Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning.


Summary: Many problems in machine learning and other fields can be (re)formulated as linearly constrained separable convex programs. In most of the cases, there are multiple blocks of variables. However, the traditional alternating direction method (ADM) and its linearized version (LADM, obtained by linearizing the quadratic penalty term) are for the two-block case and cannot be naively generalized to solve the multi-block case. So there is great demand on extending the ADM based methods for the multi-block case. In this paper, we propose LADM with parallel splitting and adaptive penalty (LADMPSAP) to solve multi-block separable convex programs efficiently. When all the component objective functions have bounded subgradients, we obtain convergence results that are stronger than those of ADM and LADM, e.g., allowing the penalty parameter to be unbounded and proving the sufficient and necessary conditions for global convergence. We further propose a simple optimality measure and reveal the convergence rate of LADMPSAP in an ergodic sense. For programs with extra convex set constraints, with refined parameter estimation we devise a practical version of LADMPSAP for faster convergence. Finally, we generalize LADMPSAP to handle programs with more difficult objective functions by linearizing part of the objective function as well. LADMPSAP is particularly suitable for sparse representation and low-rank recovery problems because its subproblems have closed form solutions and the sparsity and low-rankness of the iterates can be preserved during the iteration. It is also highly parallelizable and hence fits for parallel or distributed computing. Numerical experiments testify to the advantages of LADMPSAP in speed and numerical accuracy.

Keywords: convex programs; alternating direction method; linearized alternating direction method; proximal alternating direction method; parallel splitting; adaptive renalty

doi:10.1007/s10994-014-5469-5