Supplementary Materials for “First-Order Algorithms Converge Faster than $O(1/k)$ on Convex Problems”

Ching-pei Lee [1] **Stephen J. Wright** [1]

A. Proof of Lemma 1

Proof. The proof uses simplified elements of the proofs of Lemmas 2 and 9 of Section 2.2.1 from (Polyak, 1987). Define $s_k := k\Delta_k$ and $u_k := s_k + \sum_{i=k}^{\infty} \Delta_i$. Note that

$$s_{k+1} = (k+1)\Delta_{k+1} \leq k\Delta_k + \Delta_{k+1} \leq s_k + \Delta_k. \tag{1}$$

From (1) we have

$$\begin{aligned} u_{k+1} &= s_{k+1} + \sum_{i=k+1}^{\infty} \Delta_i \leq s_k + \Delta_k + \sum_{i=k+1}^{\infty} \Delta_i \\ &= s_k + \sum_{i=k}^{\infty} \Delta_i = u_k, \end{aligned}$$

so that $\{u_k\}$ is a monotonically decreasing nonnegative sequence. Thus there is $u \geq 0$ such that $u_k \to u$, and since $\lim_{k\to\infty} \sum_{i=k}^{\infty} \Delta_i = 0$, we have $s_k \to u$ also.

Assuming for contradiction that $u > 0$, there exists $k_0 > 0$ such that $s_k \geq u/2 > 0$ for all $k \geq k_0$, so that $\Delta_k \geq u/(2k)$ for all $k \geq k_0$. This contradicts the summability of $\{\Delta_k\}$. Therefore we have $u = 0$, so that $k\Delta_k = s_k \to 0$, proving the result. □

References

Polyak, B. T. *Introduction to Optimization*. Translation Series in Mathematics and Engineering. 1987.

[1]Department of Computer Sciences and Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, Wisconsin, USA. Correspondence to: Ching-pei Lee <ching-pei@cs.wisc.edu>, Stephen J. Wright <swright@cs.wisc.edu>.

Proceedings of the 36^{th} International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019.