

---

## Mathematics of Machine Learning

(Org: **Ben Adcock** (Simon Fraser University), **Simone Brugiapaglia** (Concordia) and/et **Giuseppe Alessio D’Inverno** (SISSA))

---

---

**MARZIA CREMONA**, Université Laval

*Selection of functional predictors and smooth coefficient estimation for scalar-on-function regression models*

In the framework of scalar-on-function regression models – in which several functional variables are employed to predict a scalar response – we propose a methodology for selecting relevant functional predictors while simultaneously providing accurate smooth (or, more generally, regular) estimates of the functional coefficients.

We suppose that the functional predictors belong to a real separable Hilbert space, while the functional coefficients belong to a specific subspace of this Hilbert space. Such a subspace can be a Reproducing Kernel Hilbert Space (RKHS) to ensure the desired regularity characteristics, such as smoothness or periodicity, for the coefficient estimates. Our procedure, called SOFIA (Scalar-On-Function Integrated Adaptive Lasso), is based on an adaptive penalized least squares algorithm that leverages functional subgradients to efficiently solve the minimization problem.

We demonstrate that the proposed method satisfies the functional oracle property, even when the number of predictors exceeds the sample size. SOFIA’s effectiveness in variable selection and coefficient estimation is evaluated through extensive simulation studies and a real-data application to GDP growth prediction.

Work in collaboration with Hedayat Fathi and Federico Severino.

---

**GIUSEPPE ALESSIO D’INVERNO**, International School for Advanced Studies (SISSA), Trieste, Italy

*Surrogate models for diffusion on graphs via sparse polynomials*

Diffusion kernels over graphs have been widely utilized as effective tools in various applications due to their ability to accurately model the flow of information through nodes and edges. However, there is a notable gap in the literature regarding the development of surrogate models for diffusion processes on graphs. In this work, we fill this gap by proposing sparse polynomial-based surrogate models for parametric diffusion equations on graphs with community structure. In tandem, we provide convergence guarantees for both least squares and compressed sensing-based approximations by showing the holomorphic regularity of parametric solutions to these diffusion equations. Our theoretical findings are accompanied by a series of numerical experiments conducted on both synthetic and real-world graphs that demonstrate the applicability of our methodology.

---

**MEHDI DAGDOUG**, McGill University

*Double Machine Learning for Nonresponse in Surveys*

Predictive models are increasingly integrated into survey strategies, supporting tasks such as model-based estimation, model-assisted estimation, and the treatment of nonresponse through imputation and reweighting. In recent decades, the rise of statistical learning has provided survey statisticians with highly flexible new tools, alongside new theoretical and computational advancements. However, incorporating statistical learning into survey estimation poses challenges for conducting valid inference. In this work, we propose an extension of the Double Machine Learning framework to survey sampling, focusing on the treatment of nonresponse through Augmented Inverse Probability Weighting (AIPW) estimators. We establish that the resulting AIPW estimators are square-root  $n$  consistent and asymptotically normal under realistic rate conditions on the statistical learning algorithms. We further propose a consistent variance estimator, enabling the construction of asymptotically valid confidence intervals. Issues related to model selection and aggregation will also be discussed. Simulation studies demonstrating the strong performance of the proposed methods will be presented. This is a joint work with David Haziza (UOttawa).

---

**SALAH IDBELOUCH**, Polytechnique Montréal

*Separable PGD-Based Solution Approximations of Parametric PDEs Using Physics-Informed Neural Networks*

Solutions of parametric initial and boundary-value problems using deep-learning approaches, such as Deep Operator Networks (DeepONets) or Green Operator Networks (GreenONets) [Aldirany et al., CAMWA, 159, 21-30, 2024] have been proposed in recent years. However, achieving high accuracy in the approximations obtained from these methods often remains a significant challenge. In this work, we consider neural-network formulations based on separable representations of the solutions and train them with PGD-like optimization. PGD-like techniques in deep learning were introduced by Ghnatios and Chinesta [Mathematics, 12, 2365, 2024]. We build on this idea, examine several variants of the alternating training strategy, and evaluate their efficiency in terms of accuracy and computational cost. The methodology is further combined with the multi-level neural-network approach [Aldirany et al., CAME, 419, 116666, 2024] to reduce numerical errors. Numerical examples on representative model problems demonstrate the efficiency of the proposed method.

---

**EMMANUEL LORIN**, Carleton University

*Some recent advances in scientific machine learning for quantum chemistry*

In this talk, I will present some recent trends in scientific machine learning applied to electronic structure calculation and time-dependent quantum dynamics. An application to the interaction of electromagnetic fields with molecules will be discussed.

---

**SINA MOHAMMAD-TAHERI**, Concordia University

*Deep greedy unfolding: sorting out the argsort operator in greedy sparse recovery algorithms*

Recent years have seen a growing interest in “unrolled neural networks” for various signal processing applications. These networks provide model-based architectures that mimic the iterations of an iterative algorithm and, when properly designed, admit recovery guarantees. However, there has been limited work on unrolling greedy and thresholding-based sparse recovery algorithms, such as Orthogonal Matching Pursuit (OMP) and Iterative Hard Thresholding (IHT), and existing efforts often lack full neural network compatibility. The primary challenge arises from the non-differentiable (discontinuous) argsort operator within their iterations, which obstructs gradient-based optimization during training. To address this issue, we approximate argsort operator by a continuous relaxation of it using a proxy called “softsort”. We then demonstrate, both theoretically and numerically, that the differentiable versions of OMP and IHT—termed “Soft-OMP” and “Soft-IHT”—serve as reliable approximations of their original counterparts, with minimal error under suitable conditions on the softsort temperature parameter and the gap between elements in the sorted vector. Finally, implementing these algorithms on neural networks, with weights as trainable parameters, reveals that unrolled Soft-OMP and Soft-IHT effectively capture hidden structures in data, establishing a connection between our approach and weighted sparse recovery.

---

**JUNXI ZHANG**, Concordia University

*Promoting Fairness in Treatment Effect Estimation via Optimal Transport*

Treatment effect estimation lies at the core of many high-stakes decision-making applications, such as precision medicine, policy design, and optimal resource allocation. However, models trained on data with endogenous bias can produce unfair treatment effect estimates across demographic subpopulations, leading to discrimination when these estimates play a key role in decision-making. To address this issue, we employ optimal transport theory to derive treatment effect estimators that satisfy group-wise fairness constraints. We establish the consistency and asymptotic properties of the proposed fair estimator that can be used for conducting inference. Furthermore, we provide a theoretical characterization of the “price of fairness” incurred by incorporating fairness constraints into the estimation process.