FIRST-ORDER REGRET FOR ONLINE CONVEX OPTIMIZATION WITH MEMORY AND ONLINE CONTROL

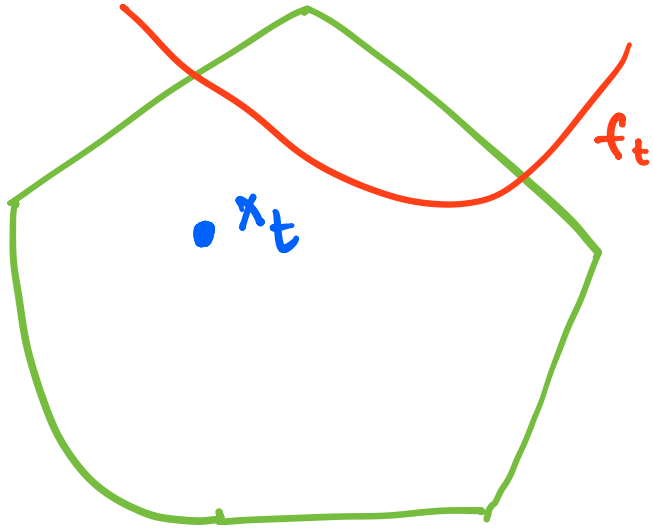
Karan Singh.
CMU
(karansingh@cmu.edu)

NEURIPS 2026.

ONLINE CONVEX OPTIMIZATION

For $t = 1 \ldots T$

1. Learner picks $x_t \in K$.
2. Nature picks $f_t : K \to \mathbb{R}_{\geq 0}$.
3. Learner suffers $f_t(x_t)$.

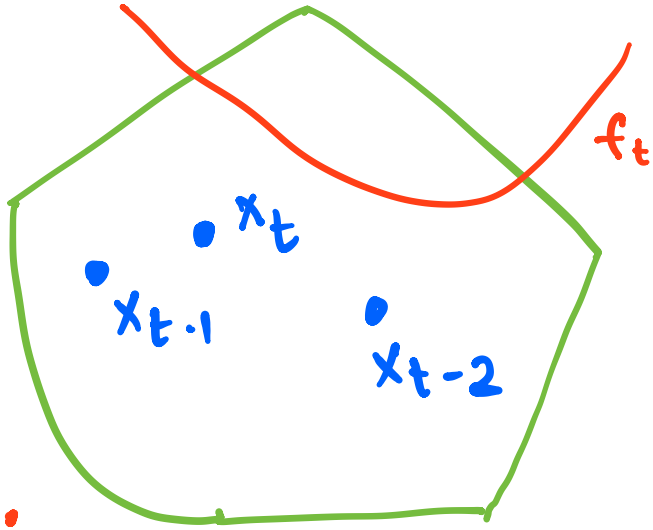


$$\text{REGRET} = \sum_{t=1}^{T} f_t(x_t) - \min_{x \in K} \sum_{t=1}^{T} f_t(x).$$

ONLINE CONVEX OPTIMIZATION WITH MEMORY

For $t = 1 \dots T$

1. Learner picks $x_t \in K$.
2. Nature picks $f_t : K \to \mathbb{R}_{\geq 0}$.
3. Learner suffers $f_t(x_t, \dots x_{t-m})$.



$x_t, \dots x_{t-m}$: m last decisions; m = memory

$$\text{REGRET} = \sum_{t=1}^{T} f_t(x_t, \dots x_{t-m}) - \min_{x \in K} \sum_{t=1}^{T} f_t(\underbrace{x, x \dots x}_{m\text{- times}})$$

ONLINE CONVEX OPTIMIZATION WITH MEMORY

$$\text{REGRET} = \sum_{t=1}^{T} f_t(x_t, \cdots x_{t-m}) - \min_{x \in K} \sum_{t=1}^{T} f_t(\underbrace{x, x \cdots x}_{m\text{- times}})$$

MERHAV WEINBERGER 2002
ARORA DEKEL TEWARI 2012
CESA-BIANCHI DEKEL SHAMIR 2013
ANAVA HAZAN MANNOR 2015

APPLICATIONS

- MDPs with Adv. Rewards
- Online Control
- Statistical Arbitrage
- Online Perform. Prediction

ALSO CAPTURES

SWITCHING COSTS : $f_t(x_t) + \mathbb{1}\{x_t \neq x_{t-1}\}$.

BOUNDED DELAY : $f_t(x_{t-d})$. d = DELAY

COMPOSITE LOSSES : $f_t(x_t) = \sum_{i=0}^{n} f_t^{(i)}(x_t)$.

ONLINE CONVEX OPTIMIZATION WITH MEMORY

$$\text{REGRET} = \sum_{t=1}^{T} f_t(x_t, \dots x_{t-m}) - \min_{x \in K} \sum_{t=1}^{T} f_t(\underbrace{x, x \dots x}_{m\text{- times}})$$

$\sqrt{T}$ Regret $\xrightarrow[\text{PAPER}]{\text{THIS}}$ $\sqrt{L^*}$ REGRET

$$L^* = \min_{x \in K} \sum_{t=1}^{T} f_t(x \dots x).$$

MAIN BARRIER

Existing analyses is strongly tied to "GLOBAL" bounds

ANAVA ET AL → GLOBAL

AGARWAL ET AL → LIPSCHITZ CONSTANT

FIRST-ORER REGRET

SMALL-LOSS BOUND

MAIN RESULTS.

THM 1. For arbitrary f_t's,

as long as $x \longmapsto f_t(x, \ldots x)$ is convex,

sampling-based algorithm

that gets $\sqrt{L^* d}$ - regret.

THM 2. For convex & smooth f_t's,

OGD based algorithm that

gets $\sqrt{L^*}$ - regret.

Thank You.

FIRST-ORDER REGRET FOR OCO WITH MEMORY AND ONLINE CONTROL

Karan Singh.
(karansingh@cmu.edu)

ATLANTA

HALL C1

WED, DEC 9

10 AM - 1 PM