

Xavier Bry : Montpellier (IMAG) (en collaboration avec Catherine Trottier et Jocelyn Chauvet)

Un outil général de modélisation linéaire généralisée en grande dimension: les composantes supervisées.

L'estimation des modèles linéaires généralisés en grande dimension nécessite leur régularisation. Pour que le prédicteur linéaire soit interprétable, il est en outre indispensable que celui-ci soit relié à un nombre réduit de dimensions d'interprétation *simple*. La pénalisation LASSO permet certes de réaliser régularisation et réduction dimensionnelle, les dimensions trouvées étant les variables sélectionnées. Mais LASSO est instable lorsque les colinéarités sont nombreuses parmi les variables explicatives, et risque de sélectionner, à la place d'une variable véritablement explicative, une variable qui lui est très corrélée. A contrario, nous proposons de conserver a priori toutes les variables explicatives, mais de chercher un nombre restreint de dimensions, appelées *composantes supervisées*, qui soient à la fois très liées aux variables explicatives (donc facilement interprétables) et fortement prédictives des réponses. La recherche de ces dimensions permet également de régulariser le modèle. La méthodologie des composantes supervisées étend, en l'assouplissant, la régression PLS à tous les modèles linéaires généralisés. Les intervalles de confiance bootstrap sur les coefficients permettent une sélection a posteriori des variables participant à la prédiction. Nous présentons l'adaptation de cette méthodologie aux GLMM et au modèle de Cox, en montrant qu'elle constitue un pont pratique entre exploration et modélisation des données.

André Mas: Montpellier (IMAG)

Optimisation par essais particuliers : quelques résultats de convergence en loi.

L'algorithme PSO (Particle Swarm Optimization) est une méta-heuristique stochastique d'optimisation introduite à la fin des années 90, généralement classée dans la famille des algorithmes génétiques. Elle est basée sur un comportement "social" de particules parcourant un domaine sur lequel on cherche à optimiser une fonction de coût. L'allure des trajectoires et le mouvement des particules ont donné leur nom à la méthode. La méthode d'optimisation PSO dispose de certaines caractéristiques : elle s'accommode d'espaces de recherche de grande dimension, ne nécessite pas de calcul de gradient. Mais sa convergence n'est pas assurée et elle convient surtout à des problèmes où les appels à la fonction à optimiser sont peu coûteux (le contraire du Computer Experiment). Dans cet exposé nous nous plaçons dans le cas où PSO converge. Notre objectif est de donner des intervalles de confiance pour le point de convergence de l'algorithme. Nous sommes amenés à considérer deux situations : un cas où les particules oscillent entre un minimum local et le minimum global ; et un second cas où un unique minimum global ou local est détecté, entraînant une convergence à vitesse exponentielle. Nous établissons des théorèmes centraux limites pour chaque particule dans ces deux contextes, l'asymptotique étant liée au nombre d'itérations. Nous obtenons aussi des inégalités de concentration pour tout l'essaim de particules à itération fixée. Les résultats théoriques sont étayés par des simulations.

Camille Mâle : Bordeaux (IMB)

Liberté asymptotique avec amalgamation de grandes matrices aléatoires.

La théorie des probabilités libres a été introduite par Voiculescu dans les années 1980 pour l'étude des algèbres de von Neumann des groupes libres. Les variables aléatoires non commutatives sont des objets abstraits, modélisés par des éléments d'une algèbre commutative équipée d'une forme linéaire. La notion de liberté remplace celle d'indépendance statistique. Les grandes matrices aléatoires jouent un rôle important en probabilités libres, puisque les limites en grande dimension de modèles élémentaires sont libres. Cependant, certains modèles élémentaires nécessitent un cadre étendue pour une analyse non commutative. La théorie des 'trafics' est un cadre mathématique qui complète celui de Voiculescu pour l'étude de ces objets. Elle vient avec une notion d'indépendance qui unifie les différentes notions d'indépendances et libertés non commutatives. Dans cet exposé, je présenterai les premiers éléments analytiques de la théorie avec la notion de liberté sur la diagonale. Ce résultat est le travail d'une collaboration avec B. Au, G. Cébron, A. Dahlqvist et F. Gabriel.

Michel Bonnefont : Bordeaux (IMB)

Estimées de valeurs propres d'ordre supérieure pour des opérateurs de diffusion en dimension 1.

Dans cet exposé, nous proposons une caractérisation variationnelle des valeurs propres d'un opérateur de diffusion en dimension 1. Cette caractérisation généralise celle obtenue par Chen et Wang pour la première valeur propre. La méthode est basée sur une itération des entrelacements entre les gradients et le générateur d'un processus de Markov.

Adrien Richou : Bordeaux (IMB) (Travail en commun avec Jean-François Chassagneux (Paris 7) et Cyril Benezet (Paris 7))

Quelques résultats nouveaux sur les EDSRs obliquement réfléchies et applications à des problèmes de switching randomisés

Dans cet exposé je présenterai quelques résultats nouveaux d'existence et d'unicité pour des équations différentielles stochastiques rétrogrades obliquement réfléchies. Ces équations apparaissent naturellement lorsque l'on traite des problèmes de contrôle stochastique de type switching. Je présenterai également un nouveau problème de contrôle stochastique appelé switching randomisé.

Mathieu Ribatet: Montpellier (IMAG)

Around the structure of max-stable processes

Les processus max-stables jouent un rôle fondamental dans la modélisation spatiale des événements rares, e.g., inondations, vagues de chaleur... Dans cet exposé nous allons repartir de zéro en nous intéressant à leurs représentations spectrales ; représentation qui n'est rien de plus qu'une construction probabiliste simple de cette classe de processus. Par la suite, nous nous intéresserons à la structure particulière induite par cette représentation spectrale, ce qui nous permettra de parler de la difficulté de simuler (conditionnellement) ces processus, d'introduire des mesures (spatiales) de dépendance adaptées mais aussi d'inférence.

Alice Cleynen: Montpellier (IMAG) (en collaboration avec Benoîte de Saporta.)

A Markov model for patient relapse detection

On s'intéresse à un patient en rémission suivi régulièrement (par exemple par une prise de sang de contrôle des marqueurs) pour détecter au plus tôt une éventuelle rechute. La quantité d'intérêt (par exemple le nombre de cellules cancéreuses dans le sang) est un processus continu qui n'est observé que ponctuellement et typiquement au travers d'un proxy (marqueurs). Le médecin souhaite être en mesure de détecter la rechute le plus tôt possible, et si possible

Le type de rechute. Nous proposons un modèle de Markov déterministe par morceau qui se ramène à un problème d'arrêt optimal pour résoudre le problème. Les équations de programmation dynamique obtenues ne sont malheureusement pas calculables numériquement, aussi nous proposons une double discrétisation des processus en Jeu par quantification afin d'établir une stratégie optimale.

Bertrand Cloez: Montpellier (INRA)

Quasi-stationary distribution: convergence and approximations.

Pour un processus de Markov possédant un état absorbant, l'état d'équilibre est trivial. Il est donc plus intéressant de regarder le processus conditionnellement à la non-absorption. Si ce dernier converge en loi vers un équilibre, on appelle cette équilibre une distribution quasi-stationnaire (QSD). Nous verrons dans cet exposé, des critères récents pour l'existence et la convergence vers les QSD (notamment développés par Champagnat et Villemonais) et des algorithmes performants pour son approximation. Ces derniers sont basés sur des méthodes de type descente de gradient stochastique au contraire des systèmes de particules de type Fleming-Viot. Nous illustrerons les résultats avec des exemples de dynamique de populations (naissances, morts...) mais aussi avec des applications en MCMC. En effet, nous verrons que l'usage des QSD permet la simulation exacte à moindre coût.

Mickael Albertus: Toulouse (IMT)

Empirical process with auxiliary information.

Dans ma thèse j'étudie le comportement asymptotique du processus empirique, objet intervenant dans de nombreuses statistiques, quand on lui injecte une information auxiliaire, c'est-à-dire une information que l'on aurait a priori. Cet objet qui est au départ centré subit une modification et n'a plus de raison de le rester. Néanmoins on s'attend à un biais asymptotiquement nul et surtout à une baisse de la variance plus ou moins importante selon l'information apportée. Je présenterai des résultats de convergence faible, d'approximation forte et l'expression des processus limites dans le cas où l'information est donnée par la probabilité des ensembles d'une partition et dans le cas de la connaissance des probabilités de plusieurs partitions (méthode du Raking-Ratio). Je montrerai également de quelle manière il est possible d'appliquer la méthode du bootstrap, méthode statistique de rééchantillonnage, permettant ici de simuler la loi du processus empirique avec informations auxiliaires. Je finirai avec une nouvelle approche de l'information auxiliaire, lorsque celle-ci est donnée par la connaissance d'espérance ou de covariances et non de probabilités.

Lorick Huang: Toulouse (IMT)

Local limit theorem for Robbins-Monro algorithms.

The Robbins-Monro algorithm is a recursive, simulation-based stochastic procedure to approximate the zeros of a function that can be written as an expectation.

It is known that under some technical assumptions, a Gaussian convergence can be established for the procedure.

Here, we are interested in the local limit theorem, that is, quantifying this convergence on the density of the involved objects.

The analysis relies on a parametrix technique for Markov chains converging to diffusions, where the drift is unbounded.

Pascal Bégout: Toulouse (TSE)

On damped second-order gradient systems

Using small deformations of the total energy, as introduced in [31], we establish that damped second order gradient systems $u''(t) + \gamma u'(t) + \nabla G(u(t)) = 0$, Turn MathJax off may be viewed as quasi-gradient systems. In order to study the asymptotic behavior of these systems, we prove that any (nontrivial) desingularizing function appearing in KL inequality satisfies $\phi(s) \geq c s^{\nu}$ whenever the original function is definable and C^2 . Variants to this result are given. These facts are used in turn to prove that a desingularizing function of the potential G also desingularizes the total energy and its deformed versions. Our approach brings forward several results interesting for their own sake: we provide an asymptotic alternative for quasi-gradient systems, either a trajectory converges, or its norm tends to infinity. The convergence rates are also analyzed by an original method based on a one-dimensional worst-case gradient system. We conclude by establishing the convergence of solutions of damped second order systems in various cases including the definable case. The real-analytic case is recovered and some results concerning convex functions are also derived.

Jérôme Bolte: Toulouse (TSE)

A family of functional inequalities: Łojasiewicz inequalities and displacement convex functions

For displacement convex functionals in the probability space equipped with the Monge-Kantorovich metric we prove the equivalence between the gradient and functional type Łojasiewicz inequalities. We also discuss the more general case of λ -convex functions and we provide a general convergence theorem for the corresponding gradient dynamics. Specialising our results to the Boltzmann entropy, we recover Otto-Villani's theorem asserting the equivalence between logarithmic Sobolev and Talagrand's inequalities. The choice of power-type entropies shows a new equivalence between Gagliardo-Nirenberg inequality and a nonlinear Talagrand inequality. Some nonconvex results and other types of equivalences are discussed.