



The Hong Kong Polytechnic University Department of Applied Mathematics

Seminar

The distribution of Ridgeless least squares interpolators

By

Dr. Qiyang Han Rutgers University

Abstract

The Ridgeless minimum $\left| \frac{2}{n - 1} \right|$ interpolator in overparametrized linear regression has attracted considerable attention in recent years. While it seems to defy conventional wisdom that overfitting leads to poor prediction, recent research reveals that its norm minimizing property induces an `implicit regularization' that helps prediction in spite of interpolation. This renders the Ridgeless interpolator a theoretically tractable proxy that offers useful insights into the mechanisms of modern machine learning methods.

This talk takes a different perspective that aims at understanding the precise stochastic behavior of the Ridgeless interpolator as a statistical estimator. Specifically, we characterize the distribution of the Ridgeless interpolator in high dimensions, in terms of a Ridge estimator in an associated Gaussian sequence model with positive regularization, which plays the role of the prescribed implicit regularization observed previously in the context of prediction risk. Our distributional characterizations hold for general random designs and extend uniformly to positively regularized Ridge estimators.

As a demonstration of the analytic power of these characterizations, we derive approximate formulae for a general class of weighted $\left| q\right|^{s} = \left(0 < q < infty \right)^{s}$ for Ridge(less) estimators that were previously available only for $\left| eq \right|^{2s}$. Our theory also provides certain further conceptual reconciliation with the conventional wisdom: given any (regular) data covariance, for all but an exponentially small proportion of the signals, a certain amount of regularization in Ridge regression remains beneficial across various statistical tasks including (in-sample) prediction, estimation and inference, as long as the noise level is non-trivial. Surprisingly, optimal tuning can be achieved simultaneously for all the designated statistical tasks by a single generalized or k-fold cross-validation scheme, despite being designed specifically for tuning prediction risk.

Biography

Qiyang Han is an Associate Professor of Statistics at Rutgers University. He received a Ph.D. in Statistics in 2018 from University of Washington under the supervision of Professor Jon A. Wellner. His research expands broadly in mathematical statistics and high dimensional probability, with a particular focus on empirical process theory and its applications to nonparametric and high dimensional statistics. He is a recipient of the NSF CAREER award in 2022, and the Bernoulli Society New Researcher Award in 2023.

Date: 6 September 2023 (Wednesday) Time: 14:00-15:00 (Hong Kong Standard Time GMT +8) Venue: M301 Speaker: Dr. Qiyang Han, Rutgers University Host: Prof. Jian Huang, The Hong Kong Polytechnic University