Online Learning in Games

DRAFT

Prof. Volkan Cevher
volkan.cevher@epfl.ch

Lecture 10: Stronger notions of regret

Laboratory for Information and Inference Systems (LIONS)
École Polytechnique Fédérale de Lausanne (EPFL)

EE-735 (Spring 2024)

lions@epfl swisscom E4S EPFL

License Information for Online Learning in Games Slides

Acknowledgements

These slides were originally prepared by Tingting Ni and Anna Maddux.

Introduction

Goals of today

1. Interval regret
   - Weakly adaptive algorithm
   - Strongly adaptive algorithm
2. Dynamic regret

A motivating example

Definition (Classic Regret)

$$R_T = \sum_{t \in T} f_t(w_t) - \min_{w \in \Omega} \sum_{t \in T} f_t(w).$$

- Is the loss of the best fixed point necessarily a good benchmark?

Example (Expert problem with two experts)

Consider two expert, where the loss of expert 1 and expert 2 is given by:

$$l_1(1) = \ldots = l_{\frac{T}{2}} = (1) = 0 \quad \text{and} \quad l_{\frac{T}{2}+1}(1) = \ldots = l_T(1) = 1$$
$$l_1(2) = \ldots = l_{\frac{T}{2}} = (2) = 1 \quad \text{and} \quad l_{\frac{T}{2}+1}(2) = \ldots = l_T(2) = 0$$

- Both experts have loss $T/2$.
- Even if the regret to the best fixed expert is 0, the loss of the algorithm is bounded by $T/2$.
- In a non-stationary environment, classic regret is not the right objective to minimize.

Interval regret

Definition (Interval regret)

$$R_{\mathcal{I}} := \sum_{t\in\mathcal{I}} f_t(w_t) - \min_{w\in\Omega} \sum_{t\in\mathcal{I}} f_t(w), \quad \text{where} \quad \mathcal{I} = [s, e] \subseteq [T].$$

- An algorithm $\mathcal{A}$ with interval regret $R_{\mathcal{I}} = \mathcal{O}(\sqrt{|T|})$ simultaneously $\forall \mathcal{I} \subseteq [T]$ is called a **weakly adaptive** algorithm.

Goal: Design an algorithm with interval regret $R_{\mathcal{I}}(\sqrt{|T|})$.

A motivating example: Follow-up

Recall the example: **"expert problem with two experts"**.

- In the two expert problem, a weakly adaptive algorithm achieves

$$R_{[1,\frac{T}{2}]} = \mathcal{O}(\sqrt{T}) \quad \text{and} \quad R_{[\frac{T}{2}+1,T]} = \mathcal{O}(\sqrt{T})$$

- $\Rightarrow$ The loss of a weakly adaptive algorithm is bounded by $\mathcal{O}(\sqrt{T})$ since:

$$\sum_{t=1}^{T} l_t(i_t) \leq \underbrace{\min_{i\in\{1,2\}} \sum_{t=1}^{T/2} l_t(i)}_{=0} + \mathcal{O}(\sqrt{T}) + \underbrace{\min_{i\in\{1,2\}} \sum_{t=T/2}^{T} l_t(i)}_{=0} + \mathcal{O}(\sqrt{T})$$

A weakly adaptive algorithm in the OCO setting: Online gradient descent(OGD)

Assumption

- *The set* $\Omega \subseteq \mathbb{R}^d$ *is convex with bounded diameter, i.e.,* $\forall y, w \in \Omega \ \|w - y\| \leq D$.
- *The function* $f_t : \Omega \to \mathbb{R}$ *is convex with bounded gradient norm, i.e.,* $\|\Delta f_t(x)\| \leq G \ \forall x \in \Omega$.

Algorithm 1 Online gradient descent algorithm

Require: Step size $\eta_t := \frac{D}{G\sqrt{t}}$.

  for $t = 1, \ldots, T$ **do**

    `Play` w_t `and observe cost` $f_t(w_t)$.

    `Update and project:`

$$
\begin{aligned}
y_{t+1} &= w_t - \eta_t \nabla f_t(w_t) \\
w_{t+1} &= \Pi_\Omega(y_{t+1})
\end{aligned}
$$

  end for

A weakly adaptive algorithm in the OCO setting: Online gradient descent(OGD) I

Theorem

The interval regret of OGD algorithm (Algorithm 1) for any interval $\mathcal{I} \subseteq [T]$ is bounded by:

$$R_{\mathcal{I}} = \mathcal{O}(\sqrt{T}).$$

Proof

For $\mathcal{I} := [s, e] \subseteq [T]$, we define

$$w_I^* := \arg\min_{w \in \Omega} \sum_{t \in \mathcal{I}} f_t(w)$$

Recall the projected gradient decent update step:

$$\begin{aligned} y_{t+1} &= w_t - \eta_t \nabla f_t(w_t) \\ w_{t+1} &= \Pi_\Omega(y_{t+1}). \end{aligned}$$

A weakly adaptive algorithm in the OCO setting: Online gradient descent(OGD) II

Proof

Then, since the projection operator on a convex set Ω is contractive the following holds:

$$\|w_{t+1} - w_I^*\|^2 \leq \|y_{t+1} - w_I^*\|^2 = \|w_t - w_I^*\|^2 + \eta_t^2 \|\nabla f_t(w_t)\|^2 - 2\eta_t (\nabla f_t(w_t))^\top (w_t - w_I^*).$$

Thus,

$$2(\nabla f_t(w_t))^T (w_t - w_I^*) \leq \frac{\|w_t - w_I^*\|^2 - \|w_{t+1} - w_I^*\|^2}{\eta_t} + \eta_t G^2 \qquad \textit{(by assumption } \|\nabla f_t(w_t)\| \leq G\textit{)} \quad (1)$$

A weakly adaptive algorithm in the OCO setting: Online gradient descent(OGD) III

Proof

Summing the inequality across the rounds we get:

$$
\begin{aligned}
2\Big(\sum_{t\in\mathcal{I}} f_t(w_t) - f_t(w_I^*)\Big) &\leq 2\sum_{t\in\mathcal{I}} (\nabla f_t(w_t))^\top (w_t - w_I^*) && \textit{(by convexity of } f_t\textit{)} \\
&= \sum_{t\in\mathcal{I}} \frac{\|w_t - w_I^*\|^2 - \|w_{t+1} - w_I^*\|^2}{\eta_t} + G^2\eta_t && \textit{(by equation (1))} \\
&= \sum_{t\in\mathcal{I}} \|w_t - w_{\mathcal{I}}^*\|^2 \Big(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\Big) + G^2 \sum_{t\in\mathcal{I}} \eta_t && \\
&\leq D^2 \sum_{t\in\mathcal{I}} \Big(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\Big) + G^2 \sum_{t\in\mathcal{I}} \eta_t && \textit{(by } \|x-y\| \leq D \textit{ for all } x, y \in \Omega\textit{)} \\
&= \frac{D^2}{\eta_e} + G^2 \sum_{t\in\mathcal{I}} \eta_t && \textit{(define } \frac{1}{\eta_{s-1}} \geq 0\textit{)}
\end{aligned}
$$

A weakly adaptive algorithm in the OCO setting: Online gradient descent(OGD) IV

Proof

By the construction of step sizes $\eta_t = \frac{D}{G\sqrt{t}}$, *we have*

$$
\begin{aligned}
2\Big(\sum_{t\in\mathcal{I}} f_t(w_t) - f_t(w_I^*)\Big) &\leq \frac{D^2}{\eta_e} + G^2 \sum_{t\in\mathcal{I}} \eta_t \\
&= DG\sqrt{e} + DG\sum_{t=s}^{e} \frac{1}{\sqrt{t}} \\
&\leq DG\sqrt{e} + 2DG\sqrt{e} && \Big(\text{By } \sum_{t=s}^{e} \frac{1}{\sqrt{t}} \leq 2\sqrt{e}\Big) \\
&\leq 3DG\sqrt{T}
\end{aligned}
$$

Interval regret: A strongly adaptive algorithm in the OCO setting

Theorem

The interval regret of OGD algorithm (Algorithm 1) for any interval $\mathcal{I} \subseteq [T]$ is bounded by:

$$R_{\mathcal{I}} = \mathcal{O}(\sqrt{T}).$$

- $\Rightarrow$ The guarantee of OGD is meaningless for intervals of length $o(\sqrt{T})$.
- An algorithm $\mathcal{A}$ with interval regret $R_{\mathcal{I}} = \mathcal{O}(\sqrt{|\mathcal{I}|})$ simultaneously $\forall \mathcal{I} \subseteq [T]$ is called a **strongly adaptive** algorithm.

Goal: Design an algorithm with interval regret $R_{\mathcal{I}}(\sqrt{|\mathcal{I}|})$.

- Any algorithm with (low) regret can be turned into a strongly adaptive algorithm using sleeping experts [**?**].

Detour: Sleeping Experts

The sleeping expert problem

At round $t = 1, \ldots, T$:

1. Environment decides which of the N experts are awake ($a_t(i) = 1$) and which are asleep ($a_t(i) = 0$).
2. a_t is revealed to learner who chooses a distribution $p_t \in \Delta(N)$ such that $p_t(i) = 0$ if $a_t(i) = 0$.
3. Environment reveals loss for awake experts $l_t(i)$ with $a_t(i) = 1$

Remark:

- The regular expert problem is a special case with $a_t(i) = 1$ for all t and i.
- We use the notation $l_t \in [0, 1]^N$, although coordinates of asleep experts are not define.

Definition (Sleeping expert regret)

The regret against a sleeping expert is defined as:

$$R_T(i) = \sum_{t: a_t(i)=1} (\langle p_t, l_t \rangle - l_t(i)).$$

Goal: Design a sleeping expert algorithm by reducing the sleeping expert problem to the regular expert problem.

Detour: Reducing the sleeping expert problem to regular expert problem

Design of a sleeping expert algorithm:

Suppose we are given an expert algorithm $\mathcal{E}$ with prediction $\hat{p}_t$ on round t.

1. Set the prediction p_t for the sleeping expert problem to $p_t(i) \propto a_t(i)\hat{p}_t(i)$.
2. - For the awake experts $(a_t(i) = 1)$ the loss $l_t(i)$ is set to be the same in the sleeping and the regular expert problem.
   - For the asleep experts $(a_t(i) = 0)$ set the loss $l_t(i) = x$ such that the loss of $\mathcal{E}$ for a round is the same as the loss of the sleeping expert algorithm, i.e., set x such that the following equality holds:

$$\sum_{i:a_t(i)=1} \hat{p}_t(i) l_t(i) + \left(\sum_{i:a_t(i)=0} \hat{p}_t(i) \right) x = \sum_{i:a_t(i)=1} p_t(i) l_t(i)$$

$$\Leftrightarrow x = \sum_{i:a_t(i)=1} p_t(i) l_t(i).$$

   - $\Rightarrow$ The "fake" loss of asleep experts is exactly the loss loss of the sleeping expert algorithm.
   - We overload the notation l_t to denote the loss vector of the sleeping and the regular expert problem, where $l_t(i) = \sum_{i:a_t(i)=1} p_t(i) l_t(i) = \langle p_t, l_t \rangle$ if $a_t(i) = 0$.

Detour: Reducing the sleeping expert problem to regular expert problem

Algorithm 2 Reducing from sleeping expert to regular expert

Require: Regular expert algorithm $\mathcal{E}$.
for $t = 1, \ldots, T$ **do**
```
Let p̂_t be the prediction of E on round t.
Observe a_t from the environment.
Play p_t such that p_t(i) ∝ a_t(i)p̂_t(i).
Observe l_t(i) for i such that a_t(i) = 1.
Set l_t(i) = ⟨p_t, l_t⟩ for i such that a_t(i) = 0.
Pass l_t to E.
```
end for

By the reduction, we have:

$$R_T(i) = \sum_{t:a_t(i)=1} (\langle p_t, l_t \rangle - l_t(i)) = \sum_{t=1}^{T} (\langle p_t, l_t \rangle - l_t(i)) = \sum_{t=1}^{T} (\langle \hat{p}_t, l_t \rangle - l_t(i)).$$

- $\Rightarrow$ This is the regret against expert i in a regular expert problem.
- If $\mathcal{E}$ (e.g. Hedge) has regret bound $\mathcal{O}(\sqrt{T \log N})$, then the sleeping expert regret $R_T(i) = \mathcal{O}(\sqrt{T \log N})$.
- Ideally, we want a sleeping expert algorithm with bound $\mathcal{O}(\sqrt{|t : a_t(i) = 1| \log N})$.

Strongly adaptive algorithms via sleeping experts in the OCO setting

Suppose we are given an OCO algorithm $\mathcal{A}$ with regular regret $\mathcal{O}(\sqrt{T})$.

- If $\mathcal{I} = [s, e]$ was known, we could run $\mathcal{A}$ starting at round s.
- **BUT:** We want to consider all intervals $\mathcal{I}$.
- **Idea:** Start a new instance of $\mathcal{A}$ at the beginning of every round and combine predictions from different instances to obtain final prediction.

$\Rightarrow$ This can be captured by the **sleeping expert problem**:

- Each instance of $\mathcal{A}$ is an expert.
- The instance that starts at round t (denoted by $\mathcal{A}_t$) is asleep for the first $t-1$ rounds and awake for the rest of the game.
- The final prediction at round t is the convex combination of the predictions from $\mathcal{A}_1, \ldots, \mathcal{A}_t$ according to the distribution decided by a slepping expert algorithm.

Strongly adaptive algorithms via sleeping experts in the OCO setting

Algorithm 3 Strongly adaptive algorithm via sleeping experts

Require: A OCO algorithm $\mathcal{A}$, a sleeping expert algorithm $\mathcal{S}$.

for $t = 1, \ldots, T$ **do**

`Start a new instance of` $\mathcal{A}$, `called` $\mathcal{A}_t$.

`Obtain predictions from` $\mathcal{A}_1, \ldots, \mathcal{A}_{t-1}$, `denoted by` $w_t^1, \ldots, w_t^t$.

`Pass` a_t `to` $\mathcal{S}$, `where` $a_t(i) = 1$ `for` $i \leq t$ `and` $a_t(i) = 0$ `for` $i > t$.

`Obtain distribution` p_t `from` $\mathcal{S}$.

`Predict` $w_t = \sum_{i=1}^{t} p_t(i) w_t^i$.

`Observe loss function` f_t, `suffer loss` $f_t(w_t)$.

`Pass` f_t `to` $\mathcal{A}_1, \ldots, \mathcal{A}_t$.

`Pass` $l_t^{\mathcal{S}}$ `to` $\mathcal{S}$, `where` $l_t^{\mathcal{S}}(i) = f_t(w_t^i)$ `for` $i \leq t$.

end for

Strongly adaptive algorithm via sleeping experts in the OCO setting

Theorem

The interval regret of Algorithm 3 for any interval $\mathcal{I} \subseteq [T]$ is bounded by:

$$R_{\mathcal{I}} = \mathcal{O}\Big(\sqrt{|\mathcal{I}| \log(T)}\Big) + \mathcal{O}\Big(\sqrt{|\mathcal{I}|}\Big).$$

Proof.

For any $w \in \Omega$ and interval $\mathcal{I} = [s, e]$ it follows:

$$\begin{aligned}
R_{\mathcal{I}} = \sum_{t \in \mathcal{I}} (f_t(w_t)) - f_t(w)) &\leq \sum_{t \in \mathcal{I}} \Big(\sum_{i=1}^{t} p_t(i) f_t(w_t^i) - f_t(w) \Big) && \text{(Jensen's inequality)} \\
&= \sum_{t \in \mathcal{I}} (\langle p_t, l_t^{\mathcal{S}} \rangle - l_t^{\mathcal{S}}(s)) + \sum_{t \in \mathcal{I}} (f_t(w_t^s) - f_t(w)) \\
&= R_e(s) + \mathcal{O}(\sqrt{|\mathcal{I}|}) && \text{(by the guarantee of } \mathcal{A}_s) \\
&= \mathcal{O}(\sqrt{|\mathcal{I}| \log(T)} + \mathcal{O}(\sqrt{|\mathcal{I}|}). && \text{(by the guarantee of } \mathcal{S})
\end{aligned}$$

□

Strongly adaptive algorithm via sleeping experts in the expert setting

Definition (Interval regret in the expert setting)

$$R_{\mathcal{I}} = \sum_{t \in \mathcal{I}} \langle w_t, l_t \rangle - \min_{i \in [N]} l_t(i),$$

where N is the number of experts.

Modify Algorithm 3 such that:

- We are given an expert algorithm $\mathcal{E}$ as input, e.g. Hedge, with regular regret $\mathcal{O}(\sqrt{T \log(N)})$.
- Replace f_t with $l_t \in [0,1]^N$
- Set $w_t^i \in \Delta(N)$ for $i \leq t$.

Theorem

The interval regret of Algorithm 3 for any interval $\mathcal{I} \subseteq [T]$ for the expert setting is bounded by:

$$R_{\mathcal{I}} = \mathcal{O}\Big(\sqrt{|\mathcal{I}| \log(TN)}\Big).$$

Proof.

$$
\begin{aligned}
R_{\mathcal{I}} &= \sum_{t\in\mathcal{I}} \langle w_t, l_t\rangle - l_t(i) = \sum_{t\in\mathcal{I}} \sum_{i=1}^{t} \langle l_t, p_t(i) w_t^i\rangle - l_t(i) \\
&= \sum_{t\in\mathcal{I}} (\langle l_t^{\mathcal{S}}, w_t\rangle - l_t^{\mathcal{S}}(s)) + \sum_{t\in\mathcal{I}} l_t(s) - l_t(i) \\
&= R_e(s) + \mathcal{O}(\sqrt{|\mathcal{I}| \ln N}) && \text{(by the guarantee of } \mathcal{E}_s\text{)} \\
&= \mathcal{O}(\sqrt{|\mathcal{I}| \ln T} + \mathcal{O}(\sqrt{|\mathcal{I}| \ln N}). && \text{(by the guarantee of } \mathcal{S}\text{)} \\
&\leq \mathcal{O}(\sqrt{|\mathcal{I}| \ln NT}) && \text{(by } \sqrt{x} + \sqrt{y} \leq \sqrt{x+y}\text{)}
\end{aligned}
$$

□

Computational efficiency of the strongly adaptive Algorithm 3

- The runtime of Algorithm 3 is $\mathcal{O}(t)$ per iteration $\Rightarrow$ Not efficient!

Goal

Improve run time to $\mathcal{O}(\log(t))$ without sacrificing the regret guarantee.

Approach [?]: Modify Algorithm 3 such that:

- Let $\mathcal{A}_t$ live for $2^{d(t)}$ rounds, where $d(t)$ is the number of 2's in t's prime factorization.
- In other words, $d(t)$ is the largest integer such that $t = b(t) * 2^{d(t)}$ for some (odd) integer $b(t)$.

t:	1	2	3	4	5	6	7	8	9	10
$b(t) * 2^{d(t)}$:	$1 * 2^0$	$1 * 2^1$	$3 * 2^0$	$1 * 2^2$	$5 * 2^0$	$3 * 2^1$	$7 * 2^0$	$1 * 2^3$	$9 * 2^0$	$5 * 2^1$
$d(t)$:	0	1	0	2	0	1	0	3	0	1

Note that:

- At any round t and integer d there is at most one expert with life time 2^d.
- At round t the longest lifetime of any awake expert is bounded by $2^{\lfloor \log_2(t) \rfloor}$.
- Thus, at any round t the total number of awake experts is at most $\lfloor \log_2(t) \rfloor + 1$.

$\Rightarrow$ The runtime of the Algorithm 3 reduces to $\mathcal{O}(\log(t))$ per iteration.

Computational efficiency of the strongly adaptive Algorithm 3 I

Theorem

The interval regret of Algorithm 3 with runtime $\mathcal{O}(\log(t))$ per iteration is bounded by:

$$R_{\mathcal{I}} = \mathcal{O}\Big(\sqrt{|\mathcal{I}|\log(T)}\Big).$$

Proof

Divide the interval $\mathcal{I} = [s, e]$ into several disjoint and consecutive subintervals. Let $\mathcal{I}_m = [s_m, e_m]$ for $m = 1, \ldots, M$ be these subintervals, where $s_1 = s$, $s_m = e_{m-1} + 1$ for $1 < m \leq M$, $e_m = s_m + 2^{d(s_m)} - 1$ for $1 < m \leq M$ and $e_M = e$. Clearly for each $\mathcal{I}_m$ ($m < M$) there is an instance $\mathcal{A}_{s_m}$ that is run solely this interval. Furthermore,

$$s_{m+1} = e_m + 1 = s_m + 2^{d(s_m)} = (b(s_m) + 1)2^{d(s_m)} = \frac{b(s_m) + 1}{2} 2^{d(s_m)+1},$$

where $\frac{b(s_m)+1}{2}$ has to be an integer since $b(s_m)$ is odd. This implies $d(s_m) + 1 \leq d(s_{m+1})$ since by construction $s_{m+1} = b(s_{m+1})2^{d(s_{m+1})}$. This in turn implies $2|\mathcal{I}_m| \leq |\mathcal{I}_{m+1}|$.

Computational efficiency of the strongly adaptive Algorithm 3 II

Proof

$$R_{\mathcal{I}} = \sum_{t\in\mathcal{I}} (f_t(w_t)) - f_t(w)) \le \sum_{t\in\mathcal{I}} \Big(\sum_{i=1}^{t} p_t(i) f_t(w_t^i) - f_t(w) \Big) \qquad \textit{(Jensen's inequality)}$$

$$= \sum_{m=1}^{M} \sum_{t\in\mathcal{I}_m} (\langle p_t, l_t\rangle - l_t(s_m)) + \sum_{m=1}^{M} \sum_{t\in\mathcal{I}_m} (f_t(w_t^{s_m}) - f_t(w))$$

$$= \sum_{m=1}^{M} R_{e_m}(s_m) + \sum_{m=1}^{M} \mathcal{O}(\sqrt{|\mathcal{I}_m|}) = \sum_{m=1}^{M} \mathcal{O}(\sqrt{|\mathcal{I}_m| \log(T)}) \qquad \textit{(by the guarantee of } \mathcal{A} \textit{ and } \mathcal{S}\textit{)}$$

$$\le \sum_{m=1}^{\infty} \mathcal{O}(\sqrt{2^{-m}|\mathcal{I}_0| \log(T)}) = \mathcal{O}(\sqrt{|\mathcal{I}|} \log(T)). \qquad \textit{(using } 2|\mathcal{I}_m| \le |\mathcal{I}_{m+1}|\textit{)}$$

Dynamic regret

Definition (Dynamic regret)

$$R_T(u_1, \ldots, u_T) = \sum_{t \in [T]} f_t(w_t) - \sum_{t \in [T]} f_t(u_t).$$

Remark: How does dynamic regret relate to interval regret?

- Suppose $\sum_{t=2}^{T} \mathbf{1}\{u_t \neq u_{t-1}\} = S - 1$.
- Define $I_i := \{[s_i, e_i] \subseteq [T] \mid \forall t \in [s_i, e_i],\, u_t \texttt{ remains unchanged}\}$, then $[T] := \cup_{i=1}^{S} I_i$.

$$\Rightarrow R_T(u_1, \ldots, u_T) = \sum_{i=1}^{S} \sum_{t \in \mathcal{I}_i} (f_t(w_t) - f_t(u_t)) = \sum_{i=1}^{S} R_{\mathcal{I}_i}.$$

Bounding dynamic regret using interval regret bounds

Assume we have a strongly adaptive algorithm with regret $\mathcal{O}(\sqrt{|\mathcal{I}|\ln T})$ for any interval $\mathcal{I}$, we have

$$
\begin{aligned}
R_T(u_1,\dots,u_T) = \sum_{i=1}^{S} R_{\mathcal{I}_i} &= \mathcal{O}\left(\sum_{i=1}^{S}\sqrt{|\mathcal{I}_i|\ln T}\right) \\
&\overset{\text{Cauchy}-\text{Schwarz}}{\leq} \mathcal{O}\left(\sqrt{S\sum_{i=1}^{S}|\mathcal{I}_i|\ln T}\right) \\
&= \mathcal{O}\left(\sqrt{ST\ln T}\right).
\end{aligned}
$$

Therefore, such regret bound is called switching regret bound and is sublinear in T as long as S is sublinear in T.

Bounding dynamic regret in the OCO setting using online gradient descent(OGD)

Definition (Path length)

Let $\mathcal{P}(u_1, \ldots, u_T)$ be the path length of the comparison sequence defined as:

$$\mathcal{P}(u_1, \ldots, u_T) = \sum_{t=1}^{T-1} \|u_t - u_{t+1}\| + 1.$$

Theorem

The dynamic regret of online gradient descent (Algorithm 1) with step size $\eta \approx \sqrt{\frac{\mathcal{P}(u_1,\ldots,u_T)}{T}}$ *is bounded by:*

$$R_T(u_1, \ldots, u_T) = \mathcal{O}\Big(\sqrt{T\mathcal{P}(u_1, \ldots, u_T)}\Big).$$

Remark:

- For a fixed comparator $u_t = x^*$, the path length is 1.
- $\Rightarrow$ This theorem recovers the $\mathcal{O}(\sqrt{T})$ standard regret bound.

Bounding dynamic regret in the OCO setting using online gradient descent(OGD) I

Proof

Recall the projected gradient decent update step:

$$
\begin{aligned}
y_{t+1} &= x_t - \eta \nabla f_t(x_t) \\
x_{t+1} &= \Pi_\Omega(y_{t+1}).
\end{aligned}
$$

Then, since the projection operator on a convex set Ω is contractive the following holds:

$$\|x_{t+1} - u_t\|^2 \leq \|y_{t+1} - u_t\|^2 = \|x_t - u_t\|^2 + \eta^2 \|\nabla f_t(x_t)\|^2 - 2\eta(\nabla f_t(x_t))^\top (x_t - u_t).$$

Thus,

$$2(\nabla f_t(x_t))^T (x_t - u_t) \leq \frac{\|x_t - u_t\|^2 - \|x_{t+1} - u_t\|^2}{\eta} + \eta G^2 \qquad \text{(by assumption } \|\nabla f_t(x_t)\| \leq G\text{)} \quad (2)$$

Bounding dynamic regret in the OCO setting using online gradient descent(OGD) II

Proof

Summing the inequality across the rounds we get:

$$
\begin{aligned}
2\Big(\sum_{t=1}^{T} f_t(x_t) - f_t(u_t)\Big) &\leq 2\sum_{t=1}^{T}(\nabla f_t(x_t))^\top(x_t - u_t) && \textit{(by convexity of } f_t\textit{)}\\
&= \sum_{t=1}^{T} \frac{\|x_t - u_t\|^2 - \|x_{t+1} - u_t\|^2}{\eta} + \eta G^2 T && \textit{(by equation (2))}\\
&= \frac{1}{\eta}\sum_{t=1}^{T}\Big(\|x_t\|^2 - \|x_{t+1}\|^2 + 2u_t^T(x_{t+1} - x_t)\Big) + \eta G^2 T\\
&\leq \frac{2}{\eta}\left(D^2 + \sum_{t=1}^{T} u_t^T(x_{t+1} - x_t)\right) + \eta G^2 T && \textit{(by } \|x-y\| \leq D \textit{ for all } x, y \in \Omega\textit{)}
\end{aligned}
$$

Bounding dynamic regret in the OCO setting using online gradient descent(OGD) III

Proof

Rearranging the terms in the last inequality, we have

$$
\begin{aligned}
2\Big(\sum_{t=1}^{T} f_t(x_t) - f_t(u_t)\Big) &\leq \frac{2}{\eta}\left(D^2 + \sum_{t=2}^{T} x_t^T(u_{t-1} - u_t) + u_T^T x_{T+1} - u_1^T x_1\right) + \eta G^2 T \\
&\leq \frac{2}{\eta}\left(3D^2 + \sum_{t=2}^{T} D\|u_{t-1} - u_t\|\right) + \eta G^2 T && \textit{(By boundness of the } \Omega\textit{)} \\
&\leq \frac{6D^2}{\eta}\left(1 + \sum_{t=2}^{T} \|u_{t-1} - u_t\|\right) + \eta G^2 T = \frac{6D^2}{\eta}\mathcal{P}(u_1, \ldots, u_T) + \eta G^2 T && \textit{(Set the diameter } D \geq 1\textit{)} \\
&= \mathcal{O}\Big(\sqrt{T\mathcal{P}(u_1, \ldots, u_T)}\Big) && \textit{(Set } \eta = \sqrt{\frac{6D^2\mathcal{P}(u_1, \ldots, u_T)}{G^2 T}}\textit{)}
\end{aligned}
$$

Variation of the loss functions

Example (Expert problem with two experts)

Consider two expert, where the loss of expert 1 and expert 2 is given by:

$$l_{t \texttt{ is an odd number}}(1) = 0.5 + \frac{1}{t} \quad \text{and} \quad l_{t \texttt{ is an even number}}(1) = 0.5,$$
$$l_{t \texttt{ is an odd number}}(2) = 0.5 \quad \text{and} \quad l_{t \texttt{ is an even number}}(2) = 0.5 + \frac{1}{t}.$$

- Choose the competitor u_t as the best decision $w_t^* := \arg\min_{w \in \Omega} f_t(w)$.
- In this case, we have $S = T$ since

$$u_{t \texttt{ is an odd number}} = \texttt{Expert 2},$$
$$u_{t \texttt{ is an even number}} = \texttt{Expert 1}.$$

- The variation of the loss functions is defined as

$$V_T := \sum_{t=2}^{T} \max_{w \in \Omega} |f_t(w) - f_{t-1}(w)| \, .$$

In this case, $V_T = \sum_{t=1}^{T} \frac{1}{t} = \mathcal{O}(\ln T)$.

Bound the dynamic regret by the variation of the loss functions

Theorem ([?])

A strongly adaptive algorithm with $R_{\mathcal{I}} = \mathcal{O}(\sqrt{|\mathcal{I}| \ln T})$ *for any interval* $\mathcal{I}$ *ensures*

$$R_T(w_1^*, \ldots, w_T^*) = \mathcal{O}(T^{\frac{2}{3}}(V_T \ln T)^{\frac{1}{3}}),$$

where $w_t^* := \arg\min_{w \in \Omega} f_t(w)$.

- Combining this regret bound with switching regret bound, we have

$$R_T(w_1^*, \ldots, w_T^*) = \min\left\{\mathcal{O}\left(\sqrt{ST \ln T}\right), \mathcal{O}(T^{\frac{2}{3}}(V_T \ln T)^{\frac{1}{3}})\right\}$$

- Back to the two experts problem, we have sublinear regret bound computed as

$$R_T(w_1^*, \ldots, w_T^*) = \min\left\{\mathcal{O}\left(T\sqrt{\ln T}\right), \mathcal{O}(T^{\frac{2}{3}}(\ln T)^{\frac{2}{3}})\right\}.$$

Proof of the theorem I

Proof

Let $\{\mathcal{I}_i\}_{i=1}^M$ *be **any partition** on the whole game* $[T]$*, and define the variation of interval* $\mathcal{I}_i$ *as*

$$V_{\mathcal{I}_i} = \sum_{t=s_i+1}^{e_i} \max_{w\in\Omega} |f_t(w) - f_{t-1}(w)|.$$

Therefore, we bound the regret as

$$\begin{aligned}
R_T &= \sum_{i=1}^M \sum_{t\in\mathcal{I}_i} (f_t(w_t) - f_t(w_t^*)) = \sum_{i=1}^M \sum_{t\in\mathcal{I}_i} \left(f_t(w_t) - f_t(w_{s_i}^*)\right) + \sum_{i=1}^M \sum_{t\in\mathcal{I}_i} \left(f_t(w_{s_i}^*) - f_t(w_t^*)\right) \\
&\leq \sum_{i=1}^M \mathcal{O}(\sqrt{|\mathcal{I}_i| \ln T}) + \sum_{i=1}^M 2|\mathcal{I}_i| V_{\mathcal{I}_i} \overset{\text{Cauchy}-\text{Schwarz}}{\leq} \mathcal{O}\left(\sqrt{M \sum_{i=1}^M |\mathcal{I}_i| \ln T}\right) + 2\max_i |\mathcal{I}_i| \sum_{i=1}^M V_{\mathcal{I}_i} \\
&= \mathcal{O}(\sqrt{MT\ln T}) + 2\max_i |\mathcal{I}_i| V_T.
\end{aligned}$$

Proof of the theorem II

Proof

Find a partition that minimize $\mathcal{O}(\sqrt{MT\ln T}) + 2\max_i |\mathcal{I}_i| V_T$ *for a fixed* M*, then we choose equal distribution such that* $\max_i |\mathcal{I}_i| = \mathcal{O}(\frac{T}{M})$*. Then*

$$
\begin{aligned}
R_T(w_1^*, \ldots, w_T^*) &= \mathcal{O}(\sqrt{MT\ln T}) + \frac{2T}{M} V_T \\
&= \mathcal{O}(T^{\frac{2}{3}} (V_T \ln T)^{\frac{1}{3}}) \qquad \left(\textit{set } M = \left(\frac{2\sqrt{T} V_T}{\sqrt{\ln T}}\right)^{\frac{2}{3}}\right)
\end{aligned}
$$

Proof of the red part

$$\sum_{t \in \mathcal{I}_i} f_t(w_{s_i}^*) - f_t(w_t^*) \leq 2|\mathcal{I}_i| V_{\mathcal{I}_i}$$

Proof.

$$\begin{aligned}
f_t(w_{s_i}^*) - f_t(w_t^*) &\leq f_t(w_{s_i}^*) - f_{s_i}(w_{s_i}^*) + f_{s_i}(w_t^*) - f_t(w_t^*) \quad (\texttt{by optimaility of } w_{s_i}^*) \\
&= \sum_{\tau=s_i+1}^{t} \left(f_\tau(w_{s_i}^*) - f_{\tau-1}(w_{s_i}^*)\right) + \sum_{\tau=s_i+1}^{t} (f_{\tau-1}(w_t^*) - f_\tau(w_t^*)) \\
&\leq 2 \sum_{\tau=s_i+1}^{t} \max_{w \in \Omega} |f_{\tau-1}(w) - f_\tau(w)| \\
&\leq 2V_{\mathcal{I}_i}.
\end{aligned}$$

□

Bound the dynamic regret for the expert problem

- For the expert problem setting, the dynamic regret against a competitor sequence $u_1, \ldots, u_T \in \Delta(N)$ can be written as

$$R_T(u_1, \ldots, u_T) = \sum_{t=1}^{T} \sum_{i=1}^{N} u_t(i) r_t(i),$$

  where $r_t(i) = \langle p_t, l_t \rangle - l_t(i)$ is the instantaneous regret.
- Similarly, we define the variation of the competitors for the expert problem setting as

$$A_T := \sum_{t=1}^{T} \sum_{i=1}^{N} [u_t(i) - u_{t-1}(i)]_+$$

  Instead of variation of the loss function as introduced before which is

$$V_T := \sum_{t=2}^{T} \max_{i \in [N]} |l_t(i) - l_{t-1}(i)|.$$

- A_T and V_T are in general distinct.

Relation between A_T and V_T: example of $A_T = \Omega(T)$ and $V_T = \mathcal{O}(1)$

Example (Expert problem with two experts)

Consider two expert, where the loss of expert 1 and expert 2 is given by:

$$l_{t \texttt{ is an odd number}}(1) = 0.5 + \frac{1}{t^2} \quad \text{and} \quad l_{t \texttt{ is an even number}}(1) = 0.5,$$

$$l_{t \texttt{ is an odd number}}(2) = 0.5 \quad \text{and} \quad l_{t \texttt{ is an even number}}(2) = 0.5 + \frac{1}{t^2}.$$

- Choose the competitor u_t as the best decision

$$u_t(1) = \begin{cases} 0, & t \text{ is an odd number} \\ 1, & t \text{ is an even number} \end{cases}$$

- In this case, we have $A_T = \Omega(T)$ since we change the best expert every round. But the variation of the loss functions is

$$V_T = \sum_{t=1}^{T} \frac{1}{t^2} = \mathcal{O}(1).$$

- Variation of competitors can be large while variation of loss remains small.

Relation between A_T and V_T: example of $A_T = \mathcal{O}(1)$ and $V_T = \Omega(T)$

Example (Expert problem with two experts)

Consider two expert, where the loss of expert 1 and expert 2 is given by:

$$l_t(1) = 0.5 \quad \text{and} \quad l_t(2) = 0.5 + 0.1t.$$

- Choose the competitor u_t as the best decision which is $u_t(1) = 1$ for $t \geq 0$.
- In this case, we have $A_T = \mathcal{O}(1)$ since we never change the best expert every round. But the variation of the loss functions is

$$V_T = \sum_{t=1}^{T} 0.1 = \Omega(T).$$

- Variation of loss can be large while variation of competitors remains small.

Dynamic regret bound for the expert problem using A_T

Theorem ([?])

For the expert problem, a strongly adaptive algorithm with $R_{\mathcal{I}} = \mathcal{O}(\sqrt{|\mathcal{I}| \ln NT})$ *for any interval* $\mathcal{I}$ *ensures*

$$R_T(u_1, \ldots, u_T) = \mathcal{O}(\sqrt{TA_T \ln NT})$$

for any competitor sequence $u_1, \ldots, u_T \in \Delta(N)$.

- The regret is sublinear when A_T is sublinear.
- Combining this regret bound with former results, we have

$$R_T(w_1^*, \ldots, w_T^*) = \min\left\{\mathcal{O}\left(\sqrt{ST \ln NT}\right), \mathcal{O}\left(T^{\frac{2}{3}}(V_T \ln NT)^{\frac{1}{3}}\right), \mathcal{O}\left(\sqrt{TA_T \ln NT}\right)\right\}$$

Proof (Proof of the theorem 18)

Fix an expert i*, and let* $\alpha_m > 0$ *and a partition* $\{\mathcal{I}_m\}_{m\in[M]}$ *of* $[T]$ *such that* $u_t(i) = \sum_{m=1}^{M} \mathbf{1}\{t \in \mathcal{I}_m\}\alpha_m$ *for any* t*. Therefore,*

$$
\begin{aligned}
&R_T(u_1, \ldots, u_T) \\
&= \sum_{i=1}^{N}\sum_{t=1}^{T} u_t(i) r_t(i) = \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{m=1}^{M} \mathbf{1}\{t \in \mathcal{I}_m\}\alpha_m r_t(i) = \sum_{i=1}^{N}\sum_{m=1}^{M} \alpha_m \sum_{t=1}^{T} \mathbf{1}\{t \in \mathcal{I}_m\} r_t(i) \\
&= \sum_{i=1}^{N}\sum_{m=1}^{M} \alpha_m R_{\mathcal{I}_m}(i) \leq \sum_{i=1}^{N}\sum_{m=1}^{M} \alpha_m \sqrt{|\mathcal{I}_m| \ln(NT)} \\
&\overset{\text{Cauchy-Schwartz}}{\leq} \sum_{i=1}^{N} \mathcal{O}\left(\sqrt{\sum_{i=1}^{M} \alpha_i} \sqrt{\sum_{i=1}^{M} \alpha_i |\mathcal{I}_m| \ln(NT)}\right) \\
&= \sum_{i=1}^{N} \mathcal{O}\left(\sqrt{\sum_{i=1}^{M} \alpha_i} \sqrt{\sum_{i=1}^{M} \alpha_i \sum_{t=1}^{T} \mathbf{1}\{t \in \mathcal{I}_m\} \ln(NT)}\right) = \sum_{i=1}^{N} \mathcal{O}\left(\sqrt{\sum_{i=1}^{M} \alpha_i} \sqrt{\sum_{t=1}^{T} u_t(i) \ln(NT)}\right).
\end{aligned}
$$

Proof (Proof of the theorem 18)

If $\sum_{i=1}^{M} \alpha_i = \sum_{t=1}^{T} [u_t(i) - u_{t-1}(i)]_+$, *then we can bound the dynamic regret as*

$$
\begin{aligned}
R_T(u_1, \ldots, u_T) &\leq \sum_{i=1}^{N} \mathcal{O}\left(\sqrt{\sum_{i=1}^{M} \alpha_i} \sqrt{\sum_{t=1}^{T} u_t(i) \ln(NT)} \right) \\
&= \sum_{i=1}^{N} \mathcal{O}\left(\sqrt{\sum_{t=1}^{T} [u_t(i) - u_{t-1}(i)]_+} \sqrt{\sum_{t=1}^{T} u_t(i) \ln(NT)} \right) \\
&\overset{\text{Cauchy}-\text{Schwartz}}{\leq} \mathcal{O}\left(\sqrt{\left(\sum_{i=1}^{N} \sum_{t=1}^{T} [u_t(i) - u_{t-1}(i)]_+ \right) \left(\sum_{i=1}^{N} \sum_{t=1}^{T} u_t(i) \ln(NT) \right)} \right) \\
&= \mathcal{O}(\sqrt{T A_T \ln(NT)}).
\end{aligned}
$$

Construction of $\alpha_t(i)$

Goal

Construct $\alpha_t(i)$ such that

$$\sum_{i=1}^{M} \alpha_i = \sum_{t=1}^{T} [u_t(i) - u_{t-1}(i)]_+.$$

- Set $t^* \in \arg\min_t u_t$ and create an interval $\mathcal{I}_1$ with $\alpha_1 = u_{t^*}$
- Recursively performing the same construction for the inputs $u_1 - u_{t^*}, \ldots, u_{t^*-1} - u_{t^*}$ and $u_{t^*+1} - u_{t^*}, \ldots, u_T - u_{t^*}$ respectively until there are no non-zero inputs left.

Here, we denote the sum of the weights of the above construction as $h(u_1, \ldots, u_T)$ and prove

$$h(u_1, \ldots, u_T) := \sum_{t=1}^{T} [u_t(i) - u_{t-1}(i)]_+.$$

Proof for such construction

Proof.

Base case: $T = 1$ holds trivially.
For $T \geq 2$**:** we have

$$
\begin{aligned}
h(u_1, \ldots, u_T) &= u_{t^*} + h(u_1 - u_{t^*}, \ldots, u_{t^*-1} - u_{t^*}) + h(u_{t^*+1} - u_{t^*}, \ldots, u_T - u_{t^*}) \\
&= u_{t^*} + (u_1 - u_{t^*}) + \sum_{t=2}^{t^*-1} [u_t - u_{t-1}]_+ + [u_{t^*+1} - u_{t^*}]_+ + \sum_{t=t^*+2}^{T} [u_t - u_{t-1}]_+ \\
&= u_1 + \sum_{t=2}^{t^*-1} [u_t - u_{t-1}]_+ + [u_{t^*+1} - u_{t^*}]_+ + \sum_{t=t^*+2}^{T} [u_t - u_{t-1}]_+ \\
&= \sum_{t=1}^{T} [u_t - u_{t-1}]_+ .
\end{aligned}
$$

$\square$

References I