Spectrum-Aware Debiasing: A modern Inference Framework with Applications to Principal Components Regression

Abstract: Debiasing methodologies have emerged as powerful tools for making statistical inferences in high-dimensional problems. Since its original introduction, the methodology underwent a major development with the introduction of debiasing techniques that adjust for degrees-of-freedom. While overcoming limitations of initial debiasing approaches, this updated method relies on Gaussian/sub-Gaussian tailed designs and independent, identically distributed samples - a key limitation. In this talk, we propose a novel debiasing formula that breaks this barrier by exploiting the spectrum of the sample covariance matrix. Our formula applies to a much broader class of designs, including some heavy-tailed distributions, as well as certain dependent data settings. Our correction term differs significantly from prior work but recovers the Gaussian-based formula as a special case. Notably, our approach does not require estimating the high-dimensional population covariance matrix yet can account for certain classes of dependence among both features and samples. We demonstrate the utility of our method for several statistical inference problems. As a by-product, our work also introduces the first debiased principal component regression estimator with formal guarantees in high dimensions.

Pragya Sur