

An ADMM Algorithm for Distributed Sparse Optimal Scoring Classification

Yuanshan Wu

Zhongnan University of Economics and Law
E-mail: shan@whu.edu.cn

Abstract: This talk discusses the classification analysis with huge sample size and high dimensionality. Due to the limited storage of a single machine, the whole dataset are usually stored across multiple ones. It makes the traditional classification methodologies unapplicable as well as introduces additional challenges such as computation complexity and communication cost. We propose a distributed sparse optimal scoring classification based on the alternating direction method of multipliers algorithm. Specifically, by imposing consensus-based constraint, the estimates in each machine are forced to be equal and the optimization problem can be separately performed on each machine, making the parallelized computation feasible. We show that at a linear rate the proposed estimates converge to the global ones which are obtained by assuming a single super-machine could store the full data. The merits of algorithms corroborating the global optimality and convergence are demonstrated through both simulated and real data sets.