

Dossier campagne de recrutement MCF 2012

JEAN-BAPTISTE MONNIER

LPMA, Université Paris Diderot - Paris 7

E-mail: monnier@math.jussieu.fr

Page Web: <https://sites.google.com/site/jrmonnier/home>

1^{er} mars 2012

Table des matières

Curriculum Vitae	2
Publications	3
Enseignements	4
Interventions et rencontres scientifiques	5
Travaux antérieurs	6
Programme de recherche	8
Rapport de thèse	11

Jean-Baptiste Monnier

LPMA, Université Paris 7

Bureau 5B01
175 rue du Chevaleret
75013 Paris, France

☎ +33601633478

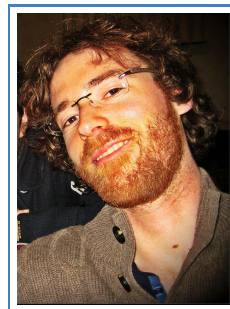
✉ j.r.monnier@gmail.com

🌐 sites.google.com/site/jrmonnier/

Né le 23/11/1983 à Paris 75012

Nationalité Française

Marié le 15/09/2012



Parcours Professionnel

Recherche Universitaire

- 2012–2012 **Post-doctorant**, LPMA, Université Paris 7, Doctorant contractuel avec monitorat.
- 2009–2011 **Thèse de doctorat en statistiques**, LPMA, Université Paris 7, Doctorant contractuel avec monitorat, Titre de thèse: “Quelques contributions en classification, régression et étude d’un problème inverse en finance”.
- Soutenue publiquement le 6/12/2011 devant le jury composé de: Jean-Yves Audibert (examinateur), Francis Bach (président), Arnaud Gloter (rapporteur), Oleg Lepski (examinateur), Dominique Picard (directrice de thèse), Peter Tankov (examinateur)
- 2004–2004 **Assistant de Recherche**, University of California at Santa-Barbara, USA, Department of Mechanical Engineering, groupe de recherche du Pr. George Homsy.

Enseignement

- 2009–2012 **Moniteur**, Université Paris 7.
- ◇ 2011–2012: TP logiciel R niveau L2 (40h), enseigné par Pr. K. Tribouley
 - ◇ 2010–2011: Économétrie niveau L3 (40h), enseigné par Pr. P. Alquier
 - ◇ 2010–2011: Statistiques Bayésiennes niveau M1 (40h), enseigné par Pr. P. Alquier
 - ◇ 2009–2010: Statistiques descriptives niveau L1,2,3 (64h), enseigné par Pr. S. Boucheron

Industrie

- 2007–2008 **Chercheur Quantitatif**, JP Morgan Chase Bank, Londres.
- 2006–2007 **Économiste Quantitatif**, Endeavour Capital Management, Londres.

Parcours Universitaire

- 2009–2011 **Thèse de doctorat en statistiques**, LPMA, Université Paris 7, Mention Très Honorable.
- 2008–2009 **M2 Modélisation Aléatoire**, LPMA, Université Paris 7, Mention TB (rang 1^{er}).
- 2005–2006 **MSc Finance and Economics**, London School of Economics, Mention TB.
- 2003–2006 **École Centrale Paris**, Spécialité mathématiques appliquées et économie.
- 2001–2003 **Classes Préparatoires**, Lycée Louis-le-Grand, Paris.

Responsabilités Administratives

- 2011–2013 **Représentant des doctorants**, LPMA, Université Paris 7.

Compétences Linguistiques et Informatiques

- Langages: Anglais (Courant), Allemand (Intermédiaire), Espagnol (Débutant), Italien (Débutant)
- Programmation: C, C++, Python. Logiciels: Matlab, Scilab, R

Publications

JEAN-BAPTISTE MONNIER

LPMA, Université Paris Diderot - Paris 7

E-mail: monnier@math.jussieu.fr

Page Web: <https://sites.google.com/site/jrmonnier/home>

En cas d'audition, je présenterai [2] et [3].

Articles parus ou à paraître

- [1] Jean-Baptiste Monnier. Non-parametric regression on the hyper-sphere with uniform design. *TEST*, 20(2) :412–446, 2010. doi : 10.1007/s11749-011-0233-7. URL <http://www.springerlink.com/content/r847j3520872323v/>.
- [2] Jean-Baptiste Monnier. Classification via local multi-resolution projections. *Electronic Journal of Statistics*, 2011. URL <http://arxiv.org/abs/1110.6427>. To appear.

Article soumis

- [3] Jean-Baptiste Monnier. Spectral analysis of restricted call and put operators and application to stable risk-neutral density recovery. Submitted, 2011. URL <http://hal.archives-ouvertes.fr/hal-00651239/en/>.

Rapport technique et manuscrit de thèse de doctorat

- [4] Jean-Baptiste Monnier. Classification via local multi-resolution projections (extended version). Technical report, LPMA, Université Paris Diderot - Paris 7, 2011. URL <http://hal.archives-ouvertes.fr/hal-00637010/en/>.
- [5] Jean-Baptiste Monnier. *Quelques contributions en classification, régression et étude d'un problème inverse en finance*. PhD thesis, Université Paris Diderot, Paris 7, 12 2011. URL <http://tel.archives-ouvertes.fr/tel-00650930/en/>.

Travaux en cours

- [6] Jean-Baptiste Monnier. Minimax risk-neutral density recovery over ellipsoids. Technical Report, 2012.
- [7] Jean-Baptiste Monnier. Minimax optimality of the local multi-resolution projection estimator over Besov spaces. Preprint, 2012. URL <http://hal.archives-ouvertes.fr/hal-00674091>.

Article paru en sciences physiques

- [8] Éric Lajeunesse, Jean-Baptiste Monnier, and George Homsy. Granular slumping on a horizontal surface. *Phys. Fluids*, 17 :103302, 2005. URL http://pof.aip.org/resource/1/phf1e6/v17/i10/p103302_s1.

Enseignements

JEAN-BAPTISTE MONNIER

LPMA, Université Paris Diderot - Paris 7

E-mail: monnier@math.jussieu.fr

Page Web: <https://sites.google.com/site/jrmonnier/home>

Dans le cadre de mon contrat doctoral, j'ai opté pour un service de monitorat à l'université Paris Diderot - Paris 7. J'ai accompli mon service de monitorat comme chargé de TD/TP dans le cadre des cours décrits ci-dessous.

- 2009-2010 : TD de statistiques descriptives en L1, L2 et L3 (64h), cours enseigné par S. Boucheron. Contenu du cours : analyse en composantes principales, classification (non) supervisée, régression, analyse des corrélations.
- 2010-2011 : TD d'économétrie en L3 (40h), cours enseigné par P. Alquier. Contenu du cours : régression linéaire, estimation et tests, séries temporelles, données de panel.
- 2010-2011 : TD de statistiques Bayésiennes en M1 (40h), cours enseigné par P. Alquier. Contenu du cours : théorie de la décision, calculs d'estimateurs de Bayes, tests Bayésiens, introduction à l'apprentissage statistique, règle des plus proches voisins, analyse en composantes principales.
- 2011-2012 : TP de statistiques descriptives sur logiciel R en L2 (40h), cours enseigné par K. Tribouley. Contenu du cours : lois de probabilité, densité, fonction de répartition, réalisation, loi des grands nombres, indicateurs de forme, estimateurs à noyau de la densité, tests d'ajustement de Shapiro et Kolmogorov Smirnov, test d'indépendance, régression linéaire.

Interventions et rencontres scientifiques

JEAN-BAPTISTE MONNIER

LPMA, Université Paris Diderot - Paris 7

E-mail: monnier@math.jussieu.fr

Page Web: <https://sites.google.com/site/jrmonnier/home>

Intervention à une conférence internationale

- 02/12/2011, Rencontres de Statistique Mathématique, CIRM, Marseille, France, invité par Christophe Pouet et Alexandre Tsybakov, *Classification via local multi-resolution projections*.

Interventions à des séminaires

- 17/06/2010, Groupe de Travail des Thésards, LPMA, Universités Paris 6-7, France, invité par Mathieu Richard, *Sparsity, minimax framework and needlets*.
- 03/10/2011, Séminaire de Probabilités et Statistiques, LAREMA, Université d'Angers, France, invité par Sébastien Lousteau, *Local multi-resolution on a random design and application to classification under the margin assumption*.
- 21/10/2011, Séminaire de Statistique, Analyse et Modélisation Multi-disciplinaire (SAMM), Université Paris 1, France, invité par Jean-Marc Bardet et Annie Millet, *Local multi-resolution regression on a random design*.
- 27/10/2011, Groupe de Travail des Thésards, LPMA/LSTA, Université Paris 6-7, France, invité par Clément Foucart et Éric Luçon, *Nonparametric regression on the hyper-sphere with uniform design*.
- 28/11/2011, Groupe de Travail de Statistique de Jussieu, Université Paris 6, France, invité par Olivier Lopez, Bertrand Michel et Tabea Rebafka, *Classification via local multi-resolution projections*.
- 16/01/2012, Groupe de Travail de Machine-Learning SMILE, ENS Paris, France, invité par Pierre Alquier et Gilles Stolz, *Classification via local multi-resolution projections*.
- 23/01/2012, Groupe de Travail de Statistique du CREST, ENSAE, Paris, France, invité par Pierre Alquier et Alexandre Tsybakov, *Classification via local multi-resolution projections*.
- 17/02/2012, Groupe de Travail de Statistique, Institut Elie Cartan, Nancy, France, invité par Sandie Ferrigno et Aurelie Muller-Gueudin, *Spectral analysis of restricted call and put operators and application to stable risk-neutral density recovery*.
- 07/03/2012, Séminaire du DGEI-DEMFI-RECFIN, Banque de France, Paris, France, invité par Jean-Stéphane Mésonnier, *Spectral analysis of restricted call and put operators and application to stable risk-neutral density recovery*.
- 20/03/2012, Séminaire de Statistique, Institut de Mathématiques, Toulouse, France, invité par Clément Marteau, *Classification via local multi-resolution projections*.

Participation à des conférences internationales

- 14-18/12/2009, Rencontres de Statistique Mathématique, CIRM, Marseille, France, organisé par Laurent Cavalier and Oleg Lepski.
- 30/08-03/09/2010, Statistiques Mathématiques et Applications, La Villa Clythia, Fréjus, France, organisé par Sylvie Huet, Pascal Massart, Sara van de Geer, Aad van der Vaart and Jean-Philippe Vert.
- 13-17/12/2010, Rencontres de Statistique Mathématique, CIRM, Marseille, France, organisé par Laurent Cavalier et Oleg Lepski.
- 28/11-02/12/2010, Rencontres de Statistique Mathématique, CIRM, Marseille, France, organisé par Alexander Goldenshluger, Alexandre Tsybakov et Christophe Pouet.

Travaux antérieurs

JEAN-BAPTISTE MONNIER

LPMA, Université Paris Diderot - Paris 7

E-mail: monnier@math.jussieu.fr

Page Web: <https://sites.google.com/site/jrmonnier/home>

Ma thèse [5] est consacrée à l'étude des problèmes de classification binaire supervisée, de régression en design aléatoire et d'estimation de la densité risque-neutre à partir des prix d'options cotés sur le marché. Elle est constituée de trois publications. La première [2] touche aux problèmes de classification binaire supervisée et de régression sur un design aléatoire à valeurs dans \mathbb{R}^d . La seconde [1] traite du problème de régression sur un design uniformément distribué à la surface de l'hyper sphère de \mathbb{R}^d . Enfin, on prouve dans la troisième [3] qu'il est possible d'élaborer une méthode simple et efficace d'estimation de la densité risque-neutre à partir d'une décomposition en valeurs singulières explicite d'opérateurs de prix restreints.

Classification via projections multi-résolution localisées

On s'intéresse dans [2] et [4] au problème de classification binaire supervisée. Celui-ci consiste à deviner le label Y associé à une co-variable $X \in \mathbb{R}^d$, étant donnée l'observation de n vecteurs (X_i, Y_i) indépendant et identiquement distribués constitués chacun d'une co-variable X_i et du label Y_i associé. On suppose que la loi du vecteur (X, Y) est inconnue et que la loi marginale de X admet une densité supportée sur un sous-ensemble \mathcal{A} de \mathbb{R}^d . Dans le cas particulier des classifieurs de type *plug-in*, résoudre le problème de classification revient à estimer la fonction de régression $\eta(X) = \mathbb{E}[Y|X]$. En supposant initialement que \mathcal{A} est connu, on montre comment il est possible de construire un estimateur de η par projections localisées sur une analyse multi-résolution (MRA). Dans un second temps, on montre comment cette procédure d'estimation se généralise au cas où \mathcal{A} est inconnu. De manière intéressante, cette nouvelle méthode d'estimation présente des performances théoriques d'optimalité minimax similaires à celles de l'estimateur par polynômes locaux (LPE). De plus, elle bénéficie de la structure de treillis de la MRA sous-jacente et présente des performances calculatoires plus intéressantes que celles du LPE, ce qui est d'un intérêt crucial dans le cadre de nombreuses applications pratiques. Finalement, on prouve que le classifieur *plug-in* associé atteint des vitesses super rapides sous une hypothèse de marge.

Régression sur l'hyper sphère \mathbb{S}^d de \mathbb{R}^{d+1}

On s'intéresse dans [1] au problème de régression sur un design uniformément distribué à la surface de l'hyper sphère \mathbb{S}^d de \mathbb{R}^{d+1} . Les procédures traditionnelles d'estimation en ondelettes ne sont plus valides dans ce contexte puisque les bases d'ondelettes vivent sur un treillis euclidien dont la géométrie n'est pas adaptée à celle de l'hyper sphère. Cependant, il est possible d'élaborer un *tight frame* de needlets de $\mathbb{L}_2(\mathbb{S}^d)$ à partir d'harmoniques sphériques, dont les éléments vérifient des propriétés de localisation similaires à celles des ondelettes euclidiennes. Nous montrons dans [1] comment les méthodes traditionnelles de régression par seuillage de coefficients d'ondelettes peuvent être transposées à l'hyper sphère grâce aux needlets. On exhibe deux estimateurs respectivement élaborés à partir d'un seuillage déterministe et stochastique et on prouve qu'il sont minimax optimaux sur une large échelle d'espaces de Besov et pour une large classe de pertes $\mathbb{L}_p(\mathbb{S}^d)$, $p \in [1, \infty]$. On en profite pour mener un vaste éventail de simulations numériques avec la procédure construite sur un seuillage aléatoire des coefficients de needlets.

Analyse spectrale d'opérateurs de prix restreints et application au recouvrement de la densité risque-neutre

Les densités risque-neutres sont d'un intérêt crucial pour les banques centrales puisqu'elles représentent le sentiment du marché par rapport au prix d'un sous-jacent à une date future. Elles sont donc un précieux indicateur de l'intérêt ou de l'impact d'une politique de régulation financière donnée. Nous

proposons dans [3] une nouvelle méthode pour estimer une densité risque-neutre (RND) directement à partir des prix à l'offre et à la vente des options de put cotés sur le marché. Plus précisément, nous proposons de voir le problème d'estimation de la RND comme un problème inverse. Nous montrons d'abord qu'il est possible de définir des opérateurs de prix d'option de call et de put "restreints" qui admettent une décomposition en valeurs singulières (SVD) explicite. Nous montrons ensuite que ce nouveau cadre théorique permet d'élaborer un algorithme de programmation quadratique simple et rapide destiné à recouvrer la RND la plus régulière compatible avec les prix d'options de put cotés sur le marché. On baptise cette méthode par le nom de "méthode d'estimation spectrale" (SRM). De manière intéressante, la SVD des opérateurs de prix restreints offre un nouveau point de vue sur le problème d'estimation de la RND à partir des prix de marché. La SRM présente de nombreux avantages sur les méthodes d'estimation de la RND existantes en ce sens que, (1) bien que totalement non paramétrique, elle est simple à mettre en oeuvre et donne lieu à un algorithme rapide puisqu'il s'agit de résoudre un unique programme quadratique ; (2) elle prend les prix à l'achat et à la vente d'options de put comme seuls paramètres d'entrée et ne requiert donc aucun pré-traitement des données de marché ; (3) elle est robuste en ce sens qu'elle fonctionne très bien avec seulement quelques prix de marché ; (4) elle retourne la densité la plus régulière donnant lieu à des prix d'option de put situés dans la fourchette bid-ask des prix correspondants cotés sur les marchés ; (5) elle retourne une formule fermée pour la RND sur l'intervalle $[0, B]$ de \mathbb{R}^+ où B est une constante positive qui peut être choisie de manière arbitraire. Nous obtenons donc l'intégralité de la queue gauche de la RND ainsi que sa partie centrale et une partie de sa queue droite. Nous confrontons cet algorithme à des données réelles et simulées et obtenons des résultats convaincants. La SRM est donc sans aucun doute une alternative intéressante aux autres méthodes d'estimation de la RND.

Programme de recherche

JEAN-BAPTISTE MONNIER

LPMA, Université Paris Diderot - Paris 7

E-mail: monnier@math.jussieu.fr

Page Web: <https://sites.google.com/site/jrmonnier/home>

Notre travail de thèse [5] s'inscrit dans le domaine de la statistique mathématique et aborde à la fois des problèmes classiques en statistique non paramétrique (tel que le problème de régression dans [1]) ainsi que plus modernes en apprentissage statistique (tel que le problème de classification dans [2]) ou en mathématiques financières (tel que le problème d'estimation de la densité risque neutre dans [3]). Tous ces travaux ont été effectués avec le souci de proposer des algorithmes de résolution performants en pratique et faciles à mettre en oeuvre. Chacun de nos travaux est donc accompagné de simulations numériques.

Notre programme de recherche consiste (i) à affiner et généraliser les résultats obtenus dans nos travaux antérieurs [1, 2] et [3] et (ii) à explorer un nouveau problème d'apprentissage statistique *en ligne*, à la frontière entre la statistique non paramétrique et la théorie des jeux : le problème du bandit stochastique à continuum de bras. Ce dernier problème trouve de nombreuses applications pratiques, notamment dans le cadre des systèmes de recommandation en ligne ou de la mise en oeuvre de politiques monétaires par les banques centrales.

Étude fine de l'estimateur par projections multi-résolutions localisées en régression

Dans le travail en cours [7], nous considérons le problème de régression en design aléatoire (voir [2]). Nous montrons dans un premier temps comment les résultats d'optimalité minimax de l'estimateur par projections multi-résolutions localisées (dit "estimateur PMRL") décrits dans [2] restent valides sur des classes de régularité bien plus larges que les espaces de Lipschitz généralisés (ou de Hölder). Comme illustré sur la Figure 1 ci-dessous, on étend l'optimalité de l'estimateur PMRL à une vaste échelle de Besov $B_{\tau,q}^s$, $s > 0$, lorsque la perte est mesurée en norme \mathbb{L}_p , $p \in [1, \infty]$ (voir [7]).

Dans un second temps, nous souhaitons travailler directement sur la version adaptative (non-linéaire) de l'estimateur PMRL obtenu via la méthode de Lepski (voir [12] et [5]). Nous espérons pouvoir prouver que cette dernière procédure est minimax optimale (en design aléatoire quelconque) sur une vaste échelle de Besov comparable au domaine d'optimalité (en design uniforme) des estimateurs plus traditionnels obtenus par seuillage de coefficients d'ondelettes (voir [5]).

Estimation de densité risque neutre

Estimation de densité risque neutre avec contrôle du reste

Comme détaillé dans [3], le problème d'estimation de la densité risque neutre (RND) à partir des prix d'options cotés sur le marché diffère des problèmes inverses traditionnellement rencontrés en statistique dans la mesure où nous ne disposons tout au plus que de 50 prix de put par maturité. Dans un tel contexte, le contrôle du reste de notre estimateur de la RND sous des hypothèses de régularité semble a priori délicat.

Toutefois, dans un souci d'exhaustivité, nous avons entamé la rédaction du rapport technique [6] dans lequel (i) nous recensons les méthodes d'estimation plus traditionnelles en statistique des problèmes inverses avec contrôle de reste; (ii) puis nous explorons dans quelle mesure ces méthodes peuvent être modifiées ou adaptées pour éventuellement donner lieu à un estimateur réaliste de la RND avec contrôle de reste sous hypothèse de régularité.

Estimation de densité risque neutre multi-variée par méthode spectrale

Nous souhaitons par ailleurs étendre les résultats de [3] à d'autres options européennes telles que les options binaires. Ce travail consiste donc à étudier les propriétés des SVD associées à ces nouveaux opérateurs de prix restreints.

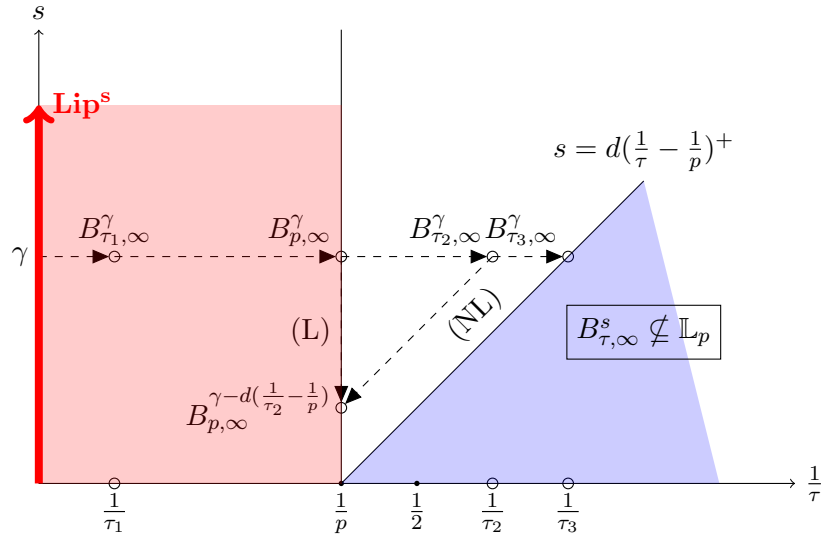


FIGURE 1 – Chaque point de cette figure représente un espace de Besov $B_{\tau, \infty}^s$ comme un point du plan de coordonnées $(s, 1/\tau)$. On suppose que la perte en régression est mesurée en norme \mathbb{L}_p . Les espaces de Lipschitz généralisés sont les espaces de Besov $B_{\infty, \infty}^s$ et correspondent donc à l'axe des ordonnées d'équation $1/\tau = 0$. Les espaces de Besov $B_{\tau, \infty}^s$ sont de plus en plus gros le long des lignes tracées en pointillés en suivant le sens des flèches. On peut montrer que l'estimateur PMRL est minimax optimal dans toute la zone $\{(1/\tau, s) : 1/\tau \leq 1/p\}$ de couleur saumonée. On souhaite de plus montrer que la version adaptative de l'estimateur PMRL reste minimax optimale dans un certain sous domaine de $\{(1/\tau, s) : 1/\tau > 1/p\} \cap \{(1/\tau, s) : s \geq d(1/\tau - 1/p)^+\}$.

Nous voudrions ensuite étudier comment cette approche spectrale pourrait être utilisée dans le cadre de l'estimation de RND bi-variées à partir de prix d'options de spread, par exemple. Là encore, il s'agit d'étudier les opérateurs de prix restreints correspondants et les propriétés des SVD associées. Remarquons en passant que la littérature sur l'estimation de densité risque neutre multi-variée est quasi inexistante.

Needlets et problèmes inverses

Problèmes inverses sur la sphère

Comme décrit dans [1] et [5], les frames de needlets sont à la sphère ce que les bases d'ondelettes sont aux espaces euclidiens. Il est bien connu que les ondelettes jouent un rôle essentiel dans la résolution de problèmes inverses (voir [10] par exemple). Nous souhaitons donc étudier dans quelle mesure les frames de needlets peuvent jouer un rôle équivalent dans le cadre de problèmes inverses sur la sphère.

Needlets en finance

On prouve dans [3] qu'il est possible d'obtenir un estimateur q_N de la densité risque neutre (RND) comme solution d'un programme quadratique (QP) dont les contraintes dépendent des prix d'options de put cotés sur le marché. Cet estimateur q_N s'exprime comme une combinaison linéaire finie de vecteurs singuliers d'un opérateur de put "restreint". Remarquons ici que ces vecteurs singuliers ne sont pas localisés, de telle sorte qu'une erreur de cotation sur un prix de marché (qui intervient comme contrainte dans QP) contamine tout l'estimateur q_N .

Les frames de needlets semblent être un outil adéquate pour palier à ce problème. Les frames de needlets sur la sphère sont construits de manière très générique à partir de la base d'harmoniques sphériques de $\mathbb{L}_2(\mathbb{S}^d)$ (voir [13, 14]). Si bien qu'il est aussi possible de construire des "frames de needlets" à partir des vecteurs singuliers des opérateurs de prix restreints introduits dans [3]. Nous souhaiterions naturellement étudier dans un premier temps comment ces "frames de needlets" pourraient permettre d'élaborer un estimateur de la RND qui soit robuste aux erreurs de cotation sur le marché, en ce sens

qu’une erreur de cotation n’aurait qu’un impact localisé sur q_N . Nous voudrions de plus étudier si il est possible d’élaborer un estimateur de la RND à partir de ces needlets avec un contrôle du reste sous une hypothèse de régularité de type Besov et pour une vaste échelle de pertes mesurées en norme \mathbb{L}_p , $p \in [1, \infty)$.

Classification et régression sur la sphère en design non uniforme

Comme détaillé dans [5], il est possible d’étendre la méthode d’estimation par projections multi-résolutions localisées (PMRL) et les résultats décrits dans [2] à la sphère de \mathbb{R}^d en utilisant le frame de needlets à support “compact” introduit dans [11]. Une telle méthode d’estimation devrait s’avérer particulièrement judicieuse dans le cadre d’applications en astrophysique où les points du design ne sont pas uniformément distribués à la surface de la sphère.

Problèmes de bandit à continuum de bras

Finalement et dans un dernier travail plus ouvert, nous souhaiterions aborder le problème du bandit à continuum de bras et en explorer les liens avec les problèmes plus classiques de statistique non paramétrique. Le problème du bandit stochastique à continuum de bras est une extension du problème du bandit stochastique à deux bras initialement introduit dans [15]. Une description formelle de ce problème peut être par exemple trouvée dans [9] et [17]. Comme détaillé dans [16], les stratégies optimales de jeu nécessitent la connaissance de la régularité de l’environnement. Un des enjeux de notre travail consisterait donc à tisser des liens entre les méthodes adaptatives en statistique non paramétrique et le problème du bandit stochastique à continuum de bras, de manière à exhiber des stratégies optimales qui ne reposent pas sur la connaissance a priori de la régularité de l’environnement.

Bibliographie

- [9] Sébastien Bubeck. *Bandits games and clustering foundations*. PhD thesis, Université Lille 1, 2010.
- [10] David Donoho. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harmon. Anal.*, 2(2) :101–126, 1995.
- [11] George Kyriazis and Pencho Petrushev. “compactly” supported frames for spaces of distributions on the ball. Preprint, 2011.
- [12] Oleg V. Lepski, Enno Mammen, and Vladimir G. Spokoiny. Optimal spatial adaptation to inhomogeneous smoothness : an approach based on kernel estimates with variable bandwidth selectors. *Ann. Stat.*, 25(3) :392–947, 1997.
- [13] Francis Narcowich, Pencho Petrushev, and John Ward. Decomposition of Besov and Triebel-Lizorkin spaces on the sphere. *J. Funct. Anal.*, 238 :530–564, 2006.
- [14] Francis Narcowich, Pencho Petrushev, and John Ward. Localized tight frames on spheres. *SIAM J. Math. Anal.*, 38(2) :574–594, 2007.
- [15] Herbert Robbins. Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc.*, 58(5) :527–535, 1952.
- [16] Gilles Stoltz Sébastien Bubeck and Jia Yuan Yu. Lipschitz bandits without the lipschitz constant. In *ALT, 22nd International Conference on Algorithmic Learning Theory, Espoo, Finland*, 2011.
- [17] Gilles Stoltz. Contributions à la prévision séquentielle de suites arbitraires. HDR, 02 2011.

Rapport sur le manuscrit de thèse de Monsieur Jean-Baptiste Monnier.

Le sujet de la thèse de M. Monnier est la statistique non-paramétrique. La thèse propose trois contributions originales sur ce sujet. Les deux premières contributions ont pour thème l'estimation non paramétrique d'une fonction de régression. La troisième contribution concerne le thème, à priori assez différent, de la calibration pour le calcul du prix d'options européennes.

L'organisation de la thèse est la suivante.

Le premier chapitre du manuscrit est une longue introduction. L'auteur y présente les résultats obtenus dans les 3 contributions, ainsi qu'une revue des travaux antérieurs sur les différents thèmes étudiés. Il convient de souligner l'extrême clarté avec laquelle l'auteur présente ses résultats dans l'introduction. L'apport nouveau de la thèse, en comparaison des résultats existants, est clairement présenté. Les difficultés mathématiques rencontrées par l'auteur, ainsi que les solutions proposées, sont remarquablement bien décrites.

Les chapitres 2, 3 et 4 sont des rappels mathématiques et ne contiennent pas de résultats nouveaux. Le but de ces 3 chapitres est de faciliter la lecture de la suite de la thèse.

Le chapitre 2 contient des rappels sur l'estimation, au sens minimax, de la fonction de régression f ,

$$Y_i = f(X_i) + \xi_i, \quad i = 0, \dots, n, \quad (1)$$

lorsque la loi des X_i est *uniforme* sur $[0, 1]^d$. L'auteur rappelle les résultats standards sur le contrôle du risque minimax $\inf_{f_n} \sup_{f \in B_{p,q}^s} (E[\|f_n - f\|_p^p])^{1/p}$, où $B_{p,q}^s$ est une boule d'un espace de Besov. Dans le chapitre 3, l'auteur rappelle les liens entre théorie de l'approximation des fonctions, analyse multirésolution et espaces de Besov. Le chapitre 4 est une discussion sur la notion de "frame" d'un espace de Hilbert. Cette notion est liée aux bases de "Needlets" utilisées dans la suite de la thèse.

Le chapitre 5 contient la première contribution originale de la thèse. Il s'agit d'un long article sur l'estimation de la fonction $f : [0, 1]^d \rightarrow \mathbb{R}$ dans le modèle de régression (1) lorsque la loi des X_i n'est plus supposée uniforme. Le but est d'obtenir un estimateur optimal lorsque la fonction de régression est dans une boule Lipschitz $B_{\infty,\infty}^s$ et sous l'hypothèse que la densité (inconnue) des X_i est uniformément minorée sur son support.

Ce problème a été résolu en utilisant des estimateurs par polynômes locaux (voir [3]). Cependant, il n'existait pas de construction d'estimateurs utilisant la structure d'une analyse multirésolution et qui atteigne la vitesse optimale du problème. Rechercher un tel estimateur est extrêmement naturel, pour pouvoir bénéficier de la souplesse qu'offre les bases d'ondelettes dans la mise en œuvre d'estimateurs.

Le premier résultat de ce chapitre est de montrer qu'il est possible de construire un tel estimateur. L'idée de M. Monnier est de découper $[0, 1]^d$ en hypercubes de côté 2^{-j} ; puis sur chaque hypercube, une régression du vecteur des observations $(Y_i)_i$ sur les vecteurs $(\phi_{j,k}(X_i))_i$ est effectuée. Ici les $\phi_{j,k}$ sont les fonctions d'échelle j de l'analyse multirésolution, dont le support intersecte l'hypercube considéré. L'estimateur de f ainsi obtenu,

est combinaison linéaire des $(\phi_{j,k}(X_i))_i$ avec, par construction, des coefficients différents sur chaque hypercube. Le choix de l'indice j dépend de la régularité de f . Une version adaptative de l'estimateur est donc ensuite introduite, en utilisant la méthode de Lepski. Le résultat principal sur cet estimateur adaptatif est que le risque pour la perte L^p décroît en $n^{-s/(2s+d)}$, qui est la vitesse optimale. Il faut souligner que l'idée de la construction de l'estimateur est assez originale. En effet, l'auteur ne cherche pas à directement estimer les coefficients d'ondelette de la fonction f par orthogonalité, ce qui serait l'idée plus naïve (mais qui ne fonctionne pas). Aussi, l'étude de l'estimateur est délicate, en particulier l'auteur doit contrôler la matrice de covariance empirique qui apparaît dans la régression sur les vecteurs $(\phi_{j,k}(X_i))_i$.

L'auteur étend ensuite la construction de l'estimateur au cas où le support de X est inconnu. Il propose ensuite une application de ses méthodes au problème de classification 'avec conditions de marge'. Ceci permet de retrouver les résultats optimaux de Audibert et Tsybakov 07 [1].

Le chapitre 6 de la thèse considère le problème de régression (1) lorsque $f : \mathbb{S}^d \rightarrow \mathbb{R}$ est une fonction définie sur l'hypersphère de dimension d . Les données X_i sont uniformément répartis sur \mathbb{S}^d . Dans ce chapitre, l'auteur introduit un estimateur de f par seuillage de coefficients sur une base $(\psi_{j,\eta})_{j \geq 0, \eta \in \mathbb{Z}_j}$ de $\mathbf{L}^2(\mathbb{S}^d)$. Les "needlets" $\psi_{j,\eta}$ sont des fonctions localisées apparaissant dans une décomposition de type Littlewood-Paley de $\mathbf{L}^2(\mathbb{S}^d)$ (voir Narcowich *et al* 06 [2]).

Le principal résultat de ce chapitre est de montrer que l'estimateur par seuillage dur atteint la vitesse optimale minimax pour toutes les normes de perte \mathbf{L}^p . L'auteur présente aussi une comparaison numérique des estimateurs obtenus en coupant les coefficients de needlets estimés par un seuil déterministe, avec ceux obtenus en prenant un seuil dépendant des données.

Le chapitre 7 contient le troisième travail de la thèse, qui s'inscrit dans le thème du calcul des prix d'options. Considérons $(X_t)_{t \geq 0}$ le prix d'un actif financier au cours du temps. Il est bien connu que le prix d'une option de maturité T et fonction de payoff π s'écrit comme $p(f) := E[\pi(X_T)] = \int \pi(x)q(x)dx$, où l'espérance est calculée sous une certaine probabilité 'risque neutre'. Le but de ce travail est d'estimer la densité q de la loi de X_T sous cette probabilité risque neutre. L'auteur suppose que l'on observe le prix (d'achat et de vente) de différentes options de vente ("put").

Les méthodes de calibrations qui permettent de calculer le prix d'options à partir d'autre prix observés ne sont pas nouvelles (voir les références dans la thèse), mais usuellement, on cherche à calibrer un modèle pour le processus stochastique $(X_t)_{t \geq 0}$ (par exemple sa fonction de volatilité). Il est naturel, mais original, de chercher à estimer directement la densité de X_T .

Ce chapitre est aussi très original par la méthode que l'auteur propose pour résoudre ce problème. En effet, M. Monnier considère, de manière abstraite, l'opérateur qui à une densité associe la famille de tout les prix des "puts" :

$$\gamma : \begin{cases} \mathbf{L}^2[0, B] \rightarrow \mathbf{L}^2[0, B] \\ f \mapsto \gamma(f)(\xi) = \int (\xi - x)^+ f(x) dx. \end{cases}$$

Une étude spectrale fine de l'opérateur compact autoadjoint $\gamma\gamma^*$ de $\mathbf{L}^2[0, B]$ permet de décomposer l'opérateur γ sur une base de fonctions adaptés. Grâce à cette décomposition en valeur singulière (complètement explicite, mais très calculatoire), l'auteur propose

d'estimer la densité, essentiellement en remplaçant γ par un opérateur de rang fini, et en cherchant la densité la plus régulière qui satisfasse la condition d'être compatible avec les prix de puts observés.

Il faut noter qu'il n'est, à priori, pas évident que chercher une décomposition en vecteurs propres de l'opérateur $\gamma\gamma^*$ puisse se faire de manière suffisamment explicite pour être exploitable. Le point de vue de considérer γ est donc à la fois nouveau et fructueux.

Ce travail de thèse contient de nombreux résultats intéressants et difficiles. Cette thèse montre que M. Monnier a une excellente maîtrise des techniques mathématiques en statistiques non paramétriques. La qualité de la rédaction de la thèse est remarquable : l'introduction et les chapitres de rappels sont extrêmement pédagogiques. Les 3 articles (chapitres 5–7) sont rédigés très clairement.

De manière plus exceptionnelle pour une thèse, certaines idées sont très originales (la construction de l'estimateur dans le chapitre 5, ou le chapitre 7 en entier). Ces idées démontrent une réelle maturité mathématique.

A mon avis, la thèse de Jean Baptiste Monnier est d'un niveau excellent, et je recommande sans réserve sa soutenance.

Fait à Evry, le 14 octobre 2011,

Arnaud Gloter
Université d'Evry Val d'Essonne
Département de Mathématiques
23 Boulevard de France
91037 Evry Cedex
Tel. : 01 64 85 35 68
Courriel : arnaud.gloter@univ-evry.fr

Références

- [1] Audibert J.Y. & Tsybakov A., Fast learning rates for plug-in classifiers. *Ann. Stat.* 35 :608–633, 2007.
- [2] Narcowich F., Petrushev P., Ward J., Localized tight frames on spheres, *J. Funct. Anal.* 238 :530–564, 2006.
- [3] Spokoiny V., Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice, *Ann. Stat.* 26 :1356–1378, 1998.

UNIVERSITÉ PARIS DIDEROT - PARIS 7

AVIS DU JURY SUR LA REPRODUCTION DE LA THESE SOUTENUE

Titre de la thèse : Quelques contributions en classification, régression et étude d'un problème inverse en finance

Nom et prénom de l'auteur : Monsieur MONNIER JEAN-BAPTISTE

Membres du jury :	M. ARNAUD GLOTER	Rapporteur du jury
	M. JEAN-YVES AUDIBERT	Membre du jury
	M. FRANCIS BACH	Membre du jury , président
	M. OLEG LEPSKI	Membre du jury
	M. PETER TANKOV	Membre du jury
	Mme DOMINIQUE PICARD	Directeur de thèse

Date de la soutenance : 6 décembre 2011

Reproduction de la thèse soutenue : (rayer les mentions inutiles)

-Thèse pouvant être reproduite en l'état .

~~-Thèse ne pouvant être reproduite.~~

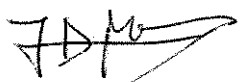
~~-Thèse pouvant être reproduite après corrections suggérées au cours de la soutenance.~~

(Il est recommandé de bien préciser au candidat les corrections qu'il doit apporter à sa thèse, dans un délai qui est réglementairement fixé à trois mois).

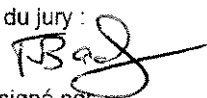
Le diplôme ne lui sera remis que dans la mesure où le président du jury aura constaté que les corrections ont été faites et que l'Institut des études doctorales en aura été informé par courrier.

Vu et pris connaissance le : 6/12/11

L'étudiant :



Le Président du jury :



Imprimé à retourner à l'Institut des études doctorales à l'issue de la soutenance, dûment rempli, daté et signé par l'étudiant et par le président du jury.

PROCES VERBAL DE SOUTENANCE DU 06/12/2011 A 15h00

ANNEE UNIVERSITAIRE 2011/2012

Etudiant : M. JEAN-BAPTISTE MONNIER né le : 23/11/1983

Version de diplôme : DOCTORAT MATHEMATIQUES APPLIQUEES

Titre des travaux : Quelques contributions en classification, régression et étude d'un problème inverse en finance

Secteur de recherche : SCIENCES

Ecole doctorale : SCIENCES MATHEMATIQUES DE PARIS CENTRE

Formation doctorale : MATHEMATIQUES APPLIQUEES

Section CNU : 26 - Mathématiques appliquées et applications

Equipe de Recherche : STATISTIQUE ET MODELES ALEATOIRES

Directeur : Mme DOMINIQUE PICARD

Lieu de soutenance : UNIVERSITE PARIS DIDEROT - PARIS 7 - salle 1C 18 - 175, rue du Chevaleret 75013 Paris

La soutenance est publique.

Résultat : *Admis*

Mention : *TRES HONORABLE*

Membres du Jury

Nom	Qualité	Etablissement	Rôle	Signature
M. ARNAUD GLOTER	<i>Professeur</i> MAITRE DE CONFERENCES	UNIVERSITE D'EVRY	Rapporteur	<i>[Signature]</i>
M. JEAN-YVES AUDIBERT	CHARGE DE RECHERCHE		Membre	<i>[Signature]</i>
M. FRANCIS BACH	DIRECTEUR DE RECHERCHE	I.N.R.I.A	Membre	<i>[Signature]</i> , président
M. OLEG LEPSKI	PROFESSEUR DES UNIVERSITES	UNIVERSITE AIX-MARSEILLE I	Membre	<i>[Signature]</i>
M. PETER TANKOV	PROFESSEUR DES UNIVERSITES	UNIVERSITE PARIS VII	Membre	<i>[Signature]</i>
Mme DOMINIQUE PICARD	PROFESSEUR DES UNIVERSITES	UNIVERSITE PARIS VII	Directeur	<i>[Signature]</i>

rapport après soutenance

Doctorat de Mathématiques Appliquées

Monsieur MONNIER JEAN-BAPTISTE

né le 23 novembre 1983 à PARIS 12 (075)

Titre des travaux : Quelques contributions en classification, régression et étude d'un problème inverse en finance
Date de soutenance : 6 décembre 2011
Etablissement soutenance : Université Paris VII
Jury : M. ARNAUD GLOTER Rapporteur du jury, ~~MAÎTRE DE CONFÉRENCES~~ **PROFESSEUR**, Université d'Evry
M. JEAN-YVES AUDIBERT Membre du jury, CHARGE DE RECHERCHE
M. FRANCIS BACH Membre du jury, DIRECTEUR DE RECHERCHE Institut Nat de Recherche en Informatique et Automatique
M. OLEG LEPSKI Membre du jury, PROFESSEUR DES UNIVERSITES Université d'Aix-Marseille I
M. PETER TANKOV Membre du jury, PROFESSEUR DES UNIVERSITES Université Paris VII
Mme DOMINIQUE PICARD Directeur de thèse, PROFESSEUR DES UNIVERSITES Université Paris VII
Ecole doctorale : SCIENCES MATHÉMATIQUES DE PARIS CENTRE
Spécialité : MATHÉMATIQUES APPLIQUÉES

Jean-Baptiste Monnier nous a présenté son travail de façon dynamique, claire et pédagogique. Il a répondu avec une grande lucidité aux nombreuses questions du jury. Il a montré une grande maîtrise de son travail et de ses perspectives.

Le jury a été impressionné par la qualité des travaux et la variété des thèmes abordés, allant de la statistique non-paramétrique aux mathématiques financières. En effet, en statistique, Jean-Baptiste

N° étudiant : 20905079

UNIVERSITÉ PARIS DIDEROT - PARIS 7

Mannier a apporté des contributions significatives à la fois à la classification, la régression non-paramétrique, et l'estimation sur les variétés. En mathématiques financières, il a proposé une méthode originale, naturelle et pertinente pour la ~~la~~ calibration de la densité usque-neutre.

Le jury est persuadé que Jean-Baptiste Mannier sera un excellent enseignant-chercheur.

M. ARNAUD GLOTER

M. JEAN-YVES AUDIBERT

M. FRANCIS BACH

M. OLEG LEPSKI

M. PETER TANKOV

Mme DOMINIQUE PICARD

sident

D Picard