# Evaluation of Online learning links with optimization and games

Statement :
https://joon-kwon.github.io/regret-ups/evaluation/adagrad-approachability.pdf

Lucas Biéchy

Department of Mathematics, University of Paris-Saclay, Orsay Mathematics Laboratory
*Email address:* lucas.biechy@universite-paris-saclay.fr

GitHub repository : https://github.com/Biechy/OnlineLearning

Let $d \geq 1$ and an outcome function $g : \mathcal{A} \times \mathcal{B} \to \mathbb{R}^d$ with $\mathcal{A}, \mathcal{B}$ two action sets. Let $\mathcal{C} \subset \mathbb{R}^d$ be a convex cone satisfying Blackwell's condition with respect to $g$ and $\alpha : \mathcal{C}^\circ \to \mathcal{A}$ be an assiociated oracle satisfying

$$x' = \lambda x \quad \text{for some } \lambda > 0 \quad \Rightarrow \quad \alpha(x) = \alpha(x')$$

Assume an adaptation of AdaGrad-Norm for this approachability problem as follows. Let $(b_t)_{t\geq 0}$ a sequence in $\mathcal{B}, x_0 = 0, a_0 = \alpha(x_0) = \alpha(0)$ and for $t \geq 0$

$$x_{t+1} = \Pi_{\mathcal{C}^\circ}\left(x_t + \frac{\gamma}{\sqrt{\sum_{s=0}^t \|r_s\|_2^2}} r_t\right), \quad a_{t+1} = \alpha(x_{t+1})$$

where $\gamma > 0$, $\Pi_{\mathcal{C}^\circ}$ denotes the Euclidean projection onto $\mathcal{C}^\circ$ and, for $t \geq 0$ $r_t = g_t(a_t, b_t)$.

**Remark.** *The associated sequence of this adaptation can be written by (3.10) (applied with $H_t = \frac{1}{\gamma_t} H = \frac{1}{2\gamma_t}\|\cdot\|_2^2$) as :*

$$x_{t+1} = \Pi_{\mathcal{C}^\circ}(x_t + \gamma_t r_t) = \arg\max_{x \in \mathcal{C}^\circ}\left\{\langle \gamma_t r_t, x\rangle - \frac{1}{2}\|x - x_t\|_2^2\right\}$$

*where the step-size $\gamma_t := \frac{\gamma}{\sqrt{\sum_{s=0}^t \|r_s\|_2^2}}$ is a nonincreasing sequence.*

Throughout this work, I will refer to the properties developed in the course (available here `https://joon-kwon.github.io/regret-ups/lecture-notes-online-learning.pdf`) in blue.

1. This adaptation of AdaGrad-Norm is parameter-free.

*Proof.* Let $(x'_t)_{t\geq 0}$ a sequence defines as $x'_{t+1} = \Pi_{\mathcal{C}^\circ}\left(x'_t + \frac{\eta}{\sqrt{\sum_{s=0}^t \|r_s\|_2^2}} r_t\right)$ with $\eta > 0$. Let us prove that $x'_t = \frac{\eta}{\gamma} x_t$ for all $t \geq 0$. It is true for $t = 0$, as $x'_0 = x_0 = 0$. Then by induction, for $t \geq 1$,

$$
\begin{aligned}
x'_{t+1} &= \Pi_{\mathcal{C}^\circ}\left(x'_t + \frac{\eta}{\sqrt{\sum_{s=0}^t \|r_s\|_2^2}} r_t\right) \\
&= \Pi_{\mathcal{C}^\circ}\left(\frac{\eta}{\gamma} x_t + \frac{\eta}{\sqrt{\sum_{s=0}^t \|r_s\|_2^2}} r_t\right) && \text{by supposition} \\
&= \Pi_{\mathcal{C}^\circ}\left(\frac{\eta}{\gamma}\left(x_t + \frac{\gamma}{\sqrt{\sum_{s=0}^t \|r_s\|_2^2}} r_t\right)\right) \\
&= \frac{\eta}{\gamma}\Pi_{\mathcal{C}^\circ}\left(x_t + \frac{\gamma}{\sqrt{\sum_{s=0}^t \|r_s\|_2^2}} r_t\right) && \text{using Proposition 5.1.4} \\
&= \frac{\eta}{\gamma} x_{t+1}
\end{aligned}
$$

Thus $\alpha(x'_t) = \alpha\left(\frac{\eta}{\gamma} x_t\right) = \alpha(x_t) = a_t$ for all $t \geq 0$, the actions chosen by dual averaging do not depend on the parameter $\gamma$. $\square$

2. The adaptation of AdaGrad-Norm algorithm guarantees for all $T \geq 0$,
(i)
$$\min_{r\in\mathcal{C}} \left\|\sum_{t=0}^{T} r_t - r\right\|_2 \leq \left(\frac{\max_{0\leq t\leq T}\|x-x_t\|_2^2}{2\gamma} + \gamma\right)\sqrt{\sum_{t=0}^{T}\|r_t\|_2^2}$$
(ii) Moreover, if $\exists R \in \mathbb{R}$ such that $\max_{0\leq t\leq T}\|x-x_t\|_2 \leq R$, $\gamma = \frac{R}{\sqrt{2}}$ yields
$$\min_{r\in\mathcal{C}} \left\|\sum_{t=0}^{T} r_t - r\right\|_2 \leq R\sqrt{2\sum_{t=0}^{T}\|r_t\|_2^2}$$
(iii) Moreover, assume that $f$ (which admits a minimizer $\mathcal{C}^\circ$ denote as $x_*$) is L-Lipschitz for $\|\cdot\|_2$ and $-r_t \in \partial f(x_t)$ for all $t \geq 0$, then
$$\min_{0\leq t\leq T} f(x) - f(x_*) \leq RL\sqrt{\frac{2}{T+1}}$$

*Proof.* (i) The regret bound for OGD with time-dependent step-sizes from Corollary 3.3.16 gives

$$\begin{aligned}
\sum_{t=0}^{T}\langle r_t, x-x_t\rangle &\leq \frac{\max_{0\leq t\leq T}\|x-x_t\|_2^2}{2\gamma_T} + \sum_{t=0}^{T}\frac{\gamma_t\|r_t\|_2^2}{2} \\
&= \frac{\max_{0\leq t\leq T}\|x-x_t\|_2^2}{2\gamma}\sqrt{\sum_{t=0}^{T}\|r_t\|_2^2} + \frac{\gamma}{2}\sum_{t=0}^{T}\frac{\|r_t\|_2^2}{\sqrt{\sum_{s=0}^{t}\|r_t\|_2^2}} \\
&\leq \left(\frac{\max_{0\leq t\leq T}\|x-x_t\|_2^2}{2\gamma} + \gamma\right)\sqrt{\sum_{t=0}^{T}\|r_t\|_2^2} \qquad \text{using Lemma 7.2.1}
\end{aligned}$$

Moreover,

$$\max_{x\in\mathcal{C}^\circ}\sum_{t=0}^{T}\langle r_t, x-x_t\rangle \geq \max_{x\in\mathcal{C}^\circ\cap B_2}\left\langle\sum_{t=0}^{T} r_t, x\right\rangle \qquad \text{using Lemma 5.4.2}$$

(applied with $\mathcal{X}_0 = \mathcal{C}^\circ \cap B_2$ where $B_2$ is the closed Euclidean unit ball)

$$= \min_{r\in\mathcal{C}}\left\|\sum_{t=0}^{T} r_t - r\right\|_{2,*} = \min_{r\in\mathcal{C}}\left\|\sum_{t=0}^{T} r_t - r\right\|_2 \qquad \text{using Proposition 5.4.1}$$

So,

$$\min_{r\in\mathcal{C}} \left\|\sum_{t=0}^{T} r_t - r\right\|_2 \leq \left(\frac{\max_{0\leq t\leq T}\|x-x_t\|_2^2}{2\gamma} + \gamma\right)\sqrt{\sum_{t=0}^{T}\|r_t\|_2^2}$$

(ii) If $\exists R \in \mathbb{R}$ such that $\max_{0\leq t\leq T} \|x - x_t\|_2 \leq R$, then

$$\min_{r\in\mathcal{C}} \left\| \sum_{t=0}^{T} r_t - r \right\|_2 \leq \left( \frac{R^2}{2\gamma} + \gamma \right) \sqrt{\sum_{t=0}^{T} \|r_t\|_2^2}$$

And in the particular case where $\gamma = \frac{R}{\sqrt{2}}$,

$$\min_{r\in\mathcal{C}} \left\| \sum_{t=0}^{T} r_t - r \right\|_2 \leq R\sqrt{2\sum_{t=0}^{T} \|r_t\|_2^2}$$

(iii) The Lemma 6.1.1 gives

$$\begin{aligned}
\min_{0\leq t\leq T} f(x) - f(x_*) &\leq \left( \sum_{t=0}^{T} \gamma_t \right)^{-1} \sum_{t=0}^{T} \langle \gamma_t r_t, x_t - x \rangle \\
&\leq \frac{1}{T+1} \left( \frac{\max_{0\leq t\leq T} \|x - x_t\|_2^2}{2\gamma} + \gamma \right) \sqrt{\sum_{t=0}^{T} \|r_t\|_2^2} && \text{with (i)} \\
&\leq \frac{1}{T+1} \left( \frac{\max_{0\leq t\leq T} \|x - x_t\|_2^2}{2\gamma} + \gamma \right) \sqrt{\sum_{t=0}^{T} L^2} && \text{by Lipschitz continuity} \\
&\leq \frac{L}{\sqrt{T+1}} \left( \frac{\max_{0\leq t\leq T} \|x - x_t\|_2^2}{2\gamma} + \gamma \right)
\end{aligned}$$

And if $\exists R \in \mathbb{R}$ such that $\max_{0\leq t\leq T} \|x - x_t\|_2 \leq R$, $\gamma = \frac{R}{\sqrt{2}}$ yields

$$\min_{0\leq t\leq T} f(x) - f(x_*) \leq LR\sqrt{\frac{2}{T+1}}$$

□

3. In the special case of regret minimization on the simplex, the algorithm can be written with the Hart–Mas-Colell reduction of Proposition 5.5.1. Thus
   - $\mathcal{A} = \Delta_d$
   - $\mathcal{B} = \mathbb{R}^d$
   - $r_t = g_t(a_t, b_t) = g(a_t, b_t)$ with $g : (a, b) \mapsto b - \langle b, a\rangle \mathbb{I}$, $\mathbb{I} := (1, ..., 1) \in \mathbb{R}^d$
   - $\mathcal{C} = \mathbb{R}^d_-$ satisfies Blackwell's condition with respect to $g$ and $\mathcal{C}^\circ = \mathbb{R}^d_+$
   - with the notation $y_+ = (\max(y_i, 0))_{1\le i\le d}$

   $$a_t := \begin{cases} \frac{x_t}{\|x_t\|_1} & \text{if } x_t \neq 0 \\ a_0 \in \Delta_d & \text{otherwise} \end{cases}, \quad x_{t+1} = (x_t + \gamma_t r_t)_+ = (x_t + \gamma_t(b_t - \langle b_t, a_t\rangle \mathbb{I}))_+$$

*Proof.* The oracle definition with the assumptions of the Hart–Mas-Colell reduction is

$$\alpha(x) := \begin{cases} \frac{x}{\|x\|_1} & \text{if } x_t \neq 0 \\ a_0 \in \Delta_d & \text{otherwise} \end{cases}, \quad x \in \mathbb{R}^d_+$$

And the algorithm defines $a_t := \alpha(x_t)$ so the first part is direct. Then because

$$\Pi_{\mathbb{R}^d_+}(y) = \arg\min_{y' \in \mathbb{R}^d_+} \|y' - y\|_2^2 = \arg\min_{y'_1,...,y'_d \ge 0} \sum_{i=1}^d (y'_i - y_i)^2 = \left(\arg\min_{y'_i \ge 0} (y'_i - y_i)^2\right)_{1\le i\le d} = (\max(y_i, 0))_{1\le i\le d} = y_+$$

and $\mathcal{C}^\circ = \mathbb{R}^d_+$ the iterates can be written as

$$x_{t+1} = \Pi_{\mathcal{C}^\circ}(x_t + \gamma_t r_t) = \Pi_{\mathbb{R}^d_+}(x_t + \gamma_t r_t) = (x_t + \gamma_t r_t)_+ = (x_t + \gamma_t(b_t - \langle b_t, a_t\rangle \mathbb{I}))_+$$

□

This new algorithm of regret minimization on the simplex has guarantees :
(i)

$$\max_{x\in\Delta_d} \sum_{t=0}^T \langle b_t, x - a_t\rangle \le \left(\frac{\max_{0\le t\le T} \|x - x_t\|_2^2}{2\gamma} + \gamma\right) \sqrt{\sum_{t=0}^T \|b_t - \langle b_t, a_t\rangle \mathbb{I}\|_2^2}$$

(ii) And if $\exists R \in \mathbb{R}$ such that $\max_{0\le t\le T} \|x - x_t\|_2 \le R$, $\gamma = \frac{R}{\sqrt{2}}$ and $\exists L \in \mathbb{R}$ such that $\|b_t\|_\infty \le L$ for all $t \ge 0$, then

$$\max_{x\in\Delta_d} \sum_{t=0}^T \langle b_t, x - a_t\rangle \le 2\sqrt{2} R L \sqrt{d(T+1)}$$

*Proof.*
(i) follows by noticing that $\Delta_d \subset \mathbb{R}_+^d \cap B_2$ and combining Proposition 5.4.1 with the guarantees developped in question 1.
(ii) Because when $\|b_t\|_\infty < L$,

$$\begin{aligned}\|b_t - \langle b_t, a_t\rangle \mathbb{I}\|_2 &\leq \|b_t\|_2 + \langle b_t, a_t\rangle \|\mathbb{I}\|_2 \\ &\leq \sqrt{d}\|b_t\|_\infty + \|b_t\|_\infty \|a_t\|_1 \|\mathbb{I}\|_2 \\ &\leq \sqrt{d}\|b_t\|_\infty + \|b_t\|_\infty \|a_t\|_1 \sqrt{d} && \text{because } \mathbb{I} := (1, ..., 1) \in \mathbb{R}^d \\ &\leq 2L\sqrt{d} && \text{because } a_t \in \Delta_d\end{aligned}$$

there is

$$\begin{aligned}\max_{x\in\Delta_d} \sum_{t=0}^{T} \langle b_t, x - a_t\rangle &\leq \left(\frac{\max_{0\leq t\leq T} \|x - x_t\|_2^2}{2\gamma} + \gamma\right) \sqrt{\sum_{t=0}^{T} \left(2L\sqrt{d}\right)^2} \\ &\leq \left(\frac{\max_{0\leq t\leq T} \|x - x_t\|_2^2}{2\gamma} + \gamma\right) \left(2L\sqrt{d(T+1)}\right)\end{aligned}$$

and if $\exists R \in \mathbb{R}$ such that $\max_{0\leq t\leq T} \|x - x_t\|_2 \leq R$, $\gamma = \frac{R}{\sqrt{2}}$ yields

$$\max_{x\in\Delta_d} \sum_{t=0}^{T} \langle b_t, x - a_t\rangle \leq 2\sqrt{2}RL\sqrt{d(T+1)}$$

□

4. The algorithm apply to regret learning for finite two-player zero-sum games represented by a matrix $A \in \mathbb{R}^{m\times n}$ can be written for all $t \geq 0$ as :

$$a_t := \begin{cases} \frac{x_t}{\|x_t\|_1} & \text{if } x_t \neq 0 \\ a_0 & \text{otherwise} \end{cases} \qquad b_t := \begin{cases} \frac{w_t}{\|w_t\|_1} & \text{if } w_t \neq 0 \\ b_0 & \text{otherwise} \end{cases}$$

$$x_{t+1} = (x_t + \gamma_t(Ab_t - \langle a_t, Ab_t\rangle \mathbb{I}))_+ \quad w_{t+1} = (w_t + \gamma_t(\langle a_t, Ab_t\rangle \mathbb{I} - A^\top a_t))_+$$

with $a_0 \in \Delta_m, b_0 \in \Delta_n, x_0 = w_0 = 0$.

For all $T \geq 0$, let $\overline{a}_T = \frac{1}{T+1}\sum_{t=0}^T a_t$ and $\overline{b}_T = \frac{1}{T+1}\sum_{t=0}^T b_t$, then

$$\delta_A(\overline{a}_T, \overline{b}_T) \leq \frac{\sqrt{n}+\sqrt{m}}{\sqrt{T+1}}\left(\frac{\max_{0\leq t\leq T}\|x-x_t\|_2^2}{2\gamma} + \gamma\right)(A_{\max} - A_{\min})$$

where $\delta_A$ is the duality gap $\delta_A : \Delta_m \times \Delta_n \to \mathbb{R}_+$ such that

$$\delta_A(a,b) = \max_{a'\in\Delta_m}\langle a', Ab\rangle - \min_{b'\in\Delta_n}\langle a, Ab'\rangle$$

And if $\exists R \in \mathbb{R}$ such that $\max_{0\leq t\leq T}\|x-x_t\|_2^2 \leq R$, $\gamma = \frac{R}{\sqrt{2}}$ yields

$$\delta_A(\overline{a}_T, \overline{b}_T) \leq \frac{\sqrt{n}+\sqrt{m}}{\sqrt{T+1}}R\sqrt{2}(A_{\max} - A_{\min})$$

*Proof.* The general regret bound from question 3. gives

$$\max_{a\in\Delta_m}\sum_{t=0}^T\langle Ab_t, a - a_t\rangle \leq \left(\frac{\max_{0\leq t\leq T}\|x-x_t\|_2^2}{2\gamma} + \gamma\right)\sqrt{\sum_{t=0}^T\|Ab_t - \langle a_t, Ab_t\rangle \mathbb{I}\|_2^2}$$

$$\max_{b\in\Delta_n}\sum_{t=0}^T\langle -A^\top a_t, b - b_t\rangle \leq \left(\frac{\max_{0\leq t\leq T}\|x-x_t\|_2^2}{2\gamma} + \gamma\right)\sqrt{\sum_{t=0}^T\|\langle a_t, Ab_t\rangle \mathbb{I} - A^\top a_t\|_2^2}$$

For all $t \geq 0$, because $a_t \in \Delta_m$ and $b_t \in \Delta_n$,

$$\|Ab_t - \langle a_t, Ab_t\rangle \mathbb{I}\|_2^2 = \sum_{i=1}^m \left((Ab_t)_i - \langle a_t, Ab_t\rangle\right)^2 \leq \sum_{i=1}^m (A_{\max} - A_{\min})^2 = m(A_{\max} - A_{\min})^2$$

and similarly

$$\left\|\langle a_t, Ab_t\rangle \mathbb{I} - A^\top a_t\right\|_2^2 \leq n(A_{\max} - A_{\min})^2$$

So,

$$\begin{aligned}
\delta_A(\bar{a}_T, \bar{b}_T) &= \frac{1}{T+1}\left(\max_{a\in\Delta_m}\sum_{t=0}^{T}\langle Ab_t, a-a_t\rangle + \max_{b\in\Delta_n}\sum_{t=0}^{T}\langle -A^\top a_t, b-b_t\rangle\right) && \text{using Lemma 9.3.1} \\
&\leq \frac{1}{T+1}\left(\frac{\max_{0\leq t\leq T}\|x-x_t\|_2^2}{2\gamma}+\gamma\right)\left(\sqrt{T+1}\right)\left(\sqrt{n}+\sqrt{m}\right)(A_{\max}-A_{\min}) \\
&= \frac{\sqrt{n}+\sqrt{m}}{\sqrt{T+1}}\left(\frac{\max_{0\leq t\leq T}\|x-x_t\|_2^2}{2\gamma}+\gamma\right)(A_{\max}-A_{\min})
\end{aligned}$$

And if $\exists R \in \mathbb{R}$ such that $\max_{0\leq t\leq T}\|x-x_t\|_2 \leq R$, $\gamma = \frac{R}{\sqrt{2}}$ yields

$$\delta_A(\bar{a}_T, \bar{b}_T) \leq \frac{\sqrt{n}+\sqrt{m}}{\sqrt{T+1}} R\sqrt{2}(A_{\max}-A_{\min})$$

□

**The second part of the question as well as question 5. is available on the GitHub :** `https://github.com/Biechy/OnlineLearning`

6. Let $\varepsilon > 0$. We can also consider an adaptation of AdaGrad-Diagonal, with the same notations as before, as follow :

$$x'_{t+1} = \left( x'_{t,i} + \frac{\gamma}{\varepsilon + \sqrt{\sum_{s=0}^{t} r_{s,i}^2}} r_{t,i} \right)_{1 \le i \le d}$$

$$x_{t+1} = \arg\min_{x \in \mathcal{C}^\circ} \left\| x - x'_{t+1} \right\|_{A_t}$$

where

$$A_t = \operatorname{diag}\left( \varepsilon + \sqrt{\sum_{s=0}^{t} r_{s,i}^2} \right)_{1 \le i \le d}$$

Like the Proposition 7.2.5, this algorithm has guarantees :
(i) For all $T \ge 0$,

$$\sum_{t=0}^{T} \langle r_t, x - x_t \rangle \le \frac{\varepsilon}{2\gamma} \|x\|_2^2 + \left( \frac{\max_{0 \le t \le T} \|x_t - x\|_\infty^2}{2\gamma} + \gamma \right) \sum_{i=1}^{d} \sqrt{\sum_{t=0}^{T} r_{t,i}^2} \text{ because } x_0 = 0$$

$$\le \frac{\varepsilon}{2\gamma} \|x\|_2^2 + \left( \frac{\max_{0 \le t \le T} \|x_t - x\|_\infty^2}{2\gamma} + \gamma \right) \sqrt{d \sum_{t=0}^{T} \|r_t\|_2^2} \text{ using (iii) of Proposition 7.2.5}$$

(ii) In particular, if $R \ge \max_{0 \le t \le T} \|x_t - x\|_\infty$, then $\gamma = \frac{R}{\sqrt{2}}$ yields,

$$\sum_{t=0}^{T} \langle r_t, x - x_t \rangle \le \frac{R\varepsilon}{\sqrt{2}} \|x\|_2^2 + R\sqrt{2d \sum_{t=0}^{T} \|r_t\|_2^2} \le R\left( \frac{\varepsilon d}{\sqrt{2}} + \sqrt{2d \sum_{t=0}^{T} \|r_t\|_2^2} \right)$$

*Proof.* Exactly the same proof of Proposition 7.2.5 but with $u_t \equiv r_t$ □

Thus, we could find :
1.This algorithm is not parameter-free because you have to choose at least $\varepsilon$.
2.Guarantee : $\min_{r \in \mathcal{C}} \left\| \sum_{t=0}^{T} r_t - r \right\|_2 \le R\left( \frac{\varepsilon d}{\sqrt{2}} + \sqrt{2d \sum_{t=0}^{T} \|r_t\|_2^2} \right)$.
3.and4. To be explore...