JIAJIN LI, University of British Columbia

Unveiling Spurious Stationarity and Hardness Results for Bregman Proximal-Type Algorithms

Bregman proximal-type algorithms, such as mirror descent, are popular in optimization and data science for effectively exploiting problem structures and optimizing them under tailored geometries. However, most of existing convergence results rely on the gradient Lipschitz continuity of the kernel, which unfortunately excludes most commonly used cases, such as the Shannon entropy. In this paper, we reveal a fundamental limitation of these methods: Spurious stationary points inevitably arise when the kernel is not gradient Lipschitz. The existence of these spurious stationary points leads to an algorithm-dependent hardness result: Bregman proximal-type algorithms cannot escape from a spurious stationary point within any finite number of iterations when initialized from that point, even in convex settings. This limitation is discovered through the lack of a well-defined stationarity measure based on Bregman divergence for non-gradient Lipschitz kernels. Although some extensions attempt to address this issue, we demonstrate that they still fail to reliably distinguish between stationary and non-stationary points for such kernels. Our findings underscore the need for new theoretical tools and algorithms in Bregman geometry, paving the way for further research.