Stein Variational Gradient Descent: The Original Heuristic and Numerical Experiments

Poster
Tuesday-16:30

Sabino Isaac Cano Paez

Universidad Nacional Autonoma de Mexico (UNAM)

Stein Variational Gradient Descent (SVGD) is a deterministic, gradient-based sampling algorithm widely used in Bayesian learning and other machine learning applications. This poster aims to explain the original heuristic behind SVGD and illustrate its performance through numerical experiments, particularly in the presence of multimodal distributions. Additionally, it will serve as a motivation for exploring alternative theoretical foundations for SVGD, such as its interpretation as a gradient flow in the space of probability measures.

Colloquium
Wed 16:15-17:15

Optimal transport for algorithmic fairness

Arthur Charpentier
Université du Québec à Montréal (UQAM)

Many industries are heavily reliant on predictions of risks based on characteristics of potential customers. Although the use of said models is common, researchers have long pointed out that such practices perpetuate discrimination based on sensitive features such as gender or race. Given that such discrimination can often be attributed to historical data biases, an elimination or at least mitigation, is desirable. With the shift from more traditional models to machine-learning based predictions, calls for greater mitigation have grown anew, as simply excluding sensitive variables in the pricing process can be shown to be ineffective. In the first part of this seminar, we propose to mitigate possible discrimination (related to so call "group fairness", related to discrepancies in score distributions) through the use of Wasserstein barycenters instead of simple scaling. To demonstrate the effects and effectiveness of the approach we employ it on real data and discuss its implications. This part will be based on recent work with François Hu and Philipp Ratz (2310.20508, 2309.06627, 2306.12912 and 2306.10155). In the second part, we will focus on another aspect of discrimination usually called "counterfactual fairness", where the goal is to quantify a potential discrimination "if that person had not been Black" or "if that person had not been a woman". The standard approach, called "ceteris paribus" (everything remains unchanged) is not sufficient to take into account indirect discrimination, and therefore, we consider a "mutates mutants" approach based on optimal transport. With multiple features, optimal transport becomes more challenging and we suggest a sequential approach based on probabilistic graphical models. This part will be based on recent work with Agathe Fernandes Machado and Ewen Gallic (2408.03425 and 2501.15549).

Branching structure and asymptotics of local times of spectrally positive Lévy processes

Weds 12:00-12:30

Jesús Contreras (CIMAT)

Centro de Investigacion en Matemáticas A.C.(CIMAT)

It has been proven recently that an analogue of the so-called Second Ray-Knight theorem holds in the $1+\alpha$-stable case, $\alpha \in (0,1)$. In this setting, the joint distribution of local times as a process indexed by the space variable is that of a non-Markovian branching process and they can be seen as a solution of a stochastic Volterra equation. In this talk we consider the more general case of a spectrally positive Lévy process of unbounded variation. Their local times can also be seen as a branching process and we study their asymptotic behavior, also in the case of the analogue of the first Ray-Knight theorem. This is joint work with Wei Xu from Beijing Institute of Technology.

Tue-10:00-10:30

Centrality and topology properties in a Tree-based Markov random field

Benjamin Côté

University of Waterloo

Abstract: The topology of the tree underlying a tree-structured Markov random field (MRF) is central to the understanding of its stochastic dynamics : it is, after all, what synthesizes the rich dependence relations within the MRF. The aim of this talk is to shed light on the influence of the tree's topology, through an extensive comparison-based analysis, on the aggregate distribution of the MRF. This is done within the framework of a recently introduced family of tree-structured MRFs with the uncommon property of having fixed marginal distributions unaffected by the dependence scheme. We establish convex orderings of sums of MRFs encrypted on trees having different topologies, leading to the devising of a new poset of trees. We present an exhaustive investigation of this new poset. Such an analysis requires, beforehand, to study the joint distribution of a MRF's component and its sum, a dependence relation we refer to as aggregate influence. We employ dependence orders to compare aggregate influences within a tree according to the position of their associated vertices. The resulting orderings are reflected through allocation-related quantities, which thus act as centrality indices. This motivates a discussion on network centrality, a notion stemming from the study of social networks, in the context of MRFs.

Proposals for Stein Operators and Kernel Stein Test for Hypothesis Testing

Poster
Tuesday-16:30

Ana Sarai Davila Martinez

UNAM

This poster is primarily based on the article A Kernel Stein Test for Comparing Latent Variable Models (Kanagawa et al., 2023) and explores proposals to improve the performance of Stein operators in statistical tests utilizing the Kernelized Stein Discrepancy (KSD). The focus is on latent variable models, which present significant statistical and computational challenges due to their complex structures and distributions. To address these challenges, we explore various Stein operators, with a particular focus on the Barker-Stein operator, the Zanella-Stein operator, and operators derived from integrated probability metrics.

The analysis emphasizes computational and simulation aspects, particularly in the robustness of tests in high-dimensional scenarios with low model separation. Additionally, this work compares the model proposed by Kanagawa et al. (2023) with alternative methodologies, such as the Maximum Mean Discrepancy (MMD) and other KSD extensions. The poster highlights the critical importance of selecting suitable Stein operators and other methodological components to achieve robust and reliable hypothesis testing, especially in the context of latent variable models.

Bibliography

- Kanagawa, H., Jitkrittum, W., Mackey, L., Fukumizu, K., & Gretton, A. (2023). A Kernel Stein Test for Comparing Latent Variable Models. arXiv. https://doi.org/10.48550/arXiv.1907.00586.

Poster
Tuesday-16:30

TBA

Sergio Alberto De León Martínez
Centro de Investigación en Matemáticas

Parameter Estimation in Stochastic SIR Models with Lévy Noise and Periodic Transmission Dynamics

12:30-13:00

Terry Easlick
Université de Montréal

This talk focuses on parameter estimation for stochastic SIR models driven by small Lévy noise with periodic transmission functions. We derive least-squares estimators and present results demonstrating their consistency and rate of convergence. These findings provide insights into periodicity in transmission dynamics and its implications for understanding and predicting the spread of infectious diseases.

12:30-13:00

Spectral Graph Analysis, Non-Commutative Probability, and Quantum Decomposition

Tulio Gaxiola

Universidad Autónoma de Sinaloa(UAS)

Spectral graph analysis studies the properties of a graph through the spectrum of associated matrices, such as the adjacency matrix or the Laplacian matrix of the graph. In this talk, we will present the connection between spectral graph analysis and non-commutative probability. In particular, we will discuss the relationship between certain types of graph products and the types of independence studied in non-commutative probability. Additionally, we will explore the quantum decomposition method, which employs quantum principles to decompose linear operators (such as adjacency matrices or Laplacians) associated with a graph.

Bound for the energy of graphs in terms of degrees and leaves

Poster
Tuesday-16:30

Samuel Gurrola Viramontes
Centro de Investigación en Matemáticas (CIMAT)

Graph energy has been a subject of increasing interest due to its connections to various fields, including chemistry, where it is used to model molecular stability, and network theory, where it provides insights into structural properties. A new bound on the energy of graphs in terms of their degrees and the number of leaves adjacent to each vertex was established. Additionally, it explored applications in random graph models like Barabási–Albert trees and Erdős–Rényi graphs.

Tues 9:30-10:30

Uniform spanning trees: theory and applications

Saraí Hernández-Torres
Instituto de MAtemáticas, UNAM (IM-UNAM)

This talk will survey recent progress in our understanding of the geometry of uniform spanning trees. Additionally, we will highlight some applications of the uniform spanning tree model. The talk is based on joint works with Omer Angel, David Croydon, Xinyi Li, and Daisuke Shiraishi

---

Robustness assessment of black-box models to feature perturbations

Thurs 11:30-12:00

Marouane Il Idrissi

Université du Québec à Montréal (UQAM)

Robustness studies of black-box models are recognized as necessary for numerical models based on structural equations and predictive models learned from data. These studies must assess the model's robustness to possible misspecification regarding its inputs (e.g., covariate shift). The study of black-box models, through the prism of uncertainty quantification (UQ), is often based on sensitivity analysis involving a probabilistic structure imposed on the inputs. In contrast, ML models are solely constructed from observed data. This work aims to unify the UQ and ML interpretability approaches by providing relevant and easy-to-use tools for both paradigms. To provide a generic and understandable framework for robustness studies, perturbations of the inputs are defined by constraining their quantiles. The Wasserstein distance between probability measures is used to solve the problem while preserving the inputs' dependence structure. It is demonstrated that this perturbation problem can be analytically solved. Ensuring regularity through isotonic polynomial approximations leads to smoother perturbations, which can be more suitable in practice. Numerical experiments on real case studies highlight the computational feasibility of such studies and provide local and global insights on the robustness of black-box models to input perturbations.

Tues 15:00-16:00

High-Frequency Statistics for Lévy Processes: A Stein's Method Perspective

Arturo Jaramillo Jil

Centro de Investigación en Matemáticas (CIMAT)

The monitoring of complex random systems evolving over time often involves observing a progressively revealed component alongside random perturbations. High-frequency statistics are widely used in fields like financial mathematics to analyze such systems. A common approach to optimal estimation in this context, in the sense of least squares, is through conditional expectation. However, a natural question arises: how effective is conditional expectation as an approximation mechanism? This talk explores this topic by framing the problem as the study of fluctuations of high-frequency statistics around their conditional expectation, when the underlying sources of randomness are Lévy processes within the domain of attraction of a stable process. Using an adapted version of Stein's method for centered mixed Gaussian variables, we establish a non-central distributional limit theorem.

The past light-cone property of the stochastic wave equation with heavy-tailed noise

Thurs 12:30-13:00

Juan Jiminez
University of Ottawa

The study of stochastic partial differential equations (SPDEs) with heavy-tailed noise is partly motivated by the need to explain the random behavior of systems with sudden changes. Heavy-tailed noise is useful for modeling discontinuous phenomena, such as financial crashes, turbulent flows, and some types of cosmic radiation. Classical models like Brownian motion were first used to describe continuous random fluctuations, but they failed to capture systems with large, unpredictable jumps. In this talk, we will prove the existence and uniqueness of a solution to the stochastic wave equation in $\mathbb{R}^d$ (for $d \leq 2$) driven by Lévy noise, which may have infinite moments of all orders. Using the past light-cone property of the wave operator, we show that a unique random field solution exists under the assumption $\int_{|z|\leq 1} |z|^p \nu(dz) < +\infty$ for some $p \in (0, 2)$, where ν is the Lévy measure of the noise. Additionally, this solution has bounded p-moments up to a stopping time related to the light-cone region. Based on: https://doi.org/10.1016/j.spa.2024.104479

Wed 10:00-11:00

GROS: A Unified Framework for Robust Aggregation in Metric Spaces with Applications to Machine Learning and Statistics

Emilien Joly

Centro de Investigación en Matemáticas (CIMAT)

In this talk, I will present GROS (General Robust Aggregation Strategy), a novel framework for robustly combining estimators in metric spaces. GROS is inspired by the median-of-means approach but extends it to a much broader class of problems, including clustering, regression, bandits, set estimation, and topological data analysis. The key idea is simple yet powerful: partition the data into K groups, compute an estimator for each group, and then aggregate these estimators using a robust minimization procedure. The resulting estimator is provably sub-Gaussian and achieves a high breakdown point, making it resilient to outliers and adversarial data. I will also discuss how GROS can be efficiently implemented in practice, with only a constant factor loss in performance compared to the theoretical ideal. Finally, I will outline future directions for applying GROS to other domains where robustness to outliers or adversarial data is critical.

Strong solutions on SDEs with singular (form-bounded) drifts via Röckner-Zhao approach

Tues 11:30-12:00

Kodjo Raphael Madou

University of McGill (McGill)

In this talk, we use the Röckner-Zhao approach to establish strong well-posedness result for SDEs featuring singular drift term, subject to certain minimal assumptions. The talk is based on joint work with Damir Kinzebulatov.

Thurs 15:00-16:00

Towards optimal privacy mechanisms under estimated sensitivity

James Melbourne

Centro de Investigación en Matemáticas (CIMAT)

We first discuss how to maximize the accuracy of responses to a query of fixed sensitivity with an oblivious privacy mechanism that meets a fixed privacy requirement. This problem can be formulated mathematically as minimizing the expected cost of a random variable under the constraint that it possesses a log-Lipschitz density function (with respect to a non-standard metric). Using rearrangement, majorization, and the identification of extreme points we show that the so-called "staircase" distribution is a minimizer in every dimension, confirming a 2015 conjecture of Geng, Kairouz, Oh, and Viswanath. Unfortunately for practitioners, query sensitivity is typically approximate, and unfortunately for the aforementioned result, the staircase distribution is in a sense made quantitative, the most fragile to sensitivity estimation errors. We suggest a well known alternative mechanism and give an optimality result in its support.

This work is joint with Mario Diaz and Shahab Asoodeh.

(Multitype) Λ-Coalescents and Branching Processes

Thurs 16:00-16:30

Imanol Nuñez

Centro de Investigación en Matemáticas (CIMAT)

The relation between λ-coalescents and continuous-state branching processes (CSBPs) has been an important question in understanding genealogies of population models. In this talk, we will first provide an overview of Λ-coalescents and their well-established relation to CSBPs, emphasizing the findings of Caballero, González Casanova, and Pérez (2024), where an explicit homeomorphism is provided between the spaces of Λ-coalescents and CSBPs. We will then discuss the multitype Λ-coalescent proposed by Johnston, Kyprianou, and Rogers (2023). Finally, we will address its relation to the multitype CSBP.

Weds 15:00-16:00

Rates on Yaglom's limit for Galton-Watson processes

Speaker: Sandra Palau

Instituto de Investigación en Matemáticas Aplicadas y en Sistemas (IIMAS-UNAM)

A Galton-Watson process in a varying environment is a discrete time branching process where the offspring distributions vary among generations. It is known that in the critical case, these processes have a Yaglom limit, that is, a suitable normalisation of the process conditioned on non-extinction converges in distribution to a standard exponential random variable. In this talk we present Stein's method to analyse the rate of convergence.

High-dimensional Optimization with Applications to Compute-Optimal Neural Scaling Laws

Thurs 9:00-10:00

Courtney Paquette

University of McGill (McGill)

Given the massive scale of modern ML models, we now only get a single shot to train them effectively. This restricts our ability to test multiple architectures and hyper-parameter configurations. Instead, we need to understand how these models scale, allowing us to experiment with smaller problems and then apply those insights to larger-scale models. In this talk, I will present a framework for analyzing scaling laws in stochastic learning algorithms using a power-law random features model, leveraging high-dimensional probability and random matrix theory. I will then use this scaling law to address the compute-optimal question: How should we choose model size and hyper-parameters to achieve the best possible performance in the most compute-efficient manner?

Tues 11:30-12:30

From magic squares, through random matrices, and to the multiplicative chaos

Elliot Paquette

University of McGill (McGill)

In 2004, motivated by connections of random matrix theory to number theory, Diaconis and Gamburd showed a fascinating connection between the enumeration problem of magic squares (squares filled integers with row and column sum constraints) and the moments of the 'secular coefficients' of random matrices, when the size of the matrix tends to infinity. These are the coefficients in the monomial expansion of a characteristic polynomial, or equivalently, the elementary symmetric polynomials of the eigenvalues of this random matrix.

It turns out that this characteristic polynomial has a limit, when the matrix size tends to infinity. It converges to a random fractal, the holomorphic multiplicative chaos. We describe this process on the unit circle, and show how it can be connected even more strongly to random matrices, and how magic square combinatorics are a type of 'signature' of this holomorphic multiplicative chaos. We'll review some open questions about these objects, and discuss some links between this and other stochastic processes such as the Gaussian multiplicative chaos, the 'circular beta-ensemble' and random multiplicative function.

Fluctuation theory for spectrally negative Lévy processes killed by an additive functional

Wed 9:00-10:00

José Luis Pérez

Centro de Investigación en Matemáticas (CIMAT)

In this talk we solve the exit problems for spectrally negative Lévy processes, which are exponentially killed with a killing intensity dependent on an additive functional of the Lévy process. Additionally, we study their associated resolvents. All identities are given in terms of new generalizations of scale functions. Our results generalize those for omega-killed spectrally negative processes obtained by Palmowski and Li. Finally, we apply these results to derive penalization results for spectrally negative Lévy processes with clocks driven by additive functionals.

This is joint work with Kei Noba, Kouji Yano, and Kohki Ibba.

Weds 10:00-11:00

Optimality of a barrier strategy in a spectrally negative Lévy model with a level-dependent intensity of bankruptcy

Jean-François Renaud

Université du Québec à Montréal (UQAM)

The stochastic control problem concerned with the maximization of dividend payments in a model based on a spectrally negative Lévy process (SNLP) has attracted a lot of research interest since the papers of Avram, Palmowski & Pistorius (2007) and Loeffen (2008). In that problem, a dividend strategy is said to be optimal if it maximises the expected present value of dividend payments made up to the time of ruin, which is a standard first-passage time below zero. In this talk, I will consider a version of this stochastic control problem in which the (controlled) process is allowed to spend time under zero, but is then subject to a level-dependent intensity of bankruptcy. In a joint paper with Dante Mata (UQAM & CRM), we were able to prove that there exists a barrier strategy that is optimal for this control problem, under a mild assumption on the Lévy measure.

Feature extraction in medical image classification using a convolutional neural network-based filter

Poster
Tuesday-16:30

Josué Rodríguez Hernández

Benemérita Universidad Autónoma de Puebla (BUAP)

Medical diagnosis based on medical imaging has revolutionized the early and accurate detection of various diseases, providing an invaluable window into the human body. Despite its effectiveness, the interpretation of these images remains a challenge, with potentially significant implications for the final diagnosis. The introduction of machine learning models, such as Convolutional Neural Networks (CNNs), has enhanced the accuracy and efficiency of this process. However, a major drawback of these models lies in their interpretability, i.e., what criteria or features were considered when making a decision based on the image?

This work proposes a CNN-based image filter that highlights the most relevant features in the images used for classification. By doing so, we aim to identify the most significant regions detected by the model, as well as patterns that indicate the class to which each image belongs. This approach not only seeks to improve diagnostic accuracy but also provides physicians with a reliable support tool, thereby reducing uncertainty in automated diagnoses.

Coauthor(s): De Celis Alonso Benito, Velázquez Castro Jorge

Poster
Tuesday-16:30

TBC

Pedro Leonardo Rodríguez Quintana
Centro de Investigación en Matemáticas (CIMAT)

On the third-order cumulants of products

Tues 16:00-16:30

Ricardo Sayle Sigarreta
Benemérita Universidad Autónoma de Puebla (BUAP)

Free independence, introduced by Voiculescu, extends the classical notion of independence to a broader algebraic framework, with free cumulants serving as multilinear objects that describe this concept. Krawczyk and Speicher addressed the problem of computing cumulants of products in terms of individual cumulants.

The theory of higher-order freeness, an extension of Voiculescu's Free Probability Theory, emerged from studying large random matrices, generalizing properties of first-order cumulants. Under this line, Mingo, Speicher, and Tan later computed second-order cumulants of products. In this work, we extend the previous result to third-order cumulants, including applications to $aa*$ when a is a third-order R-diagonal operator.

Coauthor(s): Dr. Octavio Arizmendi and Dr. Daniel Muñoz.

Poster
Tuesday-16:30

TBA

Hector Manuel Soto Ibarra
Centro de Investigación Matemáticas (CIMAT)

Machine learning
Thurs 12:00-12:30

p-Adic Boltzmann Machines

Wilson Zuñiga-Galindo
University of Texas Rio Grande Valley

The talk will survey our work on the correspondence between p-adic statistical field theories (SFTs) and neural networks (NNs). Hinton et al. constructed deep belief networks (DBNs) by stacking several restricted Boltzmann machines (RBMs). This construction aims to obtain a network with a hierarchical structure (a deep learning architecture). An RBM corresponds to a certain spin glass; we argue that a DBN should correspond to an ultrametric spin glass. A model of such a system can be easily constructed by using p-adic numbers. In our approach, a p-adic SFT corresponds to a p-adic continuous DBN, and a discretization of this theory corresponds to a p-adic discrete DBN. We established that these last machines are universal approximators. In the p-adic framework, the correspondence between SFTs and NNs is not fully developed. There are several open problems.

Bibliography

- Zúñiga-Galindo, Wilson A., A correspondence between deep Boltzmann machines and p -adic statistical field theories, Adv. Theor. Math. Phys. 28 (2024), no. 2, 679–741.
- Zúñiga-Galindo, W. A.; He, C.; Zambrano-Luna, B. A., p -Adic statistical field theory and convolutional deep Boltzmann machines, PTEP. Prog. Theor. Exp. Phys. 2023, no. 6, Paper No. 063A01, 17 pp.
- Zúñiga-Galindo, W. A., p-Adic statistical field theory and deep belief networks, Phys. A 612 (2023), Paper No. 128492, 23 pp.

Poster
Tuesday-16:30

Combinatorics in the Multivariate Markov-Krein transform

Julian Zazueta Obeso
Centro de Investigación en Matemáticas (CIMAT)

The Markov-Krein transform is a bijective transformation that links probability measures with certain signed measures on the real line. Its origins trace back to fundamental work in classical probability, and it was first introduced in 1997 by Sergei Kerov. Recently, this transform has emerged in areas related to non-commutative probability and random matrices, and these developments have led to an extension of the transform to a multivariate version. In this poster, I will provide an overview of what is known about the Markov-Krein transform in its univariate version and explain how the multivariate version has been connected to non-commutative probability, particularly highlighting the relationships between multivariate moments and the different types of cumulants: classical, boolean, and free.

Stochastic modeling: Generalised SIS non-Markovian model with waning immunity

Poster
Tuesday-16:30

Brice Arsene Zotsa Ngoufack

(UQAM)

As observed during the Covid-19 pandemic epidemic, when an individual recovers from a coronavirus infection, the immunity of this individual persists for some time, after which his/her immunity decays progressively. Usually scientists use the SIRS compartmental model to describe this process. This model assumes that once an individual has recovered, his/her immunity persists for some time, after which the individual immediately becomes fully susceptible. Moreover, this model does not take into account the randomness of the decays of immunity after each infection and the variations between individuals. Thus the goal of this presentation is to define a stochastic epidemic model with varying infectivity and with waning immunity, and to study its properties. More precisely, we present a functional law of large numbers when the size of the population tends to infinity. We also present results on the behaviour of the epidemic, more precisely the threshold for the existence of an endemic equilibrium, the stability of disease-free steady state and partial answers for the stability of the endemic equilibrium.