Finding Low Rank Solutions, Efficiently and Provably

Tuesday, November 5, 2019
3:00PM–4:00PM
Room 646A PGH

Abstract: A rank-$r$ matrix $X \in \mathbb{R}^{m \times n}$ can be written as a product $U V^T$, where $U \in \mathbb{R}^{m \times r}$ and $V \in \mathbb{R}^{n \times r}$. One could exploit this observation in optimization: e.g., consider the minimization of a convex function $f(X)$ over rank-$r$ matrices, where the set of rank-$r$ matrices is modeled via the factorization $U V^T$. Though such parameterization reduces the number of variables, and is more computationally efficient (of particular interest is the case $r \ll \min\{m, n\}$), it comes at a cost: $f(U V^T)$ becomes a non-convex function w.r.t. $U$ and $V$.

We study such parameterization for optimization of generic convex objectives $f$, and focus on first-order, gradient descent algorithmic solutions. We propose the Bi-Factored Gradient Descent (BFGD) algorithm, an efficient first-order method that operates on the $U, V$ factors. We show that when $f$ is (restricted) smooth, BFGD has local sublinear convergence, and linear convergence when $f$ is both (restricted) smooth and (restricted) strongly convex. For several key applications, we provide simple and efficient initialization schemes that provide approximate solutions good enough for the above convergence results to hold.