Proposition de parcours MA2 2022-2023

November 23, 2022

Contents

1 Probability and statistics

Ce programme consiste habituellement en deux séries de cours intriquées en probabilités (P) et en statistiques (S), dont le but est de fournir un enseignement solide et interdisciplinaire dans les aspects théoriques modernes des mathématiques de l'aléatoire. Cette année, l'accent est mis sur le volant statistique, et nous nous sommes efforcés de proposer des sujets frontières (probabilités et statistiques, EDP, apprentissage automatique).

Les étudiants feront les 3 cours proposés au premier semestre (P1, S1, S2), et 4 des 5 cours (P2, S3-S6) proposés au second semestre. Il est également possible de suivre les cours d'autres parcours, notamment en EDP.

À noter que cinq des cours orientés “statistiques” (un au premier semestre et trois au second) sont proposés par le parcours de M2 *Maths en action*, qui vient de renouveler son offre de cours.

Enfin, il est très fortement conseillé de suivre les cours de remise à niveau.

2 Refresher Course

Stochastic tools (Grégory Miermont, 15h)

1. Discrete time martingales: stopping theorems and convergence. Extensions for continuous time martingales.

2. Construction of Brownian motion. Regularity of trajectories.

3. Some properties of Brownian trajectories. Connection with the heat equation.

3 Courses

P1: Stochastic calculus (Marielle Simon, 24h)

This lecture series will present some of the most important tools allowing to build and study continuous-time stochastic processes, the central example of which is of course Brownian motion. To this end, we will have to introduce and study semimartingales, a rich class of processes for which one can develop a differential and integral calculus, and set and solve certain type of stochastic differential equations.

Just as for the familiar ordinary differential equations (or PDEs), the motivation to study such stochastic differential equations comes from the goal of understanding the global behaviour of random processes by equations describing their infinitesimal behaviour. But since we are dealing with random processes, these equations contain a random “noise”, which informally is an infinitesimal increment of Brownian motion. The main problem of their study comes from the fact that Brownian motion (and therefore the other processes of interest) have too rough trajectories (nowhere differentiable, for instance) for the usual differential and integral calculus to make sense.

In front of this obstacle, we will develop a notion of stochastic integral, due to Itô. It will give rise to a particular integral calculus, in which Itô’s formula acts as an integration by parts (or a fundamental theorem of analysis) of a new kind. This integral calculus will allow us to study the stochastic differential equations for continuous semimartingales, and will shed a new light on these processes, for instance via Lévy’s characterization of Brownian motion, or the Dubins-Schwarz theorem according to which continuous martingales are appropriate time-changes of Brownian motion. Contents:

1. Generalities on continuous-time processes
2. Continuous-time martingales, regularization. Local martingales, semimartingales. Bracket of a continuous semimartingale.
3. Stochastic integral with respect to a continuous semimartingale.
4. Itô’s formula and applications. The theorems of Lévy, Dubins-Schwarz, Girsanov. Burkholder-Davis-Gundy inequalities.
5. Stochastic differential equations. The Lipschitz case.
6. (Time allowing) Continuous-time Markov processes. Generators. Diffusions.

References
[1] Karatzas-Shreve: Brownian motion and stochastic calculus
[2] Le Gall: Brownian motion and stochastic calculus
[3] Mörters-Peres: Brownian motion
[4] Revuz-Yor: Continuous martingales and Brownian motion
[5] Varadhan: Stochastic processes

S1: Concentration of measure in probability and high-dimensional statistical learning (Guillaume Aubrun, Aurélien Garivier, Rémi Gribonval, 24h)

This course will introduce the notion of concentration of measure and highlight its applications, notably in high dimensional data processing and machine learning. The course will start from deviations inequalities for averages of independent variables, and illustrate their interest for the analysis of random graphs and random projections for dimension reduction. It will then be shown how other high-dimensional random functions concentrate, and what guarantees this concentration yields for randomized algorithms and machine learning procedures to learn from large training collections. This course will be based on a sample of the classical textbook "Concentration Inequalities" by Boucheron, Massart, Lugosi, and on "High-Dimensional Probability" by Roman Vershynin. Applications to machine learning will rely on "Understanding Machine Learning", by Shalev-Schwartz and Ben-David.

S2: Modélisation stochastique et apprentissage statistique (Aurélien Garivier et Clément Marteau, 24h)

[cours proposé également dans le master maths en action]

- Régression en grande dimension: exemples et modélisation, rappels et développement autour du modèle linéaire (modélisation, hypothèses, moindre carrés et vraisemblance, Test de Fisher/Student, etc).
  - Introduction à la sélection de modèles (construction des critères Cp/AIC/BIC, inégalités à la Birgé & Massart, comportement en grande dimension).
  - Méthode Ridge (heuristique, lien avec Tikhonov, propriété du risque, quelques mots sur le choix de lambda par estimation du risque.
  - Introduction au LASSO (construction et heuristiques, lien avec le compressed sensing, propriétés théoriques / inégalités oracles, conditions de compatibilité).
- Classification supervisée: exemples et modélisation, tour d'horizon de quelques algorithmes (kNN, SVM, réseaux de neurones, régression logistique, ...), aspects théoriques (inégalités de concentrations, théorie de Vapnik, noyaux,etc).
- Classification non-supervisée: ACP, Clustering (kmeans, méthodes hiérarchiques,...), Modèles de mélange gaussien, Spectral clustering.
- Autres risques, extrêmes, introductions aux problématiques de recherche.

P2 : Une approche des systèmes désordonnés via les équations aux dérivées partielles (Jean-Christophe Mourrat, 18h)

Le but de la mécanique statistique est de décrire le comportement à grande échelle de collections d'éléments simples, souvent appelés spins, qui interagissent localement selon des règles élémentaires et sont soumis à du désordre. Nous nous concentrerons sur la situation où les interactions locales sont choisies au hasard, auquel cas les modèles sont généralement appelés "verres de spins". De tels modèles sont déjà étonnamment difficiles à analyser lorsque tous les spins interagissent entre eux. Dans ce cours, nous revisiterons cette analyse en utilisant les outils de la théorie des équations de Hamilton-Jacobi.

Après avoir donné un aperçu global du cours, nous commencerons par nous concentrer sur le modèle très simple de Curie-Weiss. Ce sera l'occasion d'introduire les outils analytiques liés à l'étude des équations de Hamilton-Jacobi, que nous utiliserons pour identifier l'énergie libre limite du modèle. Nous transposerons ensuite cette stratégie à une première classe de modèles désordonnés, issus de l'inférence statistique. En termes de difficulté, ces modèles constituent un pont utile entre le modèle de Curie-Weiss et les verres de spins. Nous nous intéresserons enfin à ces derniers modèles, dans lesquels apparaissent des équations de Hamilton-Jacobi de dimension infinie.

S3 : Graphs and ecological networks (Clément Marteau et Thibault Espinasse, 18h)

[cours proposé également dans le master maths en action]

A graph, whose first use are mentionned in the 16th century, is a mathematical object widely used from the first appearance of network investigations, namely investigation of relationship between individual in wide sense. Ranging from social network to the internet, graphs are leading objects for the analysis of several data sets. Ecosystem relationships, from species relationship (prédation, interaction between plants and pollinating insects , etc. . .) social relationship between individuals (sociality between primates, etc. . .), offers several different possible applications of graphs modelling and network investigation.

In this course, we will investigate the framework of graph theory and network science. We will provide an introduction to modern research problems regarding ecosystems studies. We will use alternatively discrete mathematics, statistics and machine learning.

We will adress both theoretical and practical (case studies in ecology) questions.

Theretical keywords: Bases / definitions (graphs, path, etc. . .) – Metrics – Clustering methods – Spectral methods – Random graphs models – Graphical models (graphs inference) – Signal processing on graphs – Multi-level graphs (time, space, link types) – Embedding methods (optional)

Case studies : Contact network between animals. Interaction network between species in a marine and/or alpine environment. Consideration about the relevance of a graph for biodiversity support.

S4: Réseaux de neurones (Aurélien Garivier, 18h)

[cours proposé également dans le master maths en action]

The goal of this course is twofold:

- To present the principles of modern deep neural networks, as well as the technical ways to implement them for solving classification and regression problems.
- To provide a detailed overview of the mathematical foundations of modern learning techniques based on deep neural networks.

Starting with the universal approximation property of neural networks, we will then see why depth improves the capacity of networks to provide accurate function approximations for a given computational budget. Tools to address the optimization problems appearing when training networks on large collections will then be covered, and their convergence properties will be reviewed. Finally, statistical results on the generalization guarantees of deep neural networks will be presented, both in the classical underfitting scenario and in the overfitting scenario leading to the so-called "double descent" phenomenon.

S5: Transport optimal et apprentissage (Philippo Santambrigio, Ivan Gentil, Yohann de Castro, Ievgen Reedko, Julie Digne, Nicolas Bonnel, 18h)

[cours proposé également dans le master maths en action]

Le but du cours est de présenter les grandes lignes de la théorie du transport optimal et certaines de ses applications en sciences des données.

Une première partie du cours détaillera le problème de Monge-Kantorovich, sa formulation comme problème de programmation linéaire et l'utilisation de la dualité convexe, ainsi que les distances (dites distances de Wasserstein) que le transport optimal permet de définir sur l'espace des mesures de probabilité. Les géodésiques et les barycentres dans l'espace de Wasserstein, de grande importance dans l'interpolation et la comparaison des données, seront introduits également.

Une deuxième partie du cours se concentrera sur les méthodes numériques pour la résolution des problèmes de transport optimal, avec une attention particulière aux méthodes les mieux adaptées à la grande dimension et aux données non-structurées, en particulier l'algorithme de Sinkhorn.

Enfin, la troisième partie du cours présentera un choix d'applications du transport et des distances de Wasserstein en apprentissage, dont on cite comme exemples les Wasserstein GANs, l'apprentissage par transfert, les modèles de génération de données,...

S6: Problèmes inverses et parcimonie (Yohann de Castro et Rémi Gribonval, 18h)

[cours proposé également dans le master maths en action]

Sparsity and convexity are ubiquitous notions in Machine Learning and Statistics. In this course, we study the mathematical foundations of some powerful methods based on convex relaxation: L1-regularisation techniques in Statistics and Signal Processing; Nuclear Norm minimization in Matrix Completion; K-means and Graph Clustering.

These approaches turn out to be Semi-Definite representable (SDP) and hence tractable in practice. The theoretical part of the course will focus on the performance guarantees for theses approaches and for the corresponding algorithms under the sparsity assumption. The practical part of this course will present the standard SDP solvers for these learning problems.
Keywords: L1-regularisation; Matrix Completion; K-Means; Graph Clustering; Semi-Definite Programming.