Laboratoire de Matériaux Céramiques et de Mathématiques
DÉPARTEMENT MATHÉMATIQUES

Axe Probabilités et Statistique

Membres de l'axe :

Yacouba Boubacar Mainassara
Mauricio Castano-Aguirre
Nathalie Caouder
Andrés Lopez Lopera
Isabelle Massa-Turpin
Lucas Reding
Eya Zougar

Présentation :

Le Département Mathématiques du CERAMATHS exerce une activité de recherche en statistique et en probabilités appliquées. Les thématiques de recherche de l'axe Probabilités et Statistique portent sur des questions aussi bien théoriques qu'appliquées en modélisation stochastique, en méthodes statistiques et en économétrie. Les objectifs de cet axe consistent à promouvoir les techniques de l'aléatoire pour la modélisation, notamment en lien avec les sciences expérimentales telles que la physique, les sciences du vivant et de l'environnement, la science des données et pour la résolution de problèmes issus du monde économique ou industriel.

Les thèmes de recherche en statistique concernent l'analyse des séries temporelles, l'apprentissage statistique, les modèles de régressio, la modélisation statistique, la statistique non paramétrique, spatiale et spatio-temporelle, avec des applications en machine learning, en intelligence artificielle et en science du vivant. Des recherches interdisciplinaires sont également menées en mécanique, aérodynamique, biologie, neurosciences, météorologie, sociologie, économie, physique, chimie (céramique), environnement, etc.

Les thèmes de recherche en probabilités appliquées concernent les équations différentielles stochastiques, les processus de Markov, les processus gaussiens, la mécanique statistique, les modèles bayésiens, les processus de branchement et la géométrie stochastique.

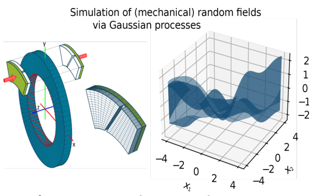
Plus précisément, les principaux sujets abordés par les chercheurs de cet axe sont les suivants

Analyse statistique des séries temporelles

Parmi la grande diversité des modèles stochastiques de séries temporelles à temps discret, on distingue, et on oppose parfois, les modèles linéaires et les modèles non linéaires. En réalité ces deux classes de modèles ne sont pas incompatibles et peuvent même être complémentaires. Les fervents partisans des modèles non linéaires, ou de la prévision non paramétrique, reprochent souvent aux modèles linéaires (par exemple de type ARMA) d'être trop restrictifs, de ne convenir qu'à un petit nombre de séries. Ceci est surtout vrai si on suppose, comme on le fait habituellement, des hypothèses fortes (d'indépendances) sur le bruit qui intervient dans l'écriture du modèle. Francq et Zakoïan (1998) ont montré qu'il existe des processus très variés admettant à la fois des représentations non linéaires et linéaires de type ARMA, pourvu que les hypothèses sur le bruit du modèle ARMA soient suffisamment peu restrictives. Les recherches que nous menons en séries temporelles consistent à revisiter, approfondir et appliquer des résultats de la Statistique moderne en faisant l'hypothèse que les termes d'erreur dans ces modèles sont dépendants (bruits faibles). Les principaux sujets de recherche concernent :

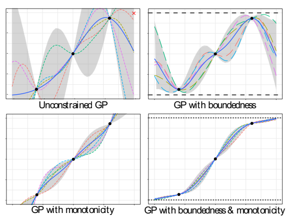
- Les propriétés asymptotiques d'estimateurs (OLS, GLS, MLE, QMLE,) ;
- Les représentations linéaires de processus non linéaires via des modèles ARMA faibles et leurs extensions ;
- Les modèles de séries temporelles à changement de régime Markovien ;
- Les séries temporelles non stationnaires : modèles à coefficients dépendant du temps, modèles à variance dépendant du temps ;
- La modélisation de la volatilité des séries financières (modèles GARCH et leurs extensions) ;
- Les modèles de séries temporelles à valeur entière, processus de Galton Watson.

Exploitation des simulations numériques à l'aide de modèles de substitution probabilistes



De nos jours, l'étude des systèmes dynamiques peut s'appuyer sur des codes de calcul (déterministes) coûteux, ce qui limite l'analyse des simulations numériques « gourmandes » en temps et/ou en ressources. Grâce aux modèles de substitution (méta-modèles) du type « data-driven », et aux avancements récents sur le design d'expérience, des données provenant des simulations peuvent être exploitées. Une exécution du code correspond à une évaluation d'une fonction sous-jacente où les paramètres d'entrée sont les conditions de simulation et les valeurs de sortie sont les quantités d'intérêt du résultat de la simulation. Les méta-modèles sont alimentés par des plans d'expérience « statiquement riches », puis utilisés pour prédire des résultats pour de nouvelles configurations de paramètres. Nous nous intéressons à l'exploitation des méta-modèles probabilistes (par exemple, les processus gaussiens) pour l'apprentissage statistique (machine learning, intelligence artificielle), la quantification d'incertitude en ingénierie (mécanique, aérodynamique) et en sciences de la vie et de la santé (biologie, neurosciences), puis pour l'évaluation et prévention des risques (génie côtier).

Modélisation d'évènements aléatoires liés à des problèmes de régression



Dans ce thème de recherche, nous nous intéressons à établir des cadres (probabilistes) par les processus gaussiens pour des problèmes de régression en ingénierie et en machine learning. Nous considérons le design de nouvelles fonctions de covariances (i.e. noyaux) inspirées par des principes physiques (par exemple, des contraintes d'inégalités, des équations différentielles) ou adaptées à des espaces géométriques d'entrée différents de l'euclidien (par exemple, l'espace des fonctions). Nous étudions également des voies à sorties multiples, multifidélités (ou multi-source), parcimonieuses et variationnelles.

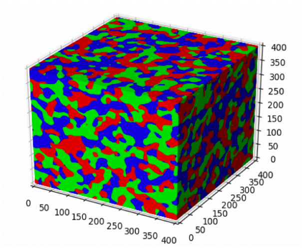
Existence de solutions optimales pour des problèmes de contrôle stochastiques

La théorie du contrôle optimal pour les équations différentielles stochastiques de type Itô est très étroitement liée aux problèmes de valeurs limites à deux points pour les équations différentielles stochastiques, appelées équations différentielles stochastiques progressives et rétrogrades (EDSPR). L'existence de contrôles optimaux donne la résolubilité de certaines EDSPR. D'autre part, certaines idées de la théorie des EDSPR peuvent aider à construire des contrôles optimaux, lorsqu'ils existent. Nous étudions donc les problèmes d'existence de contrôles optimaux stricts ou relaxés sous différentes approches et dans des contextes variés (sauts, observations partielles, réflexion, …).

Equation stochastique de la chaleur à coefficients continus par morceaux

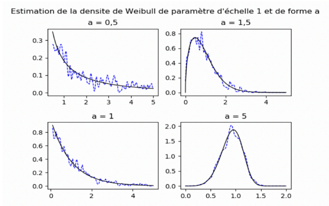
On s'intéresse à l'étude d'une nouvelle équation aux dérivées partielles stochastique avec un opérateur elliptique de second ordre sous forme de divergence, ayant un coefficient de diffusion constant par morceaux et forcée par des différents bruits gaussiens. Une telle équation pourrait être utilisée dans la modélisation mathématique des phénomènes de diffusion dans un milieu constitué par différents types de matériaux et subissant des perturbations stochastiques. On s'intéresse à l'existence de la solution Mild et on cherche les expressions explicites de ses fonctions de covariance et de variance. Certaines propriétés de régularité des trajectoires de la solution sont également analysées. En outre, on étudie les variations quartiques dans le temps et les variations quadratiques dans l'espace de la solution Mild, et, en utilisant le calcul de Stein-Malliavin, on démontre que la séquence de ses variations quadratiques spatiales recentrées et renormalisées satisfait un théorème de limite centrale presque sûr.

Géométrie des champs aléatoires gaussiens sur des espaces stratifiés



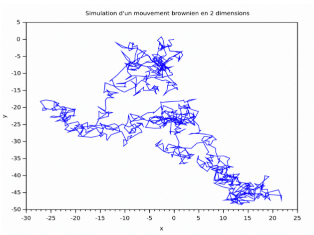
On s'intéresse à la géométrie des phases apparaissant dans un modèle d'excursion de champ gaussien multivarié sur des espaces stratifiés. En supposant que le champ gaussien et l'espace stratifié sous-jacent sont tous deux suffisamment réguliers, les phases du modèle possèdent une bonne régularité et permettent d'entreprendre des calculs de géométrie stochastique. S'intéressant plus particulièrement aux courbures moyennes des phases, on peut mentionner le résultat d'Adler et Taylor (2007) donnant les courbures de Lipschitz-Killing moyennes d'un ensemble d'excursion d'un champ gaussien. Nous sommes cependant intéressés par les mesures de courbure de Lipschitz-Killing moyennes des différentes phases les unes par rapport aux autres. L'obtention de valeurs moyennes pour ces quantités nous permettrait une meilleure compréhension de la géométrie de la micro-structure des matériaux obtenus par frittage tels que les céramiques. Au travers de cet axe de recherche, nous développons aussi un module Python de manipulation et de calcul des courbures de Lipschitz-Killing sur des espaces stratifiés permettant de simplifier l'étude de tels objets.

Estimateurs à noyaux



Dans ce thème de recherche, nous nous intéressons aux propriétés asymptotiques d'estimateurs à noyau pour des données spatiales acquises dans une région finie d'une grille d-dimensionnelle. Nous cherchons à estimer la densité de ces données supposées être identiquement distribuées mais non nécessairement indépendantes. Nous nous intéressons aussi à la question de la régression non-linéaire sous des hypothèses de dépendance similaire. Avec l'avènement de l'informatique, la quantité de données disponibles pour le traitement s'est accrue considérablement ce qui a nécessité le développement d'estimateurs plus performants en terme de calcul : ce sont les estimateurs récursifs de la densité et de la régression. L'étude des propriétés asymptotiques de ceux-ci fait aussi partie de cet axe de recherche.

Théorèmes limite « quenched »



Un problème intéressant, avec de nombreuses applications notamment en mécanique statistique, est celui de l'étude des théorèmes limite pour des processus démarrés à un point fixe. Ce problème est d'autant plus difficile que l'on perd la propriété de stationnarité dès que l'on fixe le point de départ du processus. De plus, comme le montre le contre-exemple formulé par Volný et Woodroofe (2010), l'existence d'un théorème limite « annealed » n'est pas suffisante pour affirmer la convergence en loi du processus démarré à un point fixé. Lorsqu'il y a convergence en loi du processus pour presque tous les points de départ, on parle de théorème limite conditionnel presque sûr ou théorème limite « quenched ». La question du TCL « quenched » a déjà été étudiée en détails ces dernières années cependant de nombreuses questions restent ouvertes à ce sujet.