



The Hong Kong Polytechnic University Department of Applied Mathematics

Seminar Series on Young Scholars in Optimization and Data Science

Engression: Extrapolation for Nonlinear Regression?

By

Dr. Xinwei SHEN ETH Zürich

Abstract

Extrapolation is crucial in many statistical and machine learning applications, as it is common to encounter test data outside the training support. However, extrapolation is a considerable challenge for nonlinear models. Conventional models typically struggle in this regard: while tree ensembles provide a constant prediction beyond the support, neural network predictions tend to become uncontrollable. This work aims at providing a nonlinear regression methodology whose reliability does not break down immediately at the boundary of the training support. Our primary contribution is a new method called 'engression' which, at its core, is a distributional regression technique for pre-additive noise models, where the noise is added to the covariates before applying a nonlinear transformation. Our experimental results indicate that this model is typically suitable for many real data sets. We show that engression can successfully perform extrapolation under some assumptions such as a strictly monotone function class, whereas traditional regression approaches such as least-squares regression and quantile regression fall short under the same assumptions. We establish the advantages of engression over existing approaches in terms of extrapolation, showing that engression consistently provides a meaningful improvement. Our empirical results, from both simulated and real data, validate these findings, highlighting the effectiveness of the engression method.

Biography

Dr. Xinwei Shen is a postdoctoral researcher at the Seminar for Statistics at ETH Zürich, advised by professors Peter Bühlmann and Nicolai Meinshausen. Previously, she obtained her PhD in the Department of Mathematics at Hong Kong University of Science and Technology in 2022, supervised by professor Tong Zhang. She obtained a Bachelor of Science degree from the Department of Statistics at Fudan University in 2018. Her research interests lie at the interface of statistics and machine learning. Currently, she is working on causality and robustness, distributional and generative modeling, as well as climate applications.

Date: 6 November 2023 (Monday) Time: 16:00-17:00 (Hong Kong Standard Time GMT +8) Venue: Online Talk via Zoom (Meeting ID: 960 8186 1013; Passcode: 1106) Speaker: Dr. Xinwei Shen, ETH Zürich Host: Dr. Ruijian Han, The Hong Kong Polytechnic University Click to join: https://polyu.zoom.us/j/96081861013?pwd=VnZTLzR6VERCellOazdtYW14WTIHUT09



Click to join