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

adagrad-diagonal and generalized co-coercivity

Let $d \geqslant 1$ and $A \in \mathbb{R}^{d \times d}$ be positive definite matrix.

Définition. — A map $G : \mathbb{R}^d \to \mathbb{R}^d$ is co-coercive for A if for all $x, x' \in \mathscr{X}$,

$$\langle G(x') - G(x), x' - x \rangle \geqslant \| G(x') - G(x) \|^2_{A^{-1}} .$$

Let $G : \mathbb{R}^d \to \mathbb{R}^d$.

1) Give a characterization of the co-coercivity of G involving $A - 2G$ where here, A denotes the linear map $x \mapsto Ax$.

We now assume that G is co-coercive for A and let x_* be a zero of G.

2) Let $(\gamma_t)_{t \geqslant 0}$ a sequence in $(0, 1)$, $x_0 \in \mathbb{R}^d$ and for $t \geqslant 0$,

$$x_{t+1} = x_t - \gamma_t A^{-1} G(x_t).$$

Establish a guarantee for these iterates.

3) Let $\gamma > 0$, $x_0 \in \mathscr{X}$ and for $t \geqslant 0$,

$$x_{t+1} = \left(x_{t,i} - \frac{\gamma}{\sqrt{\sum_{s=0}^{t} G(x_s)_i^2}} G(x_t)_i \right)_{1 \leqslant i \leqslant d},$$

with convention $0/0 = 0$.

a) Let $T \geqslant 0$. Prove that

$$\sum_{t=0}^{T} \langle G(x_t), x_t - x_* \rangle \leqslant \left(\frac{\max_{0 \leqslant t \leqslant T} \|x_t - x_*\|_\infty^2}{2\gamma} + \gamma \right) \sum_{t=0}^{T} \sqrt{\sum_{t=0}^{T} G(x_t)_i^2}.$$

b) Assume that A is a diagonal matrix and establish a guarantee on $\min_{0 \leqslant t \leqslant T} \|G(x_t)\|_{A^{-1}}$.