One-Step Estimation with Scaled Proximal Methods

This talk we study statistical estimators computed using iterative optimization methods that are not run until completion. Classical results on maximum likelihood estimators (MLEs) assert that a one-step estimator (OSE), in which a single Newton-Raphson iteration is performed from a starting point with certain properties, is asymptotically equivalent to the MLE. We further develop these early-stopping results by deriving properties of onestep estimators defined by a single iteration of scaled proximal methods. Our main results show the asymptotic equivalence of the likelihood-based estimator and various one-step estimators defined by scaled proximal methods. By interpreting OSEs as the last of a sequence of iterates, our results provide insight on scaling numerical tolerance with sample size. Our setting contains scaled proximal gradient descent applied to certain composite models as a special case, making our results applicable to many problems of practical interest. Additionally, our results provide support for the utility of the scaled Moreau envelope as a statistical smoother by interpreting scaled proximal descent as a quasi-Newton method applied to the scaled Moreau envelope.