On The Asymptotic Performance of ℓ_q-regularized Least Squares

Arian Maleki
Columbia University

In many application areas ranging from bioinformatics to imaging we are faced with the following question: Can we recover a sparse vector $\beta_o \in \mathbb{R}^p$ from its undersampled set of noisy observations $y \in \mathbb{R}^n$, $y = X\beta + \epsilon$. The last decade has witnessed a surge of algorithms to address this question. One of the most popular algorithms is the ℓ_q-regularized least squares given by the following formulation:

$$\hat{\beta}(\lambda, q) = \arg\min_{\beta} \frac{1}{2}\|y - X\beta\|_2^2 + \lambda\|\beta\|_q^q, \tag{1}$$

where $0 \leq q \leq 2$. Despite the non-convexity of these optimization problems for $0 < q < 1$, they are still appealing for their closer proximity to the "ideal" ℓ_0-regularized least squares. In this talk, we adopt the asymptotic framework $p \to \infty$ and $n/p \to \delta$ and analyze the properties of the global minimizer of (1) under the optimal tuning of the parameter λ. Our goal is to answer the following questions: (i) Do non-convex regularizers outperform convex regularizers? (ii) Does $q = 1$ outperform other convex optimization problems when the vector β_o is sparse?

We discuss both the predictive power and variable selection accuracy of these algorithms. If time permits, we also discuss algorithms that can provably reach to the global minima of the non-convex problems in certain regimes.

This talk is based on a joint work with Haolei Weng, Shuaiwen Wang, and Le Zheng.