Evaluation
Online learning
links with optimization and games
Université Paris–Saclay

dual averaging variants of adagrad

Let $d \geqslant 1$ an integer, $\mathcal{X} \subset \mathbb{R}^d$ a nonempty closed convex set, $(u_t)_{t\geqslant 1}$ a sequnece in $\mathbb{R}^d$, and $L > 0$. Consider

$$x_{t+1} = \operatorname*{arg\,max}_{x\in\mathcal{X}} \left\{ \left\langle \sum_{s=1}^{t} u_s, x \right\rangle - \frac{\sqrt{L^2 + \sum_{s=1}^{t} \|u_s\|_2^2}}{2} \|x\|_2^2 \right\}.$$

1) Derive a general regret bound, and derive a corollary in the case where $\|u_t\|_2 \leqslant L$ for all $t \geqslant 1$.

2) Derive guarantees for nonsmooth convex optimization (similarly to Section 7.3).

3) Derive guarantees for smooth convex optimization (similarly to Section 7.4).

4) Define and analyze a dual averaging variant of AdaGrad-Diagonal.