On Nonconvex Decentralized Gradient Descent

He Wang

This paper considers an undirected, connected network of agents, where the agents attempt to reach an optimal consensus that minimizes the sum of the local objectives in a distributed way. Such a consensus optimization problem has attracted growing interests in recent years, and a number of distributed algorithms have been developed for convex consensus optimization. Nevertheless, when it comes to the nonconvex case, a limited number of convergence results have been discussed. Motived by this, the authors study the convergence performance of two existing algorithms, DGD and Prox-DGD, for solving nonconvex smooth and nonsmooth problems respectively. In each DGD iteration, every agent locally computes its own gradient and then updates its decision variable by combining a (weighted) average of its neighbors' decisions with the negative gradient step. In each Prox-DGD iteration, every agent locally computes the gradient of the smooth part in its objective function, performs a proximal mapping of the nonsmooth part, and exchanges information with its neighbors. Surprisingly, most appealing properties of both algorithms known in the convex setting remain valid in the nonconvex setting, which allows authors to modify the existing proofs and extend them to the nonconvex case. In particular, the authors show that the iterates of both algorithms can converge to a consensus stationary solution at a sublinear rate with diminishing stepsizes. In addition, they are able to converge to a neighborhood of a consensus stationary solution if an appropriate constant stepsize is applied.