')d\lambda(z')}{\int \phi(z')d\lambda(z')}$$

et comme U est SCS, la limite sup du membre de droite est inférieure à $U(z)$ quand $\delta \to 0$. D'où l'inégalité si $u = U$ presque partout.
Soit $\phi \in C_0^\infty(C)$ dont le support est inclus dans le disque unité telle que $\phi \geq 0$ et ϕ ne dépende que de $|z|$. On pose $u_\delta = \frac{\int u(z-\delta z')\phi(z')d\lambda(z')}{\int \phi(z')d\lambda(z')}$, u_δ est dans $C^\infty(\Omega_\delta)$ et $u_\delta \to u$ en norme L^1 quand δ tend vers 0 sur les compacts de Ω.
On remarque que l'inégalité du théorème précédent est aussi vérifiée pour u_δ, donc $\Delta u_\delta \geq 0$ donc u_δ est sous-harmonique, donc $\frac{\int u(z-\epsilon z')\phi(z')d\lambda(z')}{\int \phi(z')d\lambda(z')}$ décroît quand ϵ décroît vers 0. Si on fait tendre δ vers 0, on constate que $u_\epsilon(z)$ décroît quand ϵ décroît vers 0. Donc $U(z) = \lim_{\delta\to 0} u_\epsilon(z)$ existe et est sous-harmonique par le théorème 16. De plus $u_\epsilon \to u$ dans $L^1_{loc}(\Omega)$, donc $u = U$ presque partout.

Théorème 22. *Si $0 \leq f \in C^2$ et $\log f$ est sous-harmonique, la fonction $\log(1 + f)$ est strictement sous-harmonique sauf là où $grad(f) = \Delta f = 0$.*

Preuve Comme $\log f$ est sous-harmonique, $f\Delta f - |grad(f)|^2 \geq 0$. Si on note $f_1 = 1 + f$, on a $f_1\Delta f_1 - |grad(f_1)|^2 = \Delta f + (f\Delta f - |grad(f)|^2)$, ce qui ne peut valoir 0 que si $\Delta f = 0$, ce qui implique $grad(f) = 0$

Théorème 23. *Soit $(v_k)_k$ une suite de fonctions sous-harmoniques de Ω uniformément bornées sur tout compact de Ω. On suppose que $\limsup_{k\to+\infty} v_k(z) \leq C$ pour tout $z \in \Omega$. Pour tout $\epsilon > 0$ et tout compact $K \subset \Omega$, il existe k_0 tel que $v_k(z) \leq C + \epsilon$ pour tout $z \in K$ et $k > k_0$*

Preuve Comme on peut remplacer Ω par un sous-ensemble relativement compact contenant K, on peut supposer que la suite est uniformément bornée dans Ω, ou même $v_k \leq 0$ dans Ω pour tout k. On a

$$\pi r^2 v_k(z) \leq \int_{|z-z'|<r} v_k(z')d\lambda(z')$$

par Fatou, la limite sup de ce qui est à droite est inférieure à $\pi C r^2$. Donc pour tout $z \in K$ on peut choisir k_0 tel que pour tout $k > k_0$:

$$\int_{|z-z'|<r} v_k(z')d\lambda(z') \leq \pi(C + \frac{\epsilon}{2})r^2$$

Comme $v_k \leq 0$, on a, pour $|z - w| < \delta < r$,

$$\pi(r+\delta)^2 v_k(w) \leq \int_{|z-z'|<r+\delta} v_k(z')d\lambda(z') \leq \int_{|z-z'|\leq r} v_k(z')d\lambda(z')$$

Si δ est assez petit, on a donc $v_k(w) < C + \epsilon$ si $k > k_0$ et $|w - z| < \delta$

4.2 Retour à l'équation $\frac{\partial u}{\partial \bar{z}} = f$

Etant donné une fonction sous-harmonique ϕ, on cherche à fabriquer une solution de $\frac{\partial u}{\partial \bar{z}} = f$ qui vérifie en plus

$$\int \frac{|u|^2}{(1+|z|^2)^2} e^{-\phi} \leq C \int |f|^2 e^{-\phi}$$

On peut démontrer le résultat suivant :

Théorème 24. *Pour ϕ une fonction SCS, pour Ω un ouvert de C, on pose $L^2(\Omega, e^{-\phi}) = \{g$ mesurable telle que $\int_\Omega |g|^2 e^{-\phi} < +\infty\}$ Soit ϕ une fonction C^∞ vérifiant $\Delta\phi \geq 0$, pour toute fonction $f \in L^2(\Omega, e^{-\phi})$ et $a > 0$, il existe une solution u de $\frac{\partial u}{\partial \bar{z}} = f$ telle que*

$$a \int |u|^2 (1+|z|^2)^{-a} e^{-\phi} \leq \int |f|^2 (1+|z|^2)^{2-a} e^{-\phi}$$

Preuve Soit $v \in C_0^\infty(\Omega)$ Pour un tel v, on note

$$\delta v = -e^{\phi} \frac{\partial v e^{-\phi}}{\partial z} = -\frac{\partial v}{\partial z} + v \frac{\partial \phi}{\partial z}$$

On suppose d'abord que $\Delta\phi > 0$ et on cherche une fonction $\psi \in C^\infty$ à valeurs dans R telle qu'on ait l'inégalité : (pour un $C \in R$)

$$\int |v|^2 e^{\psi - 2\phi} \leq C \int |\delta v|^2 e^{-\phi}$$

On a

$$\frac{\partial(\delta v)}{\partial\overline{z}} = -\frac{\partial^2 v}{\partial z\partial\overline{z}} + \frac{\partial v}{\partial\overline{z}}\frac{\partial\phi}{\partial z} + \frac{v}{4}\Delta\phi = \delta(\frac{\partial v}{\partial\overline{z}}) + \frac{v}{4}\Delta\phi = g$$

En intégrant par parties, sachant que le terme de bord est nul puisque v est à support compact, on a :

$$\int_\Omega |\delta v|^2 e^{-\phi} = \int_\Omega g\overline{v}e^{-\phi}$$

Donc

$$\int_\Omega |\delta v|^2 e^{-\phi} = \int_\Omega \delta(\frac{\partial v}{\partial\overline{z}})\overline{v}e^{-\phi} + \frac{1}{4}\int_\Omega |v|^2\Delta\phi e^{-\phi}$$

On a de plus $\int_\Omega \delta(\frac{\partial v}{\partial\overline{z}})\overline{v}e^{-\phi} = \int_\Omega |\frac{\partial v}{\partial\overline{z}}|^2 e^{-\phi} \geq 0$, donc on a l'inégalité voulue pour ψ telle que $e^\psi = \frac{\Delta\phi}{4}e^\phi$.

On considère désormais la forme linéaire de $L^2(\Omega, e^{-\phi})$: $\Phi : \delta v \to \int f\overline{v}e^{-\phi} =< f, v >$, montrons qu'elle est continue pour la norme induite de $L^2(\Omega, e^{-\phi})$.
Par Cauchy-Schwarz, $|\Phi(\delta v)|^2 = |\int f\overline{v}e^{-\phi}|^2 = |\int fe^{-\frac{\psi}{2}}\overline{v}e^{\frac{\psi}{2}-\phi}|^2 \leq \big(\int |f|^2 e^{-\psi}\big)\big(\int |v|^2 e^{\psi-2\phi}\big)$
Donc

$$|\Phi(\delta v)|^2 \leq \Big(\int |f|^2 e^{-\psi}\Big)\Big(\int |\delta v|^2 e^{-\phi}\Big)$$

Donc Φ est une forme linéaire continue de $L^2(\Omega, e^{-\phi})$, donc par le théorème de représentation de Riesz, il existe un unique $u \in L^2(\Omega, e^{-\phi})$ tel que, pour tout v, $\int f\overline{v}e^{-\phi} =< u, \delta v >$, or par intégration par parties on a $< u, \delta v >=< \frac{\partial u}{\partial\overline{z}}, v >$, donc on a $\frac{\partial u}{\partial\overline{z}} = f$.

Reste maintenant à prouver l'inégalité souhaitée. On vient de prouver que $|\int ue^{-\frac{\phi}{2}}\overline{\delta v}e^{-\frac{\phi}{2}}|^2 \leq \big(\int |f|^2 e^{-\psi}\big)\big(\int |\delta v|^2 e^{-\phi}\big)$, d'où

$$\int |u|^2 e^{-\phi} \leq \int |f|^2 e^{-\psi}$$

Pour conclure, il reste à appliquer ce qui précède à la fonction $\phi_1 = \phi + a\log(1 + |z|^2)$. En effet on a $\frac{\Delta\phi_1}{4} = \frac{\Delta\phi}{4} + \frac{a}{(1+|z|^2)^2} > 0$. On a donc existence d'une solution u qui vérifie

$$\int |u|^2 e^{-\phi_1} \leq \int |f|^2 e^{-\psi}$$

on a $e^{-\psi} = e^{-\phi_1}\frac{4}{\Delta\phi_1}$, or $\frac{4}{\Delta\phi_1} \leq \frac{(1+|z|^2)^2}{a}$ car $\Delta\phi \geq 0$. En remplaçant ϕ_1 par son expression on en déduit bien l'inégalité attendue :

$$a\int |u|^2(1+|z|^2)^{-a}e^{-\phi} \leq \int |f|^2(1+|z|^2)^{2-a}e^{-\phi}$$

Conclusion

Ce T.E.R. m'aura permis de découvrir des résultats intéressants de théorie des fonctions holomorphes et sur les fonctions sous-harmoniques, mais aussi de renforcer mes connaissances en matière de calcul différentiel et de topologie générale, deux domaines dont l'importance a été sous-jacente mais capitale pour démontrer ces résultats. Je tiens à remercier Monsieur Sibony dont l'aide m'aura été précieuse tout au long de ces travaux.

Annexe C

Mémoire de M2

Mémoire de M2 de Romain Mismer : Analyse bayésienne du modèle de suite gaussienne parcimonieuse

Introduction

En statistiques, de nombreux problèmes s'intéressent à un paramètre en très grande dimension sur lequel on ne dispose que d'une seule observation, par exemple après avoir moyenné, et sujette à un bruit. Le modèle auquel nous allons nous intéresser dans ce document est le suivant : on considère les observations $X = (X_1, \ldots, X_n)$ telles que, $\forall i \in \{1, \ldots, n\}$:

$$X_i = \theta_i + \varepsilon_i$$

où les ε_i représentent le bruit, on les considèrera donc comme des variables aléatoires de loi normale standard. On notera $\theta = (\theta_1, \ldots, \theta_n)$ notre paramètre d'intérêt.
Il apparaît tout de suite que, sans en connaître un minimum sur θ, il est compliqué de l'estimer efficacement.
Dans ce document, on supposera l'existence d'une certaine parcimonie dans cette suite θ. Cette hypothèse de parcimonie se rencontre dans de nombreux contextes, par exemple en astronomie, en traitement de l'image, en sélection de modèle, en data mining ou encore lorsque l'on fait de l'estimation non-paramétrique de fonction en utilisant des ondelettes.
Une approche naturelle à ces problèmes parcimonieux peut être le seuillage : si la valeur absolue d'un certain X_i dépasse un certain seuil t, alors on suppose qu'il correspond à un θ_i non nul et on l'estime alors généralement par X_i lui-même ; sinon, on l'estime comme étant zéro.
Ce document traitera d'un cadre bayésien : on considèrera une loi a priori sur les θ_i qui permettra d'induire de la parcimonie, et on étudiera certains aspects de la loi a posteriori de θ sachant X. Dans [1], Johnstone et Silverman mettent en valeur le fait que, sous certaines hypothèses, certains aspects de l'a posteriori sont de "bons" estimateurs en un certain sens. En effet, la médiane a posteriori est alors naturellement un estimateur de type seuillage et qui a en plus une propriété de rétrécissement, et la moyenne a posteriori, bien que n'ayant pas la propriété de seuillage, est tout de même un "bon" estimateur de par sa propriété de rétrécissement.
Dans une première partie on présentera donc certains résultats de [1], tout en s'intéressant à un autre aspect de la loi a posteriori, qui s'avère être la loi a posteriori elle-même ! On montrera qu'elle possède elle aussi de bonnes propriétés de convergence. Dans une seconde partie, on s'intéressera à une loi a priori plus générale, qui s'avère inclure la forme de loi a priori présentée dans la première partie. Encore une fois, la loi a posteriori aura, sous certaines hypothèses, de bonnes propriétés de convergence.

Dans une troisième partie, on s'attachera à simuler certains aspects de la loi a posteriori pour différents a priori. Enfin, une quatrième partie permettra de conclure et d'ouvrir sur quelques perspectives. En fin de document on trouvera également les preuves des principaux résultats énoncés au long de ce mémoire.

Table des matières

Chapitre 1

Une Première Forme de Loi a Priori

1.1 Présentation du cadre

Le modèle de suite gaussienne s'écrit, avec X le vecteur de $\mathbb{R}^n$ observé :

$$X_i = \theta_{0,i} + \varepsilon_i, \quad i = 1, \dots n, \tag{1.1}$$

où $\varepsilon_1, \dots, \varepsilon_n$ sont indépendantes identiquement distribuées (iid) de loi $N(0,1)$ (de densité notée ϕ), et le paramètre $\theta_0 = (\theta_{0,1}, \dots, \theta_{0,n})$ appartient à la classe $\ell_0[p_n]$ définie par

$$\ell_0[p_n] = \{\theta \in \mathbb{R}^n, \ |\{i \in \{1, \dots n\}, \ \theta_i \neq 0\}| \leq p_n\},$$

pour $0 \leq p_n \leq n$, où $|A|$ est la cardinalité de l'ensemble A.
On supposera $p_n = o(n)$ quand $n \to \infty$.
On note E_{θ_0} l'espérance par rapport à la loi sur les données engendrée par (1.1). On note $\|\cdot\|$ la norme euclidienne, $\|v\|^2 = \sum_{i=1}^n v_i^2$ pour $v \in \mathbb{R}^n$. On notera $p_{n,\theta}(X) = (2\pi)^{-n/2} e^{-\|X-\theta\|^2/2}$

Un résultat dû à Donoho, Johnstone et al [8] sur la vitesse minimax, dont la preuve est présentée en fin de document, est le suivant :

Théorème 1. *[Vitesse minimax] On note* $R_{n,2}(\ell_0[p_n]) = \inf_{\hat{\theta}} \sup_{\theta \in \ell_0[p_n]} \frac{1}{n} \sum_{i=1}^n E_\theta(\hat{\theta}_i - \theta_i)^2$.
On a alors $R_{n,2}(\ell_0[p_n]) = \frac{2p_n}{n} \log(\frac{n}{p_n})(1 + o(1))$ *quand* $n \to \infty$

1.2 Une loi a priori naturelle

On considère l'a priori suivant sur θ, un a priori produit qui traduit la parcimonie du modèle :

$$\theta \sim \Pi := \bigotimes_{i=1}^n (1-\alpha)\delta_0 + \alpha\gamma \tag{1.2}$$

avec δ_0 le Dirac en 0, γ une densité de probabilité à choisir et un paramètre $\alpha \in [0;1]$ à fixer. Plus ce paramètre est proche de 0 et plus le modèle est parcimonieux.

1.3 Premières propriétés

1.3.1 Expression de la loi a posteriori

La loi a posteriori s'écrit elle aussi comme un produit. En effet, en notant $g = \phi * \gamma$:

$$\Pi(\theta|X) = \prod_{i=1}^{n} \frac{[(1-\alpha)\delta_0(\theta_i) + \alpha\gamma(\theta_i)]\phi(X_i - \theta_i)}{(1-\alpha)\phi(X_i) + \alpha g(X_i)}$$

On obtient donc :

$$\Pi(\theta|X) = \prod_{i=1}^{n}(1-\tilde{\alpha})\delta_0(\theta_i) + \tilde{\alpha}\psi_{X_i}(\theta_i) \tag{1.3}$$

avec $\tilde{\alpha} = \frac{\alpha g(X_i)}{(1-\alpha)\phi(X_i)+\alpha g(X_i)}$ et la densité $\psi_{X_i} = \frac{\phi(X_i - \cdot)\gamma(\cdot)}{g(X_i)}$

On peut alors montrer que la médiane a posteriori est un estimateur de type seuillage, mais que ce n'est pas le cas malheureusement pour la moyenne a posteriori.
On peut désormais se demander ce que l'on pourrait choisir pour la densité γ. Le premier réflexe serait de prendre une densité gaussienne.

Cas où γ est la densité de $\mathcal{N}(0, \sigma^2)$

Dans ce cas, pour tout $i \in \{1, \ldots, n\}$, ψ_{X_i} est également la densité d'une loi normale, de moyenne $\frac{\sigma^2}{1+\sigma^2}X_i$ et de variance $\frac{\sigma^2}{1+\sigma^2}$.
On obtient ainsi pour l' estimateur de moyenne a posteriori :

$$\hat{\theta}^{moy} = \frac{\tilde{\alpha}\sigma^2}{1+\sigma^2}X$$

Si on suppose maintenant que $\sigma^2 = 1$ et en faisant l'hypothèse que $\alpha = 1$, on obtient $\tilde{\alpha} = 1$ et $\hat{\theta}^{moy} = \frac{X}{2}$
Considérons alors $\theta_0 = 0 \in \ell_0[p_n]$ et regardons $E_{\theta_0}[\|\hat{\theta}^{moy} - \theta_0\|^2]$:

$$E_{\theta_0}[\|\hat{\theta}^{moy} - \theta_0\|^2] = \frac{1}{4}\sum_{i=1}^{n} E_{\theta_0}[X_i^2] = \frac{n}{4}$$

ce qui est très loin de la vitesse minimax $\frac{2p_n}{n}\log(\frac{n}{p_n})$.. Ceci est en partie dû au fait que le choix $\alpha = 1$ est loin d'un choix optimal, puisque cela retire la parcimonie du modèle.
En revanche, quel que soit le choix de α, on a toujours $|\hat{\theta}^{moy}| \leq \frac{|X|}{2}$, ce qui n'est intuitivement jamais très "bon" pour estimer tout coefficient non nul de θ_0..

En réalité, les hypothèses des propriétés suivantes démontrées par Johnstone et Silverman excluent le cas gaussien :

Propriétés de la médiane a posteriori

Hypothèses sur le prior Dans [1], il est mis en avant qu'il est bien plus avantageux pour la densité γ d'être une densité à queues suffisamment lourdes, d'où aussi le choix dans la suite d'une densité de Laplace plutôt qu'une loi normale standard. Le fait que ceci soit nécessaire est prouvé dans [3] (pour la mesure a posteriori complète).
Précisément, on suppose désormais que

$$\sup_{u>0} |\frac{d}{du} \log \gamma(u)| = \Lambda < \infty$$

On obtient alors que $\forall u > 0, \log \gamma(u) \geq \log \gamma(0) - \Lambda u$ et donc, $\forall u > 0 : \gamma(u) \geq \gamma(0)e^{-\Lambda|u|}$, ce qui élimine le choix de γ gaussienne.
On supposera de plus que $u^2\gamma(u)$ est bornée pour tout u et qu'il existe $\kappa \in [1, 2]$ tel que, quand $y \to \infty$:

$$\frac{1}{\gamma(y)} \int_y^\infty \gamma(u)du \asymp y^{\kappa-1}$$

Propriétés de la médiane a posteriori Sous ces hypothèses, la médiane a posteriori, notée $\hat{\theta}^{med}$, a les propriétés suivantes :

Propriété 1. *$\hat{\theta}^{med} = \hat{\theta}^{med}(x, \alpha)$ est une fonction croissante en x, antisymétrique et qui possède la règle de rétrécissement :*

$$\forall x \geq 0 \; , \; 0 \leq \hat{\theta}^{med}(x, \alpha) \leq x \tag{1.4}$$

De plus c'est un estimateur de type seuillage : il existe $t = t(\alpha) > 0$ tel que

$$\theta^{med}(x, \alpha) = 0 \Leftrightarrow |x| \leq t(\alpha) \tag{1.5}$$

On a également la propriété de rétrécissement borné : il existe $b > 0$ tel que

$$\forall x, \alpha \; , \; |\hat{\theta}^{med}(x, \alpha) - x| \leq t(\alpha) + b \tag{1.6}$$

Johnstone et Silverman en déduisent de "bons" résultats de convergence (minimax) pour la médiane a posteriori, mais également pour la moyenne a posteriori qui possède malgré tout la propriété de rétrécissement borné. Ils comparent dans leurs simulations ces estimateurs à des estimateurs de seuillage classiques et obtiennent de meilleurs résultats en général, notamment pour la médiane a posteriori.
On va pouvoir désormais s'intéresser à un autre aspect de la loi a posteriori : la loi a posteriori complète.

1.4 Lien entre le seuil t et le paramètre α

On commence par établir un lien entre t et α, ce qui nous permettra de travailler non plus exclusivement avec la médiane, mais avec une loi a priori appropriée au seuil classique que nous voulons utiliser.

On commence par montrer que $P(\theta > 0|X = t) = \frac{1}{2}$:

On ne se place ici que sur une seule coordonnée.
On a $\hat{\theta}^{med}(x) = f(x)\mathbf{1}_{|x|>t}$, avec f une fonction continue.
De plus $\hat{\theta}^{med}(t) = \hat{\theta}^{med}(-t) = 0$.

$\forall x \in \mathbb{R}$, $P(\theta \leq \hat{\theta}^{med}(x)|X = x) \leq \frac{1}{2}$
En évaluant en $x = t$, on obtient $P(\theta \leq 0|X = t) \leq 1/2$, d'où $P(\theta > 0|X = t) \geq \frac{1}{2}$

$\forall \varepsilon > 0$, on a $P(\theta \leq \hat{\theta}^{med}(t) + \varepsilon|X = t) \geq \frac{1}{2}$
En faisant tendre ε vers 0, par continuité à droite, on a $P(\theta \leq 0|X = t) \geq \frac{1}{2}$. D'où $P(\theta > 0|X = t) \leq \frac{1}{2}$, et finalement $P(\theta > 0|X = t) = \frac{1}{2}$

Lien entre α et t

On prend γ Laplace standard, soit $\gamma(x) = \frac{1}{2}e^{-|x|}$.
On note $\forall x \in \mathbb{R}$, $g_+(x) = \displaystyle\int_0^{+\infty} \phi(x-u)\gamma(u)du$ et $g_-(x) = \displaystyle\int_{-\infty}^{0} \phi(x-u)\gamma(u)du$
Soit $x \in \mathbb{R}$, on a $P(\theta > 0|X = x) = \frac{\alpha g_+(x)}{(1-\alpha)\phi(x)+\alpha g(x)}$
donc $2\alpha g_+(t) = (1-\alpha)\phi(t) + \alpha g(t)$
d'où

$$\frac{1}{\alpha} = 1 + \frac{g_+(t) - g_-(t)}{\phi(t)} = 1 + 2\int_0^{+\infty} \sinh(tu)e^{-\frac{u^2}{2}}\gamma(u)du \tag{1.7}$$

On obtient ainsi un seuil fonction de α, continu, et qui décroît de $+\infty$ quand α vaut 0 à 0 quand α vaut 1.
Pour obtenir le seuil classique $t_n = \sqrt{\log n}$, on prendra donc $\alpha = \frac{1}{n}$ (on aurait préféré un seuil en $\sqrt{\log \frac{n}{p_n}}$, mais il aurait pour cela fallu prendre $\alpha = \frac{p_n}{n}$, ce qui n'est pas possible puisque p_n est inconnu).

1.5 Un premier résultat de convergence pour la loi a posteriori

On peut désormais choisir γ comme la densité d'une Laplace standard et $\alpha = \frac{1}{n}$. On obtient ainsi la loi a priori suivante :

$$\theta \sim \Pi := \bigotimes_{i=1}^{n} \frac{n-1}{n}\delta_0 + \frac{1}{n}\gamma \tag{1.8}$$

On peut ainsi énoncer ce premier théorème de convergence de la loi a posteriori, dont on trouvera une preuve en fin de document :

Théorème 2. *Pour tout $\theta_0 \in \ell_0[p_n]$, avec M choisi assez grand et Π comme dans* (1.8) *:*

$$E_{\theta_0}[\Pi(\|\theta - \theta_0\|^2 \geq M p_n \log(n)|X)] \underset{n\to\infty}{\longrightarrow} 0$$

Dans [3], Aad Van der Vaart et Ismaël Castillo utilisent une forme d'a priori qui s'avère plus générale, puisque la loi a priori étudiée dans cette première partie en est un cas particulier. On obtiendra donc le résultat de convergence de l'a posteriori de cette première partie à travers un théorème qui donne un résultat plus général.

Chapitre 2

Une Loi a Priori Plus Générale

2.1 Forme de la loi a priori

On commence par tirer une dimension s suivant une certaine loi $\pi(s)$ sur l'ensemble $\{0, 1, \dots, n\}$
On tire ensuite un support $S \subset \{1, \dots, n\}$ uniformément sur les ensembles de cardinal $s : \Pi(S) = \frac{\pi(s)}{\binom{n}{s}}$
On a alors l'a priori général suivant sur θ :

$$\theta \sim \sum_{S \subset \{1,\dots,n\}} \Pi(S) \bigotimes_{i \in S} \gamma \otimes \bigotimes_{i \notin S} \delta_0$$

avec γ une densité de probabilité.
On peut déjà remarquer que, en général, l'a posteriori n'a plus la forme d'un produit, ce qui avait l'avantage de simplifier les démonstrations dans la première partie.
Dans la suite, on dira que π est à décroissance exponentielle s'il existe $C > 0$ et $D < 1$ tels que $\pi(s) \leq D\pi(s-1)$ pour $s > Cp_n$. Si cette condition est satisfaite pour $C = 0$ alors on dira que π est à décroissance exponentielle stricte.
On peut dès lors se demander ce que l'on pourrait choisir pour le prior π sur la dimension, en voici quelques exemples :

Prior binomial

Si π est la binomiale $B(n, \alpha)$, le prior sur θ est exactement celui de la première partie : on retombe sur l'a priori et l'a posteriori sous forme de produit. Ce prior π est à décroissance exponentielle pour $\alpha \lesssim \frac{p_n}{n}$.$\alpha = \frac{p_n}{n}$ correspond à une valeur oracle, permettant d'atteindre la vitesse minimax. Si on choisit $\alpha = \frac{1}{n}$, on atteint la vitesse minimax si $p_n = n^a$, avec $a < 1$; mais si p_n est d'un ordre plus grand on pourrait rater la vitesse minimax à un facteur logarithmique près.
Une approche bayésienne naturelle serait alors de mettre un a priori sur α :

Prior β

On cherche à obtenir que $s|\alpha$ soit une $B(n,\alpha)$. Pour cela on fait suivre à α une loi β : $\alpha \sim \beta(\kappa,\lambda)$ et on prend

$$\pi(s) = \binom{n}{s}\frac{\beta(\kappa+s,\lambda+n-s)}{\beta(\kappa,\lambda)} \propto \frac{\Gamma(\kappa+s)\Gamma(\lambda+n-s)}{s!(n-s)!}$$

Pour $\kappa = 1$ et $\lambda = n+1$ on a $\pi(s) \propto \binom{2n-s}{n}$, qui a une décroissance exponentielle stricte avec $D = \frac{1}{2}$.
On peut plus généralement prendre $\kappa = 1$ et $\lambda = \kappa_1(n+1)$ et on a dans ce cas $\pi(s) \propto \binom{(\kappa_1+1)n-s}{\kappa_1 n}$, à décroissance exponentielle stricte avec $D = 1/(\kappa_1+1)$. On se réfèrera dorénavant à cette expression pour le prior β.

Complexity prior

Ce prior est de la forme $\pi(s) \propto e^{-as\log(\frac{bn}{s})}$. Il s'avère être bien adapté au problème. En effet, comme $e^{s\log(\frac{n}{s})} \leq \binom{n}{s} \leq e^{s\log(\frac{en}{s})}$, il est inversement proportionnel au nombre de modèles de taille p et paraît donc adapté pour baisser la complexité du problème. Il est à décroissance exponentielle dès que $b > 1+e$, ce qu'on suppose pour le reste de ce document.

2.2 Des résultats de convergence

Dans [2], Ghosal, Ghosh et Van der Vaart démontrent le théorème général suivant :

Théorème 3. *[Théorème général]*
Soit $\mathcal{P}$ une collection de mesures de probabilité. Soit $\Pi = \Pi_n$ une suite de prior et supposons que les observations X soient i.i.d. de densité p_0 : $dP_0(x) = p_0(x)dx$. Soit ε_n une suite de réels positifs telle que $\varepsilon_n \to 0$ et $\sqrt{n}\varepsilon_n \to \infty$ quand $n \to \infty$.
On suppose qu'il existe C et L des constantes telles que :

$$\Pi(p \in \mathcal{P}; -E_{P_0}[\log(\frac{p}{p_0}(X))] \leq \varepsilon_n^2, E_{P_0}[\log(\frac{p}{p_0}(X))^2] \leq \varepsilon_n^2) \geq e^{-Cn\varepsilon_n^2}$$

et

$$\Pi(\mathcal{P} \setminus \mathcal{P}_n) \leq Le^{-(C+4)n\varepsilon_n^2}$$

pour une suite $\mathcal{P}_n \subset \mathcal{P}$ telle qu'on puisse trouver des tests $\psi_n = \psi(X_1,\ldots,X_n)$ tels que $\forall n \in \mathbb{N}$ et $M > 0$ assez grand :

$$E_{P_0}[\psi_n] \to 0 \text{ et } \sup_{d(p,p_0)\geq M\varepsilon_n} E_P[1-\psi_n] \leq Le^{-(C+4)n\varepsilon_n^2}$$

Alors $\Pi(d(p,p_0) > M\varepsilon_n|X) \to 0$ quand $n \to \infty$ en $P_0^{\mathbb{N}}$-probabilité.

Il ne s'applique malheureusement pas dans notre cadre puisque nos observations X_i ne sont pas i.i.d.. Il existe cependant une version non i.i.d. de ce théorème, que l'on ne peut cependant utiliser telle quelle dans notre cas du fait d'une difficulté technique liée au fait que la classe $\ell_0[p_n]$ dans laquelle vit notre paramètre θ_0 n'est pas bornée.
Pour obtenir leurs résultats, Ismaël Castillo et Aad Van der Vaart démontrent au préalable un résultat sur la dimension :

Théorème 4. *[Dimension]*
Si π est à décroissance exponentielle et si γ est centrée avec un moment d'ordre 2 fini, alors il existe $M > 0$ tel que quand $n \to \infty$:

$$\sup_{\theta_0 \in \ell_0[p_n]} E_{\theta_0}[\Pi(|S_\theta| > Mp_n|X)] \to 0$$

On a alors le théorème suivant :

Théorème 5. *Si π est à décroissance exponentielle, si γ est centrée, de moment d'ordre 2 fini et de la forme e^h avec h telle que $\forall x, y \in \mathbb{R}$, $|h(x) - h(y)| \lesssim 1 + |x - y|$, alors, avec r_n tel que*

$$r_n^2 \geq (p_n \log(\frac{n}{p_n})) \vee (\log(\frac{1}{\pi(p_n)}))$$

et $M > 0$ assez grand, quand $n \to \infty$:

$$\sup_{\theta_0 \in \ell_0[p_n]} E_{\theta_0}[\Pi(\|\theta - \theta_0\|^2 > Mr_n^2|X)] \to 0$$

L'hypothèse sur γ est vérifiée si l'on prend une Laplace, et les 3 priors sur la dimension cités en exemple sont conformes aux hypothèses du théorème.
En outre, si on considère le complexity prior, les auteurs ont montré que la moyenne a posteriori atteignait elle aussi la vitesse minimax.

Au cours du stage, on a pu démontrer une version plus simple qu'on appellera ici version semi-implicite de ce théorème, en supposant que γ est une Laplace standard, une preuve se trouvant en fin de document :

Théorème 6. *Pour les prior binomial, β et le complexity prior, avec γ densité d'une Laplace standard :*

$$E_{\theta_0}[\Pi(\|\theta - \theta_0\|^2 \geq Mp_n \log(n)|X)] \underset{n\to\infty}{\longrightarrow} 0$$

avec M choisi assez grand

Chapitre 3

Algorithme et Simulations

3.1 Algorithme

3.1.1 Loi a posteriori

Soit $B = B_1 \times \cdots \times B_n$ un évènement.

On note $Q_n := \sum_{p=0}^{n} \frac{\pi_n(p)}{\binom{n}{p}} \sum_{|S|=p} \left(\prod_{i \notin S} \phi(X_i) \right) \left(\prod_{i \in S} g(X_i) \right)$

On a : $\Pi(B|X) = \frac{1}{Q_n} \sum_S \Pi(S) \left(\prod_{i \notin S} \phi(X_i) \mathbb{1}_{0 \in B_i} \right) \left(\prod_{i \in S} \int_{B_i} \phi(X_i - \theta_i) \gamma(\theta_i) d\theta_i \right)$

On pose $P(Z) = \prod_{i=1}^{n} \big(\phi(X_i) + g(X_i) Z \big)$. Le coefficient en Z^p de ce polynôme vaut :

$$\sum_{|S|=p} \left(\prod_{i \notin S} \phi(X_i) \right) \left(\prod_{i \in S} g(X_i) \right)$$

ce qui permet de calculer Q_n avec une complexité quadratique en n, alors qu'un algorithme "naïf" aurait une complexité exponentielle.

3.1.2 Moyenne a posteriori

La moyenne a posteriori est un vecteur de $\mathbb{R}^n$ qui s'écrit $\hat{\theta}^{moy} = \int \theta d\Pi_n(\theta|X)$

Pour tout $j \in \{1, \ldots, n\}$, on a :

$$\hat{\theta}_j^{moy} = \frac{1}{Q_n} \sum_{p=0}^{n} \frac{\pi_n(p)}{\binom{n}{p}} \zeta(X_j) \sum_{|S|=p, j \in S} \left(\prod_{i \notin S} \phi(X_i) \right) \left(\prod_{i \in S, i \neq j} g(X_i) \right)$$

en notant, $\forall x \in \mathbb{R}$, $\zeta(x) = \int u \phi(x-u) \gamma(u) du$.

On pose $P_j(Z) = \zeta(X_j) \prod_{i=1, i \neq j}^{n} \big(\phi(X_i) + g(X_i) Z \big)$. Le coefficient en Z^p de ce polynôme vaut :

$$\zeta(X_j)\sum_{|S|=p,j\in S}\Big(\prod_{i\notin S}\phi(X_i)\Big)\Big(\prod_{i\in S,i\neq j}g(X_i)\Big)$$

ce qui permet de calculer $\hat{\theta}_j^{moy}$.

3.1.3 Médiane a posteriori

Pour tout $u\in\mathbb{R}$, on a $\Pi(]-\infty;u]\times\mathbb{R}^{n-1}|X)=(1-q_{n,1})\mathbf{1}_{u\geq 0}+q_{n,1}\frac{\psi(X_1,u)}{g(X_1)}$ en notant, $\forall(x,u)\in\mathbb{R}^2$, $\psi(x,u)=\displaystyle\int_{-\infty}^{u}\phi(x-t)g(t)dt$ et $q_{n,1}=1-\Pi(\theta_1=0|X)$.
On a $1-q_{n,1}=\Pi(\theta_1=0|X)=\displaystyle\frac{1}{Q_n}\sum_{p=0}^{n}\frac{\pi_n(p)}{\binom{n}{p}}\sum_{|S|=p,1\notin S}\Big(\prod_{i\notin S}\phi(X_i)\Big)\Big(\prod_{i\in S}g(X_i)\Big)$,
ce que l'on calcule à l'aide du coefficient en Z^p de

$$P_2(Z)=g(X_1)Z\prod_{i=2}^{n}\big(\phi(X_i)+g(X_i)Z\big)$$

On note $H_{n,1}(u)=\frac{\psi(X_1,u)}{g(X_1)}\forall u\in\mathbb{R}$, $H_{n,1}^{-1}(u)$ son inverse avec la convention qu'elle vaut $-\infty$ pour $u\leq 0$ et $+\infty$ pour $u\geq 1$, la première coordonnée de la médiane a posteriori s'écrit alors :

$$\hat{\theta}_1^{med}=\big(H_{n,1}^{-1}(1-\frac{1}{2q_{n,1}})\vee 0\big)+\big(H_{n,1}^{-1}(\frac{1}{2q_{n,1}})\wedge 0\big)$$

et de même pour les autres coordonnées.

3.1.4 Risque quadratique

On note $\hat{R}=\int\|\theta-\theta^o\|^2d\Pi(\theta|X)$.
On a $\hat{R}=\displaystyle\sum_{i=1}^{n}\int(\theta_i-\theta_i^o)^2d\Pi(\theta|X)$,
soit $\hat{R}=\displaystyle\sum_{i=1}^{n}\Big(\int\theta_i^2d\Pi(\theta|X)-2\theta_i^o\hat{\theta}_i^{moy}+\theta_i^{o2}\Big)$
On calcule alors les coordonnées du moment d'ordre 2 a posteriori de la même façon qu'on a calculé celles de la moyenne a posteriori, mais en remplaçant simplement les $\zeta(X_j)$ par des $\zeta_2(X_j)=\int u^2\phi(X_j-u)\gamma(u)du$. Leurs expressions sont rassemblées dans la propriété ci-dessous :

Propriété 2. *Pour tout $x\in\mathbb{R}$, on a*

$$\zeta(x)=\frac{1}{2}\Big[e^{(\frac{1}{2}-x)}\big((x-1)\int_{1-x}^{+\infty}\phi(u)du+\phi(x-1)\big)+e^{(\frac{1}{2}+x)}\big((x+1)\int_{1+x}^{+\infty}\phi(u)du-\phi(x+1)\big)\Big]$$

et $\zeta_2(x) = \dfrac{1}{2}\Big[e^{(\frac{1}{2}-x)}\big((x^2+2-2x)\int_{1-x}^{+\infty}\phi(u)du + (x-1)\phi(x-1)\big) + e^{(\frac{1}{2}+x)}\big((x^2+2+2x)\int_{1+x}^{+\infty}\phi(u)du - (x+1)\phi(x+1)\big)\Big]$

3.2 Simulations

On a effectué des simulations avec un θ_0 nul sur les $n - p_n$ premières coordonnées et valant une constante A sur les p_n dernières. On a comparé, avec $n = 500$ et p_n et A variables, risque de la médiane, de la moyenne et risque quadratique pour un prior 1 complexity avec $a = 0,1$ et $b = 4$ et un prior 2 β avec $\kappa_1 = 0,1$:

pn		25			50			100	
A	3	4	5	3	4	5	3	4	5
Médiane 1	166	97	73	206	137	133	250	266	280
Médiane 2	109	78	74	158	141	145	252	272	286
Moyenne 1	131	91	81	174	156	156	266	290	305
Moyenne 2	106	93	95	164	169	175	270	294	310
Risque 1	188	180	176	312	333	334	541	590	610
Risque 2	230	233	239	355	384	389	559	600	619

On voit que dans tous les cas, notre "risque quadratique" donne des performances beaucoup moins bonnes qu'à la fois médiane et moyenne, et même pire : qu'il dépasse la valeur de n pour $p_n = 100$... Cet aspect précis de la loi a posteriori semble donc ne pas aussi bien se comporter que la médiane et la moyenne a posteriori.

Chapitre 4

Conclusion et Perspectives

On a vu au long de ce document que la méthode bayésienne consistant à prendre des lois a priori qui tiennent en compte de la parcimonie du modèle fournit des estimateurs efficaces, en particulier la médiane a posteriori qui est de type seuillage, mais également d'autres aspects de la loi a posteriori comme la moyenne ou encore la loi a posteriori elle-même qui a de bons résultats de convergence.
Les simulations mettent en avant le fait que le risque quadratique a posteriori n'obtient pas d'aussi bons résultats que la médiane ou même la moyenne, voire même des résultats pas vraiment satisfaisants. Une des perspectives maintenant serait d'étudier théoriquement pourquoi ces résultats n'étaient pas aussi bons qu'on aurait pu espérer.
On pourra voir que les aspects de la loi a posteriori n'ont pas tous le même comportement asymptotique et qu'ils n'ont pas forcément tous de bonnes propriétés de convergence. On peut en effet dresser un parallèle entre modes a posteriori, quand ils sont bien définis, et estimateurs pénalisés en sélection de modèles comme dans [7]. A l'avenir, on pourra également se pencher sur le cas où les ε_i ne suivent plus une loi normale standard, mais une loi normale centrée de variance σ^2 inconnue.
On pourra aussi regarder un modèle de la forme $X = M\theta + \varepsilon$, où M est une matrice de design.
Dans [1], on utilise une approche par le bayésien empirique avec le maximum de vraisemblance pour obtenir des résultats de convergence sur la médiane et la moyenne a posteriori. On pourra s'intéresser aux mêmes types de résultats mais pour l'a posteriori tout entier, on pourra pour cela étudier [6] et voir si leurs résultats s'appliquent dans notre cas précis.

Chapitre 5

Preuves

5.1 Deux lemmes de concentration

Lemme 1. *En notant $\mathcal{N}(0,1)$ la loi normale standard, on a, $\forall t \in \mathbb{R}_+^*$:*

$$P(|\mathcal{N}(0,1)| > t) \leq e^{-\frac{t^2}{2}}$$

Lemme 2. *On note $\mathcal{A}_n = \{\max |\varepsilon_i| \leq \sqrt{4 \log n}\}$. On a $P(\mathcal{A}_n^c) = O(\frac{1}{n})$quand $n \to \infty$.*

5.2 Preuve du résultat de vitesse minimax

Théorème 7. *Vitesse minimax On note $R_{n,2}(\ell_0[p_n]) = \inf_{\hat{\theta}} \sup_{\theta \in \ell_0[p_n]} \frac{1}{n} \sum_{i=1}^{n} E_\theta(\hat{\theta}_i - \theta_i)^2$. On a alors $R_{n,2}(\ell_0[p_n]) = \frac{2p_n}{n} \log(\frac{n}{p_n})(1 + o(1))$ quand $n \to \infty$*

Démonstration

I) Majoration de $R_{n,2}(\ell_0[p_n])$

Soit $\theta \in \ell_0[p_n]$
Considérons l'estimateur oracle

$$\hat{\theta}_n = (X_i 1\!\!1_{|X_i| > t_n})_{i \in \{1,\ldots,n\}}, \quad t_n = \sqrt{2 \log \frac{n}{p_n}},$$

On s'intéresse à

$$R(\theta) = E_\theta \|\hat{\theta} - \theta\|^2 = \sum_{i=1}^{n} E_\theta(\hat{\theta}_i - \theta_i)^2.$$

Commençons par étudier

$$\begin{aligned}
E_\theta(\hat\theta_1-\theta_1)^2 &= E_\theta\left[(\hat\theta_1-\theta_1)^2 1\!\!1_{\hat\theta_1=0}\right]+E_\theta\left[(\hat\theta_1-\theta_1)^2 1\!\!1_{\hat\theta_1\neq 0}\right]\\
&= \theta_1^2 E_\theta\left[1\!\!1_{|X_1|\le t_n}\right]+E_\theta\left[(X_1-\theta_1)^2 1\!\!1_{|X_1|>t_n}\right]\\
&= \theta_1^2 P\left[|\theta_1+\varepsilon_1|\le t_n\right]+E\left[\varepsilon_1^2 1\!\!1_{|\theta_1+\varepsilon_1|>t_n}\right]\\
&= \qquad r(\theta_1) \qquad + \qquad q(\theta_1).
\end{aligned}$$

Distinguons trois cas :

a) $\theta_1=0$

b) θ_1 est 'petit', soit $|\theta_1|\le\sqrt{2\log\frac{n}{p_n}}+C_n$,

c) θ_1 est 'grand', soit $|\theta_1|>\sqrt{2\log\frac{n}{p_n}}+C_n$, avec $C_n=\sqrt{2\log(\log(n))}$ à choisir.

a) Cas $\theta_1=0$

$r(\theta_1)=0$ et $q(\theta_1)=\int\limits_{|x|>t_n} x^2e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}=\frac{2t_n}{\sqrt{2\pi}}e^{-\frac{t_n^2}{2}}+\int\limits_{|x|>t_n} e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}$ après intégration par parties.

Or $t_n=\sqrt{\log\frac{n}{p_n}}$ d'où $e^{-\frac{t_n^2}{2}}=\frac{p_n}{n}$

d'où $q(\theta_1)=2\frac{p_n}{n}\sqrt{\frac{2\log\frac{n}{p_n}}{2\pi}}+P(|\mathcal{N}(0,1)|>t_n)$

et on a donc dans ce cas $r(\theta_1)+q(\theta_1)\le\frac{p_n}{n}\sqrt{2\log\frac{n}{p_n}}\left(\frac{2}{\sqrt{2\pi}}+\frac{1}{\sqrt{2\log\frac{n}{p_n}}}\right)$

b) Cas $|\theta_1|\le\sqrt{2\log\frac{n}{p_n}}+C_n$

On a $q(\theta_1)=E\left[\varepsilon_1^2 1\!\!1_{|\theta_1+\varepsilon_1|>t_n}\right]\le 1$

et $r(\theta_1)=\theta_1^2P\left[|\theta_1+\varepsilon_1|\le t_n\right]\le 2\log\frac{n}{p_n}+C_n^2+2C_n\sqrt{\log(\frac{n}{p_n})}$

d'où $r(\theta_1)+q(\theta_1)\le\log\frac{n}{p_n}\left(2+\frac{1+C_n^2+2C_n\sqrt{\log(\frac{n}{p_n})}}{\log\frac{n}{p_n}}\right)$ dans ce cas

c) Cas $|\theta_1|>\sqrt{2\log\frac{n}{p_n}}+C_n$

On a $q(\theta_1)\le 1$ et $r(\theta_1)\le\theta_1^2P[|\varepsilon_1|\ge|\theta_1|-t_n]$ par inégalité triangulaire.
On a donc $r(\theta_1)\le\theta_1^2e^{-\frac{(|\theta_1|-t_n)^2}{2}}$
Distinguons maintenant selon que $|\theta_1|\ge\sqrt{8\log\frac{n}{p_n}}$ ou non.

Si c'est le cas, on a $|\theta_1| - t_n = \frac{|\theta_1|}{2} + \frac{|\theta_1|}{2} - t_n \geq \frac{|\theta_1|}{2} + (\frac{\sqrt{8}}{2} - \sqrt{2}) \log \frac{n}{p_n} \geq \frac{|\theta_1|}{2}$
d'où $r(\theta_1) \leq \theta_1^2 e^{-\frac{\theta_1^2}{2}} \leq 1$
Sinon, $r(\theta_1) \leq 8 \log(\frac{n}{p_n}) e^{-\frac{D_n^2}{2}} \leq 8 \frac{\log(\frac{n}{p_n})}{\log(n)}$ d'où $r(\theta_1) + q(\theta_1) \leq 2 + 8 \frac{\log(\frac{n}{p_n})}{\log(n)}$

Conclusion

On sait qu'il y a au plus p_n composantes de θ non nulles, on obtient donc :

$$R(\theta) \leq n\sqrt{2\log\frac{n}{p_n}}\frac{p_n}{n}(\frac{2}{\sqrt{2\pi}}+\frac{1}{\sqrt{2\log\frac{n}{p_n}}})+p_n\log\frac{n}{p_n}(2+\frac{1+C_n^2+2C_n\sqrt{\log(\frac{n}{p_n})}}{\log\frac{n}{p_n}})+p_n(2+8\frac{\log(\frac{n}{p_n})}{\log(n)})$$

On obtient donc $R(\theta) \leq 2p_n \log \frac{n}{p_n}(1 + o(1))$
D'où $R_{n,2}(\ell_0[p_n]) \leq 2\frac{p_n}{n} \log \frac{n}{p_n}(1 + o(1))$

II) Minoration de $R_{n,2}(\ell_0[p_n])$

Une première version de la preuve est la suivante :
Soit Π un prior sur $\mathbb{R}^n$.

$$\inf_T \int_\Theta E_\theta[\|T(X) - \theta\|^2] d\Pi(\theta) \leq \inf_T \sup_\Theta E_\theta[\|T(X) - \theta\|^2] \int_\Theta d\Pi(\theta)$$

d'où $R := nR_{n,2}(\ell_0[p_n]) \geq \inf_T \int_\Theta E_\theta[\|T(X) - \theta\|^2] d\Pi(\theta)$.
On choisit alors le prior suivant,
en notant $\alpha_n = \frac{p_n - \sqrt{4\log(n)p_n}}{n}$ et $M_n = \sqrt{2\log\frac{1}{\alpha_n}} - C_n$ avec C_n une suite qui tend vers l'infini plus lentement que $\sqrt{2\log\frac{1}{\alpha_n}}$:

$$\Pi \sim \bigotimes_{i=1}^n (1 - \alpha_n)\delta_0 + \alpha_n \delta_{M_n}$$

C'est un prior sur $\Theta_1 = \{0, M_n\}^n$. On peut restreindre l'infimum sur T à la classe S des estimateurs à valeurs dans $[-2M_n; 2M_n]$.
En effet, si T est un estimateur et $\theta \in \Theta_1$: $\|T(X) - \theta\|^2 \geq \|\tilde{T}(X) - \theta\|^2$, avec $\forall i$, $\tilde{T}_i = T_i$ si $|T_i| \leq 2M_n$ et $2M_n sign(T_i)$ sinon.
On a $R \geq \inf_{T \in S}(\int_{\Theta_1} E_\theta[\|T(X) - \theta\|^2] d\Pi(\theta) - \int_{\Theta_1 \setminus \Theta} E_\theta[\|T(X) - \theta\|^2] d\Pi(\theta))$
Donc $R \geq \inf_T(\int_{\Theta_1} E_\theta[\|T(X) - \theta\|^2] d\Pi(\theta)) - \sup_{T \in S}(\int_{\Theta_1 \setminus \Theta} E_\theta[\|T(X) - \theta\|^2] d\Pi(\theta)) = I - II$

Pour le second terme : $II \leq \sup_{T\in S}(2\int_{\Theta_1\setminus\Theta} E_\theta[\|T(X)\|^2]d\Pi(\theta)+2\int_{\Theta_1\setminus\Theta} E_\theta[\|\theta\|^2]d\Pi(\theta))$
donc $II \leq 10nM_n^2\Pi(\Theta_1 \setminus \Theta)$
On note $V_n = n\alpha_n(1-\alpha_n)$.
Pour tout $x>0$, par Bernstein, on a $\Pi(||S_\theta| - n\alpha_n| \geq x) \leq 2e^{\frac{-x^2}{2(V_n+\frac{x}{3})}}$.
donc $II = 10nM_n^2\Pi(|S_\theta| > p_n) \leq 10nM_n^2\Pi(||S_\theta| - n\alpha_n| \geq \sqrt{4\log(n)p_n}) \leq 20nM_n^2 e^{-\frac{p_n \log n}{V_n+\frac{1}{3}\sqrt{4p_n\log n}}}$
(c'est ici que l'on se sert du fait que α_n est légèrement inférieur à p_n/n)

Or $V_n + \frac{1}{3}\sqrt{4p_n\log n} \leq p_n - \sqrt{4p_n \log n} + \frac{1}{3}\sqrt{4p_n\log n}$

donc $V_n + \frac{1}{3}\sqrt{4p_n\log n} \leq p_n - \frac{2}{3}\sqrt{4p_n\log n} \leq p_n$

d'où $II \leq 20nM_n^2 e^{-2\log n} = \frac{20M_n^2}{n}$

Pour le premier terme :
On a, en notant $\bar\theta$ la moyenne a posteriori :

$$\inf_T \int E_\theta[\|T-\theta\|^2]d\Pi(\theta) = \int E_\theta[\|\bar\theta - \theta\|^2]d\Pi(\theta) = \sum_{i=1}^n \int E_\theta[(\bar\theta_i - \theta_i)^2]d\Pi(\theta_i)$$

Soit $i \in \{1,\ldots,n\}$, $A_i := \int E_\theta[(\bar\theta_i - \theta_i)^2]d\Pi(\theta_i) \geq \Pi(\theta_i = M_n)E_{M_n}[(\bar\theta_i - M_n)^2]$

Donc $A_i \geq \alpha_n E_{M_n}[(\bar\theta_i - M_n)^2 1\!\!1_{|\varepsilon_i|<K_n}]$, avec K_n une suite qui tend vers l'infini plus telle que $4C_n \geq K_n$

De plus, un rapide calcul montre que $\bar\theta_i = \frac{\alpha_n M_n}{\alpha_n + (1-\alpha_n)e^{\frac{M_n^2}{2} - X_i M_n}}$
Donc $\bar\theta_i = \frac{M_n}{1+(1-\alpha_n)e^{\frac{M_n^2}{2} - X_iM_n + \log(\frac{1}{\alpha_n})}}$
Sous P_{M_n}, on a $X_i = M_n + \varepsilon_i$, d'où $\bar\theta_i = \frac{M_n}{1+(1-\alpha_n)e^{-\frac{M_n^2}{2} - \varepsilon_i M_n + \log(\frac{1}{\alpha_n})}}$
D'où, sous P_{M_n}, $\bar\theta_i - M_n = -M_n \frac{1}{1+\frac{\alpha_n}{1-\alpha_n}e^{\frac{M_n^2}{2}+\varepsilon_i M_n}}$
or, sur l'évènement $\{|\varepsilon_i| \leq K_n\}$,

$$\frac{M_n^2}{2} + \varepsilon_i M_n \leq \log\frac{1}{\alpha_n} - C_n\sqrt{2\log\frac{1}{\alpha n}} + \frac{C_n^2}{2} + K_n(\sqrt{2\log\frac{1}{\alpha n}} - C_n)$$

donc $\frac{\alpha_n}{1-\alpha_n}e^{\frac{M_n^2}{2}+\varepsilon_i M_n} \leq \frac{1}{1-\alpha_n}e^{-C_n\sqrt{2\log\frac{1}{\alpha n}} + \frac{C_n}{2}\sqrt{2\log\frac{1}{\alpha n}} + \frac{C_n}{4}(\sqrt{2\log\frac{1}{\alpha n}} - C_n)}$

d'où, $\frac{\alpha_n}{1-\alpha_n}e^{\frac{M_n^2}{2}+\varepsilon_i M_n} \leq \frac{1}{1-\alpha_n}e^{-\frac{C_n}{4}\sqrt{2\log\frac{1}{\alpha n}} - \frac{C_n^2}{4}}$ quantité qui tend vers 0 quand $n \to \infty$.

Donc $A_i \geq \alpha_n M_n^2(1+o(1))P(|\varepsilon_i| < K_n)$

or $P(|\varepsilon_i| < K_n) \to 1$ quand $n \to \infty$, et finalement $I \geq n\alpha_n M_n^2(1+o(1))$

D'où $R \geq 2p_n \log(\frac{n}{p_n})(1+o(1))$, et finalement on a bien $R_{n,2} \geq 2\frac{p_n}{n}\log(\frac{n}{p_n})(1+o(1))$, ce qui achève la démonstration.

Remarque
Une version alternative à la preuve de la minoration pourrait utiliser le lemme de Fano suivant dont une preuve est fournie par Bin Yu en [5]

Lemme 3. *Soit $\mathcal{P}$ une famille de mesures de probabilités de paramètre d'intérêt $\theta(P)$, à valeurs dans un espace pseudo-métrique $(\mathcal{D}, d)$. Soit n un entier naturel supérieur ou égal à 2, et $\mathcal{M}_n = \{P_{\theta_1}, \ldots, P_{\theta_n}\} \subset \mathcal{P}$ tel qu'on ait $\forall j \neq k : d(\theta_j, \theta_k) \geq m_n$ et $K(P_{\theta_j}, P_{\theta_k}) \leq M_n$*
Alors $\max_{j \in \{1,\ldots,n\}} E_{\theta_j}[d(\hat{\theta}, \theta_j)] \geq \dfrac{m_n}{2}(1 - \dfrac{M_n + \log 2}{\log n})$ pour tout estimateur $\hat{\theta}$.

5.3 Preuve du premier théorème de convergence de la loi a posteriori

Le résultat

Pour Π comme dans (1.8),

$$E_{\theta_0}[\Pi(\|\theta - \theta_0\|^2 \geq M p_n \log(n)|X)] \underset{n\to\infty}{\longrightarrow} 0 \text{ avec } M \text{ choisi assez grand}$$

Démonstration

Dans tout ce qui suit, on travaillera sur l'évènement $\mathcal{A} = \left\{\max_{i \in \{1,\ldots,n\}} |\varepsilon_i| \leq \sqrt{4\log(n)}\right\}$, étant donné que $P(\mathcal{A}^c) = O\left(\frac{1}{n}\right)$

On note $S = \{i \in \{1,\ldots,n\}; \theta_i \neq 0\}$ et $S_0 = \{i \in \{1,\ldots,n\}; \theta_{0,i} \neq 0\}$
On commence par décomposer selon qu'il y a ou non du signal :

$$\|\theta - \theta_0\|^2 = \|\theta_{S_0} - \theta_{0,S_0}\|^2 + \|\theta_{S_0^c}\|^2$$

et donc :

$$\Pi\left(\|\theta - \theta_0\|^2 \geq M p_n \log(n)|X\right) \leq \underbrace{\Pi\left(\|\theta_{S_0} - \theta_{0,S_0}\|^2 \geq \frac{M}{2} p_n \log(n)|X\right)}_{I} + \underbrace{\Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2} p_n \log(n)|X\right)}_{II}$$

I : Présence de signal

On va distinguer si le signal est "grand" ou "petit". Pour cela on considère l'ensemble $S_{0,G} = \{i \in \{1,\dots,n\}; \theta_{0,i} > \sqrt{C_1 \log(n)}\}$, avec $C_1 > 0$ une constante à choisir.

$$I \leq \underbrace{\Pi\left(\|\theta_{S_{0,G}} - \theta_{0,S_{0,G}}\|^2 \geq \frac{M}{4} p_n \log(n) | X\right)}_{i} + \underbrace{\Pi\left(\|\theta_{S_{0,G}^c} - \theta_{0,S_{0,G}^c}\|^2 \geq \frac{M}{4} p_n \log(n) | X\right)}_{ii}$$

$$\begin{aligned}
i &= \Pi\left(\sum_{i \in S_{0,G}} (\theta_i - \theta_{0,i})^2 \geq \tfrac{M}{4} p_n \log(n) | X\right) \\
&\leq \sum_{i \in S_{0,G}} \Pi\left((\theta_i - \theta_{0,i})^2 \geq \tfrac{M}{4|S_{0,G}|} p_n \log(n) | X\right) \\
&\leq \sum_{i \in S_{0,G}} \Pi\left(|\theta_i - \theta_{0,i}| \geq \sqrt{\tfrac{M}{4} \log(n)} | X\right) \\
&\leq \sum_{i \in S_{0,G}} \Pi\left(|\theta_i - X_i| \geq \sqrt{\tfrac{M}{4} \log(n)} - |\varepsilon_i| | X\right) \\
&\leq \sum_{i \in S_{0,G}} \Pi\left(|\theta_i - X_i| \geq \underbrace{\left(\sqrt{\frac{M}{4}} - 2\right)}_{=L} \sqrt{\log(n)} | X\right)
\end{aligned}$$

Lemme 4. *On choisit $M = 44^2 = 1936$ et $C_1 = 36$. En probabilité :*

$$\Pi\left(|\theta_i - X_i| \geq L\sqrt{\log(n)} | X\right) = O\left(\frac{1}{n}\right)$$

et par conséquent $i = O\left(\frac{p_n}{n}\right)$

$$\begin{aligned}
ii &= \Pi\left(\sum_{i \in S_{0,G}^c} (\theta_i - \theta_{0,i})^2 \geq \tfrac{M}{4} p_n \log(n) | X\right) \\
&\leq \sum_{i \in S_{0,G}^c} \Pi\left((\theta_i - \theta_{0,i})^2 \geq \tfrac{M}{4|S_{0,G}^c|} p_n \log(n) | X\right) \\
&\leq \sum_{i \in S_{0,G}^c} \Pi\left(|\theta_i - \theta_{0,i}| \geq \sqrt{\tfrac{M}{4} \log(n)} | X\right) \\
&\leq \sum_{i \in S_{0,G}^c} \Pi\left(|\theta_i| \geq \sqrt{\tfrac{M}{4} \log(n)} - |\theta_{0,i}| | X\right) \\
&\leq \sum_{i \in S_{0,G}^c} \Pi\left(|\theta_i| \geq \underbrace{\left(\sqrt{\frac{M}{4}} - \sqrt{C_1}\right)}_{=l} \sqrt{\log(n)} | X\right)
\end{aligned}$$

Lemme 5. *En probabilité :*

$$\Pi\left(|\theta_i| \geq l\sqrt{\log(n)}|X\right) = O\left(\frac{1}{n}\right)$$

et par conséquent $ii = O\left(\frac{p_n}{n}\right)$

Nos deux termes sont des $O\left(\frac{p_n}{n}\right)$ et tendent donc bien vers 0 quand $n \to \infty$.

II : Pas de signal

On traite ensuite le second terme :

$$II = \Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n \log(n)|X\right)$$

Pour tout $c > 0$, on a :

$$II \leq \underbrace{\Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n \log(n), |S \cap S_0^c| \leq cp_n|X\right)}_{i} + \underbrace{\Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n \log(n), |S \cap S_0^c| > cp_n|X\right)}_{ii}$$

Lemme 6. *En probabilité :* $ii = o_{P_{\theta_0}}(1)$

Pour le premier terme, on a : $i = \sum_{|S^*|\leq cp_n} \Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n \log(n), S \cap S_0^c = S^*|X\right)$
soit $i = \sum_{|S^*|\leq cp_n} \Pi\left(S \cap S_0^c = S^*|X\right) \Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n \log(n)|X, S \cap S_0^c = S^*\right)$
donc $i \leq \max_{|S^*|\leq cp_n} \left(\Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n \log(n)|X, S \cap S_0^c = S^*\right)\right) \sum_{|S^*|\leq cp_n} \Pi\left(S \cap S_0^c = S^*|X\right)$
d'où $i \leq \max_{|S^*|\leq cp_n} \left(\Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n \log(n)|X, S \cap S_0^c = S^*\right)\right)$

De plus,$\forall S^*$ tel que $|S^*| \leq cp_n$, on a :

$$\Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n \log(n)|X, S \cap S_0^c = S^*\right) \leq \sum_{i\in S^*} \Pi\left(\theta_i^2 \geq \frac{M}{2|S^*|}p_n \log(n)|X, S \cap S_0^c = S^*\right)$$

donc $\Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n \log(n)|X, S \cap S_0^c = S^*\right) \leq \sum_{i\in S^*} \Pi\left(|\theta_i| \geq \sqrt{\frac{M}{2c}\log(n)}|X, S \cap S_0^c = S^*\right)$

Lemme 7. $\Pi\left(\theta|X, S = S^*\right) = \prod_{i\notin S^*} \delta_0(\theta_i) \prod_{i\in S^*} \psi_{X_i}(\theta_i)$

Ainsi, $\forall i \in S^*$:

$$\Pi\left(|\theta_i| \geq \sqrt{\frac{M}{2c}\log(n)}|X, S \cap S_0^c = S^*\right) = \int_{|u|\geq\sqrt{\frac{M}{2c}\log(n)}} \psi_{X_i}(u)du$$

Or :

$$\begin{aligned}
\int_{|u|\geq\sqrt{\frac{M}{2c}\log(n)}} \psi_{X_i}(u)du &= \int_{|u|\geq\sqrt{\frac{M}{2c}\log(n)}} \frac{\phi(X_i-u)\gamma(u)}{g(X_i)}du \\
&\lesssim \int_{|u|\geq\sqrt{\frac{M}{2c}\log(n)}} \exp\left(-\frac{(X_i-u)^2}{2} - |u| + |X_i|\right) du
\end{aligned}$$

Or $|X_i| = |\varepsilon_i| \leq 2\sqrt{\log(n)}$ et en prenant $c = 2$, on a $\sqrt{\frac{M}{2c}} = 22$ et donc $|X_i| \leq \frac{|u|}{2}$. Ainsi $|X_i - u| \geq |u| - |X_i| \geq \frac{|u|}{2}$. D'où :

$$\begin{aligned}
\int_{|u|\geq\sqrt{\frac{M}{2c}\log(n)}} \psi_{X_i}(u)du &\lesssim \int_{|u|\geq 22\sqrt{\log(n)}} \exp\left(-\frac{u^2}{2} - \frac{|u|}{2}\right) du \\
&\lesssim \int_{|u|\geq 22\sqrt{\log(n)}} \exp\left(-\frac{u^2}{2}\right) du \\
&\lesssim P(|N(0,1)| \geq 22\sqrt{\log(n)}) \\
&\lesssim e^{-\frac{(22\sqrt{\log(n)})^2}{2}} \\
&\lesssim \frac{1}{n^{242}}
\end{aligned}$$

Finalement, $i = O\left(\frac{p_n}{n}\right)$ et on a donc $E_{\theta_0}[II] \underset{n\to\infty}{\longrightarrow} 0$

D'où $E_{\theta_0}[\Pi(\|\theta - \theta_0\|^2 \geq Mp_n \log(n)|X)] \underset{n\to\infty}{\longrightarrow} 0$ **CQFD**

Preuve des Lemmes

On prend $M = 1936$ et $C_1 = 36$

Preuve du Lemme 4

Enoncé :
Si $\theta_{0,i} > \sqrt{C_1 \log(n)}$,

$$\Pi\left(|\theta_i - X_i| \geq L\sqrt{\log(n)}|X\right) = O\left(\frac{1}{n^{\frac{L^2}{8}}}\right)$$

Démonstration.

$$\Pi\left(|\theta_i - X_i| \geq L\sqrt{\log(n)}|X\right) = (1-\tilde{\alpha}) + \tilde{\alpha}\int_{|u-X_i|\geq L\sqrt{\log(n)}} \psi_{X_i}(u)du$$

D'une part :

$$\begin{aligned} 1-\tilde{\alpha} &= \tfrac{(1-\alpha)\phi(X_i)}{(1-\alpha)\phi(X_i)+\alpha g(X_i)} \\ &= \tfrac{n-1}{n-1+\frac{g}{\phi}(X_i)} \\ &\leq \tfrac{n-1}{\frac{g}{\phi}(X_i)} \end{aligned}$$

On a $|X_i| = |\theta_{0,i} + \varepsilon_i| \geq (\sqrt{C_1} - 2)\sqrt{\log(n)} = 4\sqrt{\log(n)}$ et $\frac{g}{\phi}$ est une fonction croissante, donc $1 - \tilde{\alpha} \leq \frac{n-1}{\frac{g}{\phi}(4\sqrt{\log(n)})}$ De plus, $\forall x > 0$, on a :

$$\begin{aligned} \tfrac{g}{\phi}(x) &= \frac{1}{2}\int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}+ux-u}du \\ &\geq \tfrac{1}{2}\int_{x-1}^{x} e^{-\frac{u^2}{2}+ux-u}du \\ &\geq \tfrac{1}{2}e^{-\frac{x^2}{2}+x^2-x} \text{ (car } u \mapsto e^{-\frac{u^2}{2}+ux-u} \text{ est décroissante sur } [x-1;x]) \\ &\geq \tfrac{1}{2}e^{\frac{x^2}{2}-x} \end{aligned}$$

Ainsi $1 - \tilde{\alpha} \leq \frac{2(n-1)}{e^{\frac{(4\sqrt{\log(n)})^2}{2}-4\sqrt{\log(n)}}} = \frac{2(n-1)e^{4\sqrt{\log(n)}}}{n^8} \leq \frac{2(n-1)n^4}{n^8}$
D'où $1 - \tilde{\alpha} = O\left(\frac{1}{n}\right)$

D'autre part :

$$\begin{aligned} \tilde{\alpha}\textstyle\int_{|u-X_i|\geq L\sqrt{\log(n)}} \psi_{X_i}(u)du &\lesssim \int_{|u-X_i|\geq L\sqrt{\log(n)}} \frac{e^{-\frac{(X_i-u)^2}{2}-|u|}}{g(X_i)}du \\ &\lesssim \int_{|u-X_i|\geq L\sqrt{\log(n)}} e^{-\frac{(X_i-u)^2}{2}-|u|+|X_i|}du \\ &\lesssim \int_{|u-X_i|\geq L\sqrt{\log(n)}} e^{-\frac{(X_i-u)^2}{2}-|X_i-u|}du \\ &= e^{\frac{1}{2}}\textstyle\int_{|u-X_i|\geq L\sqrt{\log(n)}} e^{-\frac{(X_i-u)^2}{2}-|X_i-u|-\frac{1}{2}}du \\ &\lesssim \textstyle\int_{|v|\geq L\sqrt{\log(n)}} e^{-\frac{(v-1)^2}{2}}dv \end{aligned}$$

On a $L = \sqrt{\frac{M}{4}} - 2 = 20$ et $|v| - 1 \geq \frac{|v|}{2}$, d'où :

$$\begin{aligned}
\tilde{\alpha} \int_{|u-X_i| \geq L\sqrt{\log(n)}} \psi_{X_i}(u)du &\lesssim \int_{|v| \geq L\sqrt{\log(n)}} e^{-\frac{v^8}{2}} dv \\
&\lesssim P(|N(0,1)| \geq \frac{L}{2}\sqrt{\log(n)}) \\
&\lesssim \frac{1}{n^{\frac{L^2}{8}}} = \frac{1}{n^{50}}
\end{aligned}$$

□

Preuve du Lemme 5

Enoncé :
Si $\theta_{0,i} \leq \sqrt{C_1 \log(n)}$,

$$\Pi\left(|\theta_i| \geq l\sqrt{\log(n)}|X\right) = O\left(\frac{1}{n^{\frac{l^2}{8}}}\right)$$

Démonstration.

$$\begin{aligned}
\Pi\left(|\theta_i| \geq l\sqrt{\log(n)}|X\right) &= \tilde{\alpha} \int_{|u| \geq l\sqrt{\log(n)}} \psi_{X_i}(u)du \\
&\leq \int_{|u| \geq l\sqrt{\log(n)}} \psi_{X_i}(u)du
\end{aligned}$$

$$\begin{aligned}
\int_{|u| \geq l\sqrt{\log(n)}} \psi_{X_i}(u)du &= \int_{|u| \geq l\sqrt{\log(n)}} \frac{\phi(X_i - u)\gamma(u)}{g(X_i)} du \\
&\lesssim \int_{|u| \geq l\sqrt{\log(n)}} \exp\left(-\frac{(X_i - u)^2}{2} - |u| + |X_i|\right) du
\end{aligned}$$

Or $|X_i| = |\theta_{0,i} + \varepsilon_i| \leq (\sqrt{C_1} + 2)\sqrt{\log(n)}$ et $l = \left(\sqrt{\frac{M}{4}} - \sqrt{C_1}\right)$. Comme $M = 1936$ et $C_1 = 6$, on a $\sqrt{C_1} + 2 = 8$ et $l = 16$ donc $|X_i| \leq \frac{|u|}{2}$. Ainsi $|X_i - u| \geq |u| - |X_i| \geq \frac{|u|}{2}$. D'où :

$$\begin{aligned}
\int_{|u| \geq l\sqrt{\log(n)}} \psi_{X_i}(u)du &\lesssim \int_{|u| \geq l\sqrt{\log(n)}} \exp\left(-\frac{u^2}{4} - \frac{|u|}{2}\right) du \\
&\lesssim \int_{|u| \geq l\sqrt{\log(n)}} \exp\left(-\frac{u^2}{2}\right) du \\
&\lesssim P(|N(0,1)| \geq l\sqrt{\frac{\log(n)}{2}}) \\
&\lesssim e^{-\frac{(l\sqrt{\log(n)})^2}{4}} \\
&\lesssim \frac{1}{n^{\frac{l^2}{4}}} = \frac{1}{n^{64}}
\end{aligned}$$

□

Preuve du Lemme 6

Enoncé :
$\Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n\log(n), |S\cap S_0^c| > cp_n|X\right) = o_{P_{\theta_0}}(1)$

Démonstration. On note B_n l'évènement $\{\theta \in \mathbb{R}^n; |S\cap S_0^c| > cp_n\}$
On peut choisir comme on veut les coordonnées de θ qui sont dans le support S_0 de θ_0 et on a ainsi : $B_n = \mathbb{R}^{S_0} \times B'_n$

$$\Pi\left(\|\theta_{S_0^c}\|^2 \geq \frac{M}{2}p_n\log(n), |S\cap S_0^c| > cp_n|X\right) \leq \Pi\left(B_n|X\right)$$

En notant $\theta_2 = \{\theta_i, i\in S_{\theta_0}^c\}$,$X_2 = \{X_i, i \in S_{\theta_0}^c\}$ et $n_2 = n - p_n$, on a :

$$\Pi(B_n|X) = \frac{\int_{B'_n} p_{n_2,\theta_2(X_2)d\Pi(\theta_2)}}{\int p_{n_2,\theta_2(X_2)d\Pi(\theta_2)}} =: \frac{N}{D}$$

On a $D \geq \displaystyle\int_{\{\theta_2=0\}} p_{n_2,\theta_2}(X_2)d\Pi(\theta_2) = (1-\alpha)p_{n_2,0}(X_2) = \frac{n-1}{n}p_{n_2,0}(X_2)$

D'où $\Pi(B_n|X) \leq \displaystyle\frac{n}{n-1}\int_{B'_n}\frac{p_{n_2,\theta_2}}{p_{n_2,0}}(X_2)d\Pi(\theta_2)$

Donc $E_{\theta_0}\left[\Pi(B_n|X)\right] \leq \displaystyle\frac{n}{n-1}\int_{B'_n} E_{\theta_0}\left[\frac{p_{n_2,\theta_2}}{p_{n_2,0}}(X_2)\right]d\Pi(\theta_2)$

Or $E_{\theta_0}\left[\frac{p_{n_2,\theta_2}}{p_{n_2,0}}(X_2)\right] = 1$

Ainsi $E_{\theta_0}\left[\Pi(B_n|X)\right] \leq \displaystyle\frac{n}{n-1}\Pi(B'_n) \leq \frac{n}{n-1}\frac{1}{cp_n}E[|S\cap S_0^c|]$

Or $|S\cap S_0^c| \sim B(n_2,\alpha)$, donc $E_{\theta_0}\left[\Pi(B_n|X)\right] \leq \displaystyle\frac{n}{n-1}\frac{1}{cp_n}(n-p_n)\alpha$

D'où $E_{\theta_0}\left[\Pi(B_n|X)\right] \leq \displaystyle\frac{n-p_n}{n-1}\frac{1}{cp_n} \underset{n\to\infty}{\longrightarrow} 0$

□

Preuve du Lemme 7

Enoncé :
$\Pi\left(\theta|X, S=S^*\right) = \prod_{i\notin S^*}\delta_0(\theta_i)\prod_{i\in S^*}\psi_{X_i}(\theta_i)$

Démonstration. Il suffit de vérifier sur les évènements produits.
Soit $B \subset \mathbb{R}^n$, qu'on écrit $B = B_1 \times .. \times B_n$.

$$\Pi\left(\theta\in B|X, S=S^*\right) = \frac{\Pi\left(\theta \in B, S = S^*|X\right)}{\Pi\left(S=S^*|X\right)}$$

On a d'une part : $\Pi\left(S=S^*|X\right) = \tilde{\alpha}^{|S^*|}(1-\tilde{\alpha})^{n-|S^*|}$

Et d'autre part : $\Pi\left(\theta\in B, S=S^*|X\right) = \left(\displaystyle\prod_{i\in S^*}\int_{B_i}\tilde{\alpha}\psi_{X_i}(\theta_i)d\theta_i\right)\left(\prod_{i\notin S^*}(1-\tilde{\alpha})\mathbb{1}_{\{0\in B_i\}}\right)$

donc $\Pi\left(\theta \in B, S = S^*|X\right) = \tilde{\alpha}^{|S^*|}(1-\tilde{\alpha})^{n-|S^*|}\left(\prod_{i\in S^*}\int_{B_i}\psi_{X_i}(\theta_i)d\theta_i\right)\left(\prod_{i\notin S^*}\mathbb{1}_{\{0\in B_i\}}\right)$
d'où $\Pi\left(\theta \in B|X, S = S^*\right) = \left(\prod_{i\in S^*}\int_{B_i}\psi_{X_i}(\theta_i)d\theta_i\right)\left(\prod_{i\notin S^*}\mathbb{1}_{\{0\in B_i\}}\right)$

□

5.4 Preuve du théorème version semi-implicite

Enoncé

Pour les prior binomial, β et le complexity prior, avec γ densité d'une Laplace standard :

$$E_{\theta_0}[\Pi(\|\theta-\theta_0\|^2 \geq Mp_n\log(n)|X)] \underset{n\to\infty}{\longrightarrow} 0$$

avec M choisi assez grand

Démonstration

Soit B un évènement.

$$\begin{aligned}
\Pi(B|X) &= \frac{\int_B e^{-\frac{1}{2}\|X-\theta\|^2}d\Pi(\theta)}{\int e^{-\frac{1}{2}\|\varepsilon+\theta_0-\theta\|^2}d\Pi(\theta)} \\
&= \frac{\int_B e^{-\frac{1}{2}\|\varepsilon+\theta_0-\theta\|^2}d\Pi(\theta)}{\int e^{-\frac{1}{2}\|\varepsilon+\theta_0-\theta\|^2}d\Pi(\theta)}\text{sous } P_{\theta_0} \\
&= \frac{\int_B e^{\varepsilon^T(\theta-\theta_0)-\frac{1}{2}\|\theta_0-\theta\|^2}d\Pi(\theta)}{\int e^{\varepsilon^T(\theta-\theta_0)-\frac{1}{2}\|\theta_0-\theta\|^2}d\Pi(\theta)} \\
&= \frac{N}{D}
\end{aligned}$$

On commence par prouver les deux lemmes suivants :

Lemme 8. *On a* $D \geq \Pi(S_0)e^{-\|\theta_0\|_1-p_n}$

Lemme 9. *On suppose que* $B \subset \{|S| \leq Cp_n\}$ *et on se place sur l'évènement* $\mathcal{A}_n = \{\max|\varepsilon_i| \leq \sqrt{4\log n}\}$.
Alors, $\forall\theta \in B$ *et* $\forall A > 0$, *on a* $|\varepsilon^T(\theta-\theta_0)| \leq 2A(1+C)p_n\log n + \frac{1}{A}\|\theta-\theta_0\|^2$

On choisit alors $A = 4$ et on obtient, sur $\mathcal{A}_n$:

$$\begin{aligned}\Pi(B|X) &= \frac{N}{D} \\ &\leq \frac{e^{8(1+C)p_n \log n} \int_B e^{-\frac{1}{4}\|\theta_0-\theta\|^2} d\Pi(\theta)}{\Pi(S_0)e^{-\|\theta_0\|_1 - p_n}}\end{aligned}$$

On montre ensuite le résultat suivant :

Lemme 10. *On suppose que $B \subset \{|S| \leq Cp_n\} \cap \{\|\theta - \theta_0\|^2 > Mp_n \log n\}$.*
Alors $\int_B e^{-\frac{1}{4}\|\theta_0-\theta\|^2} d\Pi(\theta) \leq e^{C_2 p_n - \frac{M}{16} p_n \log n - \|\theta_0\|_1}$

On a de plus $\Pi(S_0) = \frac{\pi(p_n)}{\binom{n}{p_n}}$, on obtient ainsi que, pour un certain $M' > 0$, sur $\mathcal{A}_n$: $\Pi(B|X) \leq \frac{\binom{n}{p_n}}{\pi(p_n)} e^{-M' p_n \log n}$ pour $B \subset \{|S| \leq Cp_n\} \cap \{\|\theta - \theta_0\|^2 > Mp_n \log n\}$. On vérifie aisément que cette quantité tend vers 0 quand $n \to \infty$ pour les 3 a priori.

On peut donc conclure :

$$\begin{aligned}\Pi(\|\theta - \theta_0\|^2 \geq Mp_n \log(n)|X) &= \Pi(\|\theta - \theta_0\|^2 \geq Mp_n \log(n) \cap \{|S| \leq Cp_n\}|X) \\ &\quad + \Pi(\|\theta - \theta_0\|^2 \geq Mp_n \log(n) \cap \{|S| > Cp_n\}|X) \\ &\leq \Pi(\|\theta - \theta_0\|^2 \geq Mp_n \log(n) \cap \{|S| \leq Cp_n\} \cap \mathcal{A}_n|X) \\ &\quad + P(\mathcal{A}_n^c) + \Pi(\{|S| > Cp_n\}|X)\end{aligned}$$

Les deux premiers termes tendent bien en probabilité vers 0, il en est de même du dernier terme pour les 3 différents prior d'après le théorème sur la dimension démontré par Castillo et Van der Vaart, étant donné la décroissance exponentielle des 3 prior.

Preuve des lemmes

Lemme 8

$$D = \int e^{\varepsilon^T(\theta-\theta_0) - \frac{1}{2}\|\theta_0-\theta\|^2} d\Pi(\theta)$$

On a $D \geq \displaystyle\int_{S=S_0} e^{\varepsilon^T(\theta-\theta_0) - \frac{1}{2}\|\theta_0-\theta\|^2} d\Pi(\theta) =: I$

En notant $dG_{S_0}(\theta) = \displaystyle\prod_{i \in S_0} \gamma(\theta_i) d\theta_i$, on a $I = \Pi(S_0) \displaystyle\int e^{\varepsilon^T(\theta-\theta_0) - \frac{1}{2}\|\theta_0-\theta\|^2} dG_{S_0}(\theta)$

En notant $dG_{S_0}^{\theta_0}(\theta) = \displaystyle\prod_{i \in S_0} \gamma(\theta_i - \theta_{0,i}) d\theta_i$, on réécrit $I = \Pi(S_0) \displaystyle\int e^{\varepsilon^T(\theta-\theta_0) - \frac{1}{2}\|\theta_0-\theta\|^2 - \log(\frac{dG_{S_0}^{\theta_0}}{dG_{S_0}}(\theta)} dG_{S_0}^{\theta_0}(\theta)$

et on peut ensuite appliquer Jensen à la fonction exponentielle pour obtenir :

$$I \geq \Pi(S_0)\exp\Big(\int \varepsilon^T(\theta-\theta_0)dG_{S_0}^{\theta_0}(\theta) - \frac{1}{2}\int \|\theta_0-\theta\|^2 dG_{S_0}^{\theta_0}(\theta) - \int \log(\frac{dG_{S_0}^{\theta_0}}{dG_{S_0}}(\theta)dG_{S_0}^{\theta_0}(\theta)\Big)$$

Soit $I \geq \Pi(S_0)\exp\big(-\frac{1}{2}\int \|\theta_0-\theta\|^2 dG_{S_0}^{\theta_0}(\theta) - K(G_{S_0}^{\theta_0}, G_{S_0})\big)$
De plus on a :

$$\begin{aligned}
\frac{1}{2}\int \|\theta_0-\theta\|^2 dG_{S_0}^{\theta_0}(\theta) &= \frac{1}{2}\sum_{i\in S_0}\int (\theta_i-\theta_{0,i})^2\gamma(\theta_i-\theta_{0,i})d\theta_i \\
&= \frac{1}{2}\sum_{i\in S_0}\int u^2\gamma(u)du \\
&= \frac{1}{2}\sum_{i\in S_0} 2 \\
&\leq p_n
\end{aligned}$$

et

$$\begin{aligned}
K(G_{S_0}^{\theta_0}, G_{S_0}) &= \int (\sum_{i\in S_0}\log(\frac{\gamma(\theta_i-\theta_{0,i})}{\gamma(\theta_i)}))dG_{S_0}^{\theta_0}(\theta) \\
&= \int (\sum_{i\in S_0}\log(e^{-|\theta_i-\theta_{0,i}|+|\theta_i|}))dG_{S_0}^{\theta_0}(\theta) \\
&\leq \int (\sum_{i\in S_0}\log(e^{|\theta_{0,i}|}))dG_{S_0}^{\theta_0}(\theta) \\
&\leq \int (\sum_{i\in S_0}|\theta_{0,i}|)dG_{S_0}^{\theta_0}(\theta) \\
&\leq \int \|\theta_0\|_1 dG_{S_0}^{\theta_0}(\theta) \\
&\leq \|\theta_0\|_1
\end{aligned}$$

Et finalement on a bien $D \geq \Pi(S_0)e^{-\|\theta_0\|_1-p_n}$

Lemme 9

Soient $\theta \in B$ et $A > 0$.
Par Cauchy-Schwarz, on a $|\varepsilon^T(\theta-\theta_0)| \leq \sqrt{4\log n}\|\theta-\theta_0\|_1$
On a $2\times\sqrt{4\log n}\|\theta-\theta_0\|_1 = 4\sum_{i=1}^{n}\sqrt{\log n}|\theta_i-\theta_{0,i}|$
soit $2\times\sqrt{4\log n}\|\theta-\theta_0\|_1 = 4\sqrt{\log n}(\sum_{i\in S}|\theta_i-\theta_{0,i}|\frac{A}{A} + \sum_{i\in S_0\setminus S}|\theta_{0,i}|\frac{A}{A})$
donc $2\times\sqrt{4\log n}\|\theta-\theta_0\|_1 \leq (\sum_{i\in S}4A^2\log n + \frac{|\theta_i-\theta_{0,i}|^2}{A^2} + \sum_{i\in S_0\setminus S}4A^2\log n + \frac{|\theta_{0,i}|^2}{A^2})$

d'où $2 \times \sqrt{4 \log n}\|\theta - \theta_0\|_1 \leq 4A^2(1+C)p_n \log n + \sum_{i=1}^{n} \frac{|\theta_i - \theta_{0,i}|^2}{A^2}$
Et finalement, $\forall A > 0$, $|\varepsilon^T(\theta - \theta_0)| \leq A(1+C)p_n \log n + \frac{\|\theta-\theta_0\|^2}{A}$

Lemme 10

$$\int_B \exp(-\frac{\|\theta - \theta_0\|^2}{2})d\Pi(\theta) \leq \sum_{|S|\leq Cp_n} \Pi(S) \int_{\|\theta-\theta_0\|^2 > Mp_n \log n} \prod_{i\in S_0 \setminus S} e^{-\frac{\theta_{0,i}^2}{2}} \prod_{i \in S} e^{-\frac{(\theta_i - \theta_{0,i})^2}{4}} e^{-|\theta_i|} d\theta_i$$

or $\exp(-\frac{1}{4}\|\theta - \theta_0\|^2 - \|\theta\|_1) = \exp(-\frac{1}{4}\|\theta - \theta_0\|^2\| - \|\theta\|_1 + \theta_0\|_1 - \|\theta_0\|_1)$
et donc $\exp(-\frac{1}{4}\|\theta - \theta_0\|^2 - \|\theta\|_1) \leq \exp(-\frac{1}{4}\|\theta - \theta_0\|^2 + \|\theta - \theta_0\|_1 - \|\theta_0\|_1)$
de plus $\forall x > 0, -\frac{x^2}{4} + x \leq -\frac{x^2}{8} + 2$
donc $-\dfrac{1}{4}\|\theta - \theta_0\|^2 + \|\theta - \theta_0\|_1 \leq -\dfrac{1}{8}\|\theta - \theta_0\|^2 + 2\sum_{i \in S \cup S_0} 1$
d'où $-\dfrac{1}{4}\|\theta - \theta_0\|^2 + \|\theta - \theta_0\|_1 \leq -\dfrac{1}{8}\|\theta - \theta_0\|^2 + 2(C+1)p_n$
Ainsi $\displaystyle\int_B \exp(-\frac{\|\theta - \theta_0\|^2}{2})d\Pi(\theta) \leq \sum_{|S|\leq Cp_n} \Pi(S) e^{-\|\theta_0\|_1} e^{2(C+1)p_n} \int_{\|\theta-\theta_0\|^2 > Mp_n \log n} \prod_{i \in S_0 \cup S} e^{-\frac{(\theta_i - \theta_{0,i})^2}{4}} d\theta_i$
Or $\displaystyle\int_{\|\theta-\theta_0\|^2 > Mp_n \log n} \prod_{i \in S_0 \cup S} e^{-\frac{(\theta_i - \theta_{0,i})^2}{4}} d\theta_i \leq (2\pi)^{(C+1)p_n} e^{-\frac{M}{16}p_n \log n}$
et $\displaystyle\sum_{|S|\leq Cp_n} \Pi(S) \leq 1$, d'où :

$$\int_B \exp(-\frac{\|\theta - \theta_0\|^2}{2})d\Pi(\theta) \leq \exp((C+1)p_n(2 + \log(2\pi) - \frac{M}{16}p_n \log n - \|\theta_0\|_1)$$

5.5 Preuve du théorème général

Le résultat

Soit $\mathcal{P}$ une collection de mesures de probabilité. Soit $\Pi = \Pi_n$ une suite de prior et supposons que les observations X soient i.i.d. de densité p_0 : $dP_0(x) = p_0(x)dx$. Soit ε_n une suite de réels positifs telle que $\varepsilon_n \to 0$ et $\sqrt{n}\varepsilon_n \to \infty$ quand $n \to \infty$.
On suppose qu'il existe C et L des constantes telles que :

$$\Pi(p \in \mathcal{P}; -E_{P_0}[\log(\frac{p}{p_0}(X))] \leq \varepsilon_n^2, E_{P_0}[\log(\frac{p}{p_0}(X))^2] \leq \varepsilon_n^2) \geq e^{-Cn\varepsilon_n^2}$$

et

$$\Pi(\mathcal{P} \setminus \mathcal{P}_n) \leq Le^{-(C+4)n\varepsilon_n^2}$$

pour une suite $\mathcal{P}_n \subset \mathcal{P}$ telle qu'on puisse trouver des tests $\psi_n = \psi(X_1, \ldots, X_n)$ tels que $\forall n \in \mathbb{N}$ et $M > 0$ assez grand :

$$E_{P_0}[\psi_n] \to 0 \text{ et } \sup_{d(p,p_0) \geq M\varepsilon_n} E_P[1-\psi_n] \leq Le^{-(C+4)n\varepsilon_n^2}$$

Alors $\Pi(d(p,p_0) > M\varepsilon_n | X) \to 0$ quand $n \to \infty$ en $P_0^{\mathbb{N}}$-probabilité.

Démonstration

Tout d'abord, comme $E_{P_0}[\Pi(d(p,p_0) > M\varepsilon_n|X)\psi_n] \leq E_{P_0}[\psi_n]$ qui tend vers 0 quand $n \to \infty$, on doit juste prouver la convergence en probabilité vers 0 de

$$v := \Pi(d(p,p_0) > M\varepsilon_n|X)(1-\psi_n) = \frac{\int_{d(p,p_0)\geq M\varepsilon_n} \prod_{i=1}^n \frac{p}{p_0}(X_i) d\Pi(p)(1-\psi_n)}{\int_{\mathcal{P}} \prod_{i=1}^n \frac{p}{p_0}(X_i) d\Pi(p)}$$

On utilise pour cela le lemme suivant :

Lemme 11. *Pour tous $\varepsilon > 0$ et ν mesure de probabilité sur*

$$B = \{p \in \mathcal{P}; -E_{P_0}[\log(\frac{p}{p_0}(X))] \leq \varepsilon^2, E_{P_0}[\log(\frac{p}{p_0}(X))^2] \leq \varepsilon^2\}$$

on a, pour tout $c > 0$:

$$P_0^{\mathbb{N}}(\int_B \prod_{i=1}^n \frac{p}{p_0}(X_i) d\nu(p) \leq e^{-(1+c)n\varepsilon^2}) \leq \frac{1}{c^2 n\varepsilon^2}$$

On note $B_n = \{p \in \mathcal{P}; -E_{P_0}[\log(\frac{p}{p_0}(X))] \leq \varepsilon_n^2, E_{P_0}[\log(\frac{p}{p_0}(X))^2] \leq \varepsilon_n^2\}$

et $A_n = \{\int_{B_n} \prod_{i=1}^n \frac{p}{p_0}(X_i) d\Pi(p) \geq \Pi(B_n) e^{-2n\varepsilon_n^2} \geq e^{-(2+C)n\varepsilon_n^2}\}$

On peut alors appliquer notre lemme avec $\varepsilon = \varepsilon_n$, $c = 1$ et $\nu = \frac{\Pi_{|B_n}}{\Pi(B_n)}$ pour obtenir que $P_0^{\mathbb{N}}(A_n) \geq 1 - \frac{1}{n\varepsilon_n^2}$, qui tend vers 1 quand $n \to \infty$.
donc $\forall \varepsilon > 0$:

$$P_0^{\mathbb{N}}(v > \varepsilon) \leq P_0^{\mathbb{N}}(A_n^c) + P_0^{\mathbb{N}}(e^{(2+C)n\varepsilon_n^2}(1-\psi_n) \int_{d(p,p_0)\geq M\varepsilon_n} \prod_{i=1}^n \frac{p}{p_0}(X_i) d\Pi(p) > \varepsilon)$$

Pour gérer le second terme, on s'intéresse à $A = E_{P_0}[(1-\psi_n) \int_{d(p,p_0)\geq M\varepsilon_n} \prod_{i=1}^n \frac{p}{p_0}(X_i) d\Pi(p)]$

On a $A \leq E_{P_0}[\int_{p\in\mathcal{P}\setminus\mathcal{P}_n; d(p,p_0)\geq M\varepsilon_n} \prod_{i=1}^n \frac{p}{p_0}(X_i) d\Pi(p)] + E_{P_0}[(1-\psi_n) \int_{p\in\mathcal{P}_n; d(p,p_0)\geq M\varepsilon_n} \prod_{i=1}^n \frac{p}{p_0}(X_i) d\Pi(p)]$

donc $A \leq \int_{p\in\mathcal{P}\setminus\mathcal{P}_n; d(p,p_0)\geq M\varepsilon_n} (\int_{p_0>0} p(x)dx)^n d\Pi(p) + \sup_{p\in\mathcal{P}_n; d(p,p_0)\geq M\varepsilon_n} E_P[(1-\psi_n)]$

d'où $A \leq \Pi(\mathcal{P} \setminus \mathcal{P}_n) + Le^{-(C+4)n\varepsilon_n^2} \leq 2Le^{-(C+4)n\varepsilon_n^2}$

Finalement, $\forall \varepsilon > 0$:
$P_0^{\mathbb{N}}((1-\psi_n)\int_{d(p,p_0)\geq M\varepsilon_n}\prod_{i=1}^n \frac{p}{p_0}(X_i)d\Pi(p) > \varepsilon e^{-(2+C)n\varepsilon_n^2}) \leq \frac{2L}{\varepsilon}e^{-2n\varepsilon_n^2}$ qui tend bien vers 0 quand $n \to \infty$, ce qui conclut la preuve.

Preuve du lemme 11

Par Jensen, $\displaystyle\int_B \prod_{i=1}^n \frac{p}{p_0}(X_i)d\nu(p) \geq \sum_{i=1}^n \int_B \log\frac{p}{p_0}(X_i)d\nu(p)$
donc, en notant $\sqrt{n}(P_n - P_0)$ le processus empirique, on a :

$$A := P_0^{\mathbb{N}}(\int_B \prod_{i=1}^n \frac{p}{p_0}(X_i)d\nu(p) \leq e^{-(1+c)n\varepsilon^2}) \leq P_0^{\mathbb{N}}(\sum_{i=1}^n \int_B \log\frac{p}{p_0}(X_i)d\nu(p) \leq -(1+c)n\varepsilon^2)$$

$$A \leq P_0^{\mathbb{N}}(\sqrt{n}\int\int_B \log\frac{p}{p_0}(X_i)d\nu(p)d(P_n-P_0) \leq -\sqrt{n}(1+c)\varepsilon^2 - \sqrt{n}\int\int_B \log\frac{p}{p_0}(x)d\nu(p)dP_0(x))$$

Par Fubini, $\sqrt{n}\int\int_B \log\frac{p}{p_0}(x)d\nu(p)dP_0(x) = \sqrt{n}\int_B \big(-E_{P_0}[\log\frac{p}{p_0}]\big)d\nu(p) \leq \sqrt{n}\varepsilon^2$

D'où $A \leq P_0^{\mathbb{N}}(\sqrt{n}\displaystyle\int\int_B \log\frac{p}{p_0}(X_i)d\nu(p)d(P_n - P_0) \leq -\sqrt{n}c\varepsilon^2)$

Donc $A \leq \dfrac{1}{c^2 n\varepsilon^4} Var_{P_0}\displaystyle\int \log\frac{p}{p_0}(X)d\nu(p)$,

soit par Jensen : $A \leq \dfrac{1}{c^2 n\varepsilon^4} E_{P_0}\displaystyle\int \big(\log\frac{p}{p_0}(X)\big)^2 d\nu(p)$

Et finalement $A \leq \frac{1}{c^2 n\varepsilon^2}$, ce qui conclut la preuve.

5.5.1 Un résultat d'existence de tests point contre complémentaire de boule

On sait comment construire des tests de P_0 contre la boule $\{P; d(P,P_1) \leq \eta\}$ pour un P_1 donné. On considèrera dans ce paragraphe une distance d telle que pour toute paire $(P_0; P_1) \in \mathcal{P}^2$ il existe des tests ω_n tels qu'il existe $K > 0$ telle que pour tout n entier naturel :

$$E_{P_0}[\omega_n] \leq e^{-Knd^2(P_0,P_1)} \leq \text{ et } \sup_{d(P,P_1)\leq d(P_0,P_1)/2} E_P[1-\omega_n] \leq e^{-Knd^2(P_0,P_1)} \tag{5.1}$$

On peut par exemple prendre pour$d$ la distance en variation totale ou d'Hellinger.
On notera $D(\varepsilon, \mathcal{P}, d)$ le nombre maximal de points de $\mathcal{P}$ dont la distance entre chaque paire est au moins ε. On a alors le résultat suivant :

Théorème 8. *On suppose l'existence d'une fonction décroissante $D(\varepsilon)$ telle que pour tout $\varepsilon_n \geq 0$ et tout $\varepsilon > \varepsilon_n$ on ait : $D(\frac{\varepsilon}{2}, \{P; \varepsilon \leq d(P,P_0) \leq 2\varepsilon\}, d) \leq D(\varepsilon)$*

Alors $\forall \varepsilon > \varepsilon_n$, il existe des tests ϕ_n (qui dépendent de $\varepsilon > 0$) tels qu'il existe $K > 0$ telle que $\forall j \in \mathbb{N}^*$:

$$E_{P_0}[\phi_n] \leq D(\varepsilon)\frac{e^{-Kn\varepsilon^2}}{1-e^{-Kn\varepsilon^2}} \text{ et } \sup_{d(P,P_0)\geq j\varepsilon} E_P[1-\phi_n] \leq e^{-Kn\varepsilon^2 j^2}$$

Démonstration. Soit $j \in \mathbb{N}^*$.
On considère l'ensemble $S_j = \{P; j\varepsilon < d(P,P_0) \leq (j+1)\varepsilon\}$, dont on tire un ensemble maximal S'_j de points $j\varepsilon/2$-séparés. Par hypothèse, cet ensemble contient au plus $D(j\varepsilon)$ points. Tout point P de S_j est à distance au moins $j\varepsilon/2$ d'un de ces points.
Soit alors $P_1 \in S'_j$, il existe des tests $\omega_{P_1,n}$ vérifiant (5.5.1).
$\forall n \in \mathbb{N}$, on pose $\phi_n = \sup_{j\in\mathbb{N}^*} \sup_{P_1\in S'_j} \omega_{P_1,n}$
Ainsi $E_{P_0}[\phi_n] \leq \sum_{j\in\mathbb{N}^*} \sum_{P_1\in S'_j} e^{-Kn\varepsilon^2 j^2}$
donc $E_{P_0}[\phi_n] \leq \sum_{j\in\mathbb{N}^*} D(j\varepsilon) e^{-Kn\varepsilon^2 j^2}$
d'où $E_{P_0}[\phi_n] \leq D(\varepsilon) \sum_{j\in\mathbb{N}^*} e^{-Kn\varepsilon^2 j} = D(\varepsilon)\frac{e^{-Kn\varepsilon^2}}{1-e^{-Kn\varepsilon^2}}$
On a de plus : $\sup_{P\in\bigcup_{i\geq j} S_i} E_P[1-\phi_n] \leq \sup_{i\geq j} e^{-Kn\varepsilon^2 i^2} \leq e^{-Kn\varepsilon^2 j^2}$ □

Bibliographie

[1] Iain M. Johnstone et Bernard W. Silverman, *Needles and straw in haystacks : empirical Bayes estimates of possibly sparse sequences*, (2004).Ann.Statist.**32** 1594-1649.

[2] Subhashis Ghosal, Jayanta K. Ghosh et Aad W. van der Vaart, *Convergence rates of posterior distributions*, (2000).Ann.Statist.**28(2)** 500-531.

[3] Ismaël Castillo et Aad W. van der Vaart, *Needles and straw in haystacks : posterior concentration for possibly sparse sequences* , (2012).Ann.Statist.**40(4)** 2069-2101.

[4] Ismaël Castillo et Aad W. van der Vaart, *Supplement to 'Needles and straw in haystacks : posterior concentration for possibly sparse sequences'* , (2012).Ann.Statist.**40(4)**.

[5] Bin Yu, *Assouad, Fano and Le Cam*

[6] Judith Rousseau et Botond Szabo, *Asymptotic behaviour of the empirical bayes posteriors associated to maximum marginal likelihood estimator*

[7] Christophe Giraud, *Introduction to high dimensional statistics*

[8] David L. Donoho, Iain M. Johnstone, J.C. Hoch et A.S. Stern, *Maximum entropy and the nearly black object* , (1992)J.R.Stat.Soc.Ser.B Stat. Methodol. **54** 41-81.