

Continuous Time Frank-Wolfe Does Not Zig-Zag, But Multistep Methods Do Not Accelerate

Yifan Sun
Stony Brook University
yifan.sun@stonybrook.edu

The Frank-Wolfe algorithm has regained much interest in its use in structurally constrained machine learning applications. However, one major limitation of the Frank-Wolfe algorithm is the slow local convergence property due to the zig-zagging behavior. We observe that this zig-zagging phenomenon can be viewed as an artifact of discretization, as when the method is viewed as an Euler discretization of a continuous time flow, that flow does not zig-zag. For this reason, we propose multistep Frank-Wolfe variants based on discretizations of the same flow whose truncation errors decay as $O(\Delta^p)$, where p is the method's order. This strategy "stabilizes" the method, and allows tools like line search and momentum to have more benefit. However, in terms of a convergence rate, our result is ultimately negative, suggesting that no Runge-Kutta-type discretization scheme can achieve a better convergence rate than the vanilla Frank-Wolfe method. We believe that this analysis adds to the growing knowledge of flow analysis for optimization methods, and is a cautionary tale on the ultimate usefulness of multistep methods.

Joint work with: Zhaoyue Chen, Mokhwa Lee.