UMD-based extension of Adagrad-Norm and application to games

Abbadi Omar

April 2024

In this project we study a generalization of the Adagrad algorithm based on UMD iterates, then study special cases, such as dual averaging (DA) and online mirror descent (OMD), with different regularizers. Then we apply this to regret learning in zero-sum games and conclude with some numerical experiments.

Contents

Let $d \geqslant 1, \mathcal{X} \subset \mathbb{R}^d$ a nonempty closed convex set, $\mathrm{K} > 0, \|\cdot\|$ a norm on $\mathbb{R}^d, h$ a regularizer on $\mathcal{X}$ that is K-strongly convex for $\|\cdot\|$.

Let $\gamma > 0$. For $(u_t)_{t\geqslant 0}$ a sequence in $\mathbb{R}^d$, let $(x_t, y_t)_{t\geqslant 0}$ be a sequence of strict UMD iterates associated with regularizer b and dual increments $(\gamma_t u_t)_{t\geqslant 0}$, where

$$\gamma_t = \frac{\gamma}{\sqrt{\Sigma_{s=0}^t \|u_s\|_*^2}}, \quad t \geqslant 0$$

with convention $0/0 = 0$.

Suppose in all results that follow that we have $u_0 \neq 0$. Indeed, if not, let $t = \min\{t \geqslant 0 \mid u_t \neq 0\}$. Define γ_t as follows:

$$\gamma_t = \begin{cases} \frac{\gamma}{\|u_\tau\|_*} & \text{if } t < \tau, \\ \frac{\gamma}{\sqrt{\sum_{k=0}^t \|u_t\|_*^2}} & \text{if } t \geqslant \tau. \end{cases}$$

then we get the same algorithms and guarantees by considering this step-size instead.

1 UMD extension

I derived two guarantees for this question.

For the first one we see $(x_t, y_t)_{\geqslant 0}$ is a strict UMD with dual increments $(\gamma_t u_t)_{\geqslant 0}$ and constant regularizer h.

We get by UMD lemma (Lemma 2.4.1) :

$$\begin{aligned} &\langle \gamma_t u_t, x - x_t \rangle \leqslant D_t - D_{t+1} + D_t^* \\ &\text{where } D_t = D_t(x, x_t; y_t) \\ &\text{and } D_t^* = D_{h^*}(y_t + \gamma_t u_t; y_t) \end{aligned}$$

Therefore: $\langle u_t, x - x_t \rangle \leqslant \frac{1}{\gamma_t}(D_t - D_{t+1} + D_t^*)$

$$\text{So } \langle u_t, x - x_t \rangle \leqslant \frac{1}{\gamma_t} D_t - \frac{1}{\gamma_{t+1}} D_{t+1} + \left(\frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \right) D_{t+1} + \frac{1}{\gamma_t} D_t^*$$

$$\begin{aligned}
\text{And } \sum_{t=0}^{T} \langle u_t, x - x_t \rangle &\leqslant \frac{1}{\gamma_0} D_0 + \sum_{t=0}^{T} \left(\frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \right) D_{t+1} + \sum_{t=0}^{T} \frac{1}{\gamma_t} D_t^* \\
&\leqslant \frac{1}{\gamma_0} D_0 + \frac{1}{\gamma_{T+1}} \max_{0 \leqslant t \leqslant T} D_{t+1} + \sum_{t=0}^{T} \frac{1}{\gamma_t} D_t^* \\
&\quad \left(\text{since } \frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \geqslant 0 \text{ and } D_t \geqslant 0 \text{ by convexity of } h\right)
\end{aligned}$$

and $D_t^* \leqslant \frac{1}{2K} \|u_t\|_*^2$ (by strong convexity of h).
and by lemma 7.2.1, we get $\sum_{t=0}^{T} \frac{1}{\gamma_t} D_t^* \leqslant \frac{\gamma}{K} \sqrt{\sum_{t=0}^{T} \|u_t\|_*^2}$.

For the second guarantee, we view $\left(x_t, \frac{y_t}{\gamma_t} \right)_{t \geqslant 0}$ as a strict UMD system with dual increments $(u_t)_{t \geqslant 0}$ and time-dependent regularizers: $h_t = \frac{1}{\gamma_t}$ and $h_{t+\frac{1}{2}}(x) = h\left(\frac{x}{\gamma_t} \right)$ for all x. Then, since $\frac{y_t}{\gamma_t} \in \partial f_t(x_t)$, we have:

$$\forall x \quad \frac{h(x)}{\gamma_t} - \frac{h(x_t)}{\gamma_t} \geqslant \left\langle \frac{y_t}{\gamma_t}, x - x_t \right\rangle, \text{ so } y_t \in \partial h(x_t).$$

and $x_{t+1} = \nabla h^* \left(\gamma_t \left(\frac{y_t}{\gamma_t} + u_t \right) \right) = \nabla h^* \left(y_t + \gamma_t u_t \right)$ since $h(ax)^* = h^* \left(\frac{x^*}{a} \right)$ for any constant a.
So by Lemma 2.5.2, $\langle u_t, x - x_t \rangle = D_t - D_{t+1} + D_t^* + D_{t+1/2}^{\Delta} + \Delta h_t(x)$. (These are defined in the lecture notes)

Hence:

$$\sum_{t=0}^{T} \langle u_{t,x} - x_t \rangle \leqslant D_0 + \sum_{t=0}^{T} \left(D_t^* + D_{t+1/2} + \Delta h_t(x) \right)$$

2 DA and OMD extension

2.1 General regularizer

For DA, we consider the algorithm

$$\begin{cases} z_t = y_t/\gamma_t \\ x_t = \nabla h^* (\gamma_t z_t) \\ z_{t+1} = z_t + u_t \end{cases}$$

We then have indeed $x_t = \nabla h^* (y_t)$ and $y_{t+1} = y_t + \gamma_t u_t$. And , since $(\gamma_t)_{\geqslant 0}$ is nonincreasing and positive, we can apply Proposition 3.2.6 :

$$\begin{aligned} \sum_{t=0}^{T} \langle u_t, x - x_t \rangle &\leqslant \frac{h(x) - \min h}{\gamma_T} + \frac{\gamma}{K} \sqrt{\sum_{t=0}^{T} \|u_t\|_*^2} \\ &= \left(\frac{h(x) - \min h}{\gamma} + \frac{\gamma}{K} \right) \sqrt{\sum_{t=0}^{T} \|u_t\|_*^2} \text{ (using Lemma 7.2.1).} \end{aligned}$$

For OMD, assuming h is a mirror map, since again $(\gamma_t)_{\geqslant 0}$ is nonincreasing and positive, we can then apply the OMD iteration algorithm and apply Proposition 3.3.14 to get the regret bound :

$$\begin{aligned} \sum_{t=0}^{T} \langle u_t, x - x_t \rangle &\leqslant \max_{0 \leqslant t \leqslant T} \frac{D_h (x, x_t)}{\gamma_T} + \frac{1}{2K} \sum_{k=0}^{T} \gamma_T \|u_T\|_*^2 \\ &\leqslant \left(\frac{1}{\gamma} \max_{0 \leqslant t \leqslant T} D_h (x, x_t) + \frac{\gamma}{K} \right) \sqrt{\sum_{t=0}^{T} \|u_t\|_*^2} \text{ (using Lemma 7.2.1)} \end{aligned}$$

2.2 Euclidean regularizer

We let $h = \frac{1}{2} \| \cdot \|^2$ which is 1-strongly convex, we have $D_h(x, y) = \|x - y\|^2$.

For DA we thus get :

$$\begin{aligned} \sum_{t=0}^{T} \langle u_t, x - x_t \rangle &\leqslant \frac{1}{2} \frac{\|x\|^2}{\gamma_T} + \gamma \sqrt{\sum_{t=0}^{T} \|u_t\|_*^2} \\ &= \left(\frac{\|x\|^2}{2\gamma} + \gamma \right) \sqrt{\sum_{t=0}^{T} \|u_t\|_*^2} \end{aligned}$$

For OMD, we get:

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{1}{2\gamma_T} \max_{0 \leqslant t \leqslant T} \|x - x_t\|^2 + \gamma \sqrt{\sum_{t=0}^{T} \|u_t\|_*^2}$$
$$= \left(\frac{1}{2\gamma} \max_{0 \leqslant t \leqslant T} \|x - x_t\|^2 + \gamma \right) \sqrt{\sum_{t=0}^{T} \|u_t\|_*^2}$$

We find the same result as for Proposition 7.2.3.

2.3 Entropic regularizer

For the entropic regularizer

$$h_{\text{ent}}(x) = \begin{cases} \sum_{i=1}^{d} x_i \log x_i & \text{if } x \in \Delta_d, \text{ with convention } 0 \cdot \log 0 = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

we can apply the exponential weight algorithm with $(u_t)_{\geqslant 0}$ and $(\gamma_t)_{\geqslant 0}$ and associated iterates, defined in the simplex Δ_d as:

$$x_t = \left(\frac{\exp\left(\gamma_t \sum_{s=0}^{t-1} u_{s,i}\right)}{\sum_{j=1}^{d} \exp\left(\gamma_t \sum_{s=0}^{t-1} u_{s,j}\right)} \right)_{1 \leqslant i \leqslant d}, t \geqslant 0.$$

Then by Proposition 3.4.4 $(x_t)_{\geqslant 0}$ is a DA with regularizer h_{ent} parameters $(\gamma_t)_{\geqslant 0}$ and dual increments $(u_t)_{\geqslant 0}$ and since the former is nonincreasing and positive we get the following guarantee :

$$\mathbb{E}\left[\max_{1 \leqslant i \leqslant d} \sum_{t=0}^{T} (u_{t,i} - u_{t,i_t}) \right] \leqslant \frac{\log d}{\gamma_T} + \sum_{t=0}^{T} \frac{D_{h_{\text{ent}}^*}(y_t + \gamma_t u_t, y_t)}{\gamma_t}$$

3 Application to games

3.1 Regret learning

We will consider a variant of Optimistic OMD with time-dependant step-sizes :

Let $(x_t, y_t)_{t\geqslant 0}$ be a sequence of strict UMD iterates associated with regularizer h and dual iterates $(2\gamma_t u_t - \gamma_{t-1}u_{t-1})_{t\geqslant 0}$ (with convention $u_{-1} = 0$). Then for all $T \geqslant 0, \alpha > 0$, and $x \in \operatorname{dom} h$,

$$\sum_{t=0}^{T} \langle u_t, x - x_t\rangle \leqslant D_h(x, x_0; y_0) + \frac{\alpha}{2}\|x - x_T\|^2 + \frac{1}{2\alpha}\|u_0\|_*^2$$
$$+ \frac{1}{2\alpha}\sum_{t=0}^{T-1}\|\gamma_{t+1}u_{t+1} - \gamma_t u_t\|_*^2 + \left(\frac{\alpha}{2} - \frac{K}{2}\right)\sum_{t=0}^{T-1}\|x_{t+1} - x_t\|^2.$$

This follows from Lemma 9.4.1. by taking $(\gamma_t u_t)_{\geqslant 0}$ instead of $(u_t)_{\geqslant 0}$.

We can the consider a time-dependant step-size version of optimistic exponential weights for two-player zero-sum games : For $t \geqslant 0$, let

$$y_t = \left(\sum_{s=0}^{t-1}\gamma_s A b_s + \gamma_{t-1}A b_{t-1}\right), \quad z_t = -\left(\sum_{s=0}^{t-1}\gamma_s A^\top a_s + \gamma_{t-1}A^\top a_{t-1}\right),$$

and

$$a_t = \left(\frac{\exp(y_{t,i})}{\sum_{i'=1}^m \exp(y_{t,i'})}\right)_{1\leqslant i\leqslant m}, \quad b_t = \left(\frac{\exp(z_{t,j})}{\sum_{j'=1}^m \exp(z_{t,j'})}\right)_{1\leqslant j\leqslant n}$$

Then $(a_t)_{t\geqslant 0}$ (resp. $(b_t)_{t\geqslant 0}$) corresponds to the exponential weights algorithm associated with dual increments $(2\gamma_t A b_t - \gamma_{t-1}A b_{t-1})_{t\geqslant 0}$ (resp. $\big(2\gamma_t A^\top a_t - \gamma_{t-1}A^\top a_{t-1}\big)_{t\geqslant 0}$) (with convention $b_{-1} = 0$ and $a_{-1} = 0$).

Because $y_0 = 0$, if h is the entropic regularizer on Δ_m,

$$D_h(a, a_0; y_0) = h(a) - h(a_0) \leqslant \log m,$$

by Proposition 3.4 .3 . Similarly, if $\tilde{h}$ is the entropic regularizer on Δ_n, because $z_0 = 0$,

$$D_{\tilde{h}}(b, b_0; z_0) \leqslant \log n.$$

Then, because the entropic regularizer is 1 -strongly convex for $\|\cdot\|_1$ by

Proposition 2.2.6, applying Lemma 9.4.1 gives regret bounds

$$\sum_{t=0}^{T} \langle Ab_t, a - a_t \rangle \leqslant \log m + \frac{1}{4} \|a - a_T\|_1^2 + \|2\gamma_0 A b_0\|_\infty^2 + \sum_{t=0}^{T-1} \|\gamma_{t+1} A b_{t+1} - \gamma_t A b_t\|_\infty^2 - \frac{1}{4} \sum_{t=0}^{T-1} \|a_{t+1} - a_t\|_1^2,$$

$$\sum_{t=0}^{T} \left\langle -A^\top a_t, b - b_t \right\rangle \leqslant \log n + \frac{1}{4} \|b - b_T\|_1^2 + \|2\gamma_0 A^\top a_0\|_\infty^2 \sum_{t=0}^{T-1} \left\|\gamma_{t+1} A^\top a_{t+1} - \gamma_t A^\top a_t\right\|_\infty^2 - \frac{1}{4} \sum_{t=0}^{T-1} \|b_{t+1} - b_t\|_1^2 .$$

Note that for all $t \geqslant 0$,

$$\|\gamma_{t+1} A b_{t+1} - \gamma_t A b_t\|_\infty^2 \leqslant \|A\|_\infty^2 \|\gamma_{t+1} b_{t+1} - \gamma_t b_t\|_1^2 ,$$
$$\left\|\gamma_{t+1} A^\top a_{t+1} - \gamma_t A^\top a_t\right\|_\infty^2 \leqslant \|A\|_\infty^2 \|\gamma_{t+1} a_{t+1} - \gamma_t a_t\|_1^2 .$$

Besides,

$$\|2\gamma_0 A b_0\|_\infty^2 \leqslant 4\|A\|_\infty^2 \|\gamma_0 b_0\|_1^2 = 4\gamma_0^2 \|A\|_\infty^2,$$

similarly $\left\|2A^\top a_0\right\|_\infty^2 \leqslant 4\gamma_0^2 \|A\|_\infty^2$, and

$$\|a - a_T\|_1^2 \leqslant (\|a\|_1 + \|a_T\|_1)^2 = 4,$$

and similarly $\|b - b_T\|_1^2 \leqslant 4$

Hence, by summing :

$$(T+1) \cdot \delta_A \left(\bar{a}_T, \bar{b}_T\right) \leqslant \log m + \log n + 2 + 8\gamma_0^2 \|A\|_\infty^2 + \|A\|_\infty^2 (\|\gamma_{t+1} b_{t+1} - \gamma_t b_t\|_1^2 + \|\gamma_{t+1} a_{t+1} - \gamma_t a_t\|_1^2) - \frac{1}{4} \left(\sum_{t=0}^{T-1} \|a_{t+1} - a_t\|_1^2 + \sum_{t=0}^{T-1} \|b_{t+1} - b_t\|_1^2 \right)$$

3.2 Numerical experiments

Numerical experiment are to be found in the notebook associated to this project. In it, I programmed DA and OMD with Euclidean regularizers (on the unit ball) and entropic regularizer (exponential weights algorithm) with the dual iterates sampled randomly. For each of them I visualized the evolution of the regret compared to the theoretical guarantee as a function of the horizon T and also the evolution of the iterates in the plane in the case of $d = 2$ and $\gamma = 1$.

I also ran simulations for the classical exponential weights algorithm and regret marching (RM and RM+).

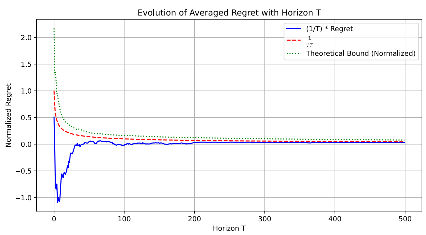

Figure 1: Evolution of regret for Euclidean DA

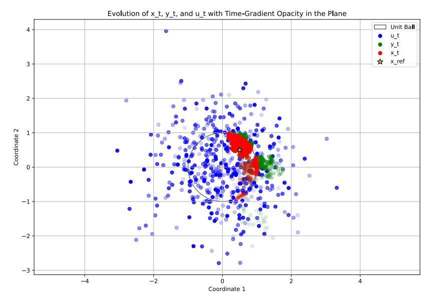

Figure 2: Evolution of iterates for Euclidean DA

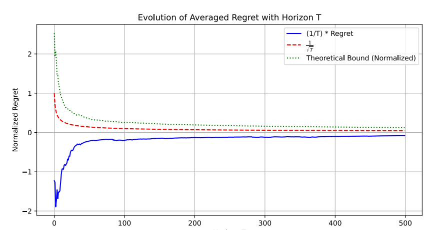

Figure 3: Evolution of regret for Euclidean OMD

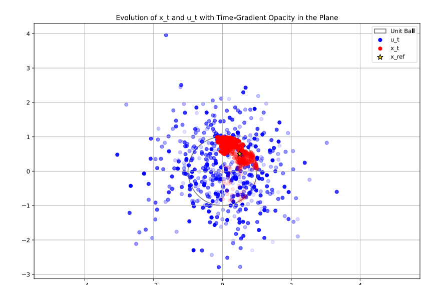

Figure 4: Evolution of iterates for Euclidean OMD

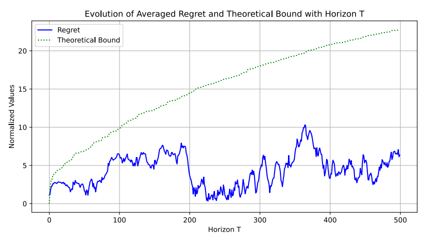

Figure 5: Evolution of regret for exponential weights algorithm with parameters $(\gamma_t)_{\geqslant 0}$

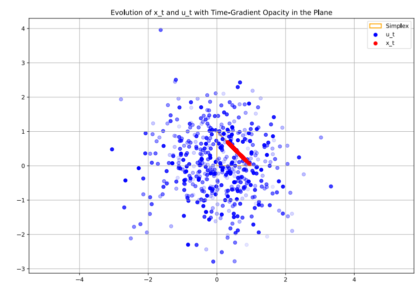

Figure 6: Evolution of iterates for exponential weights algorithm parameters $(\gamma_t)_{\geqslant 0}$

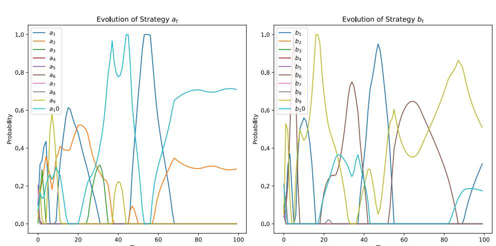

Figure 7: Evolution of strategies in RM (10 dimensions)

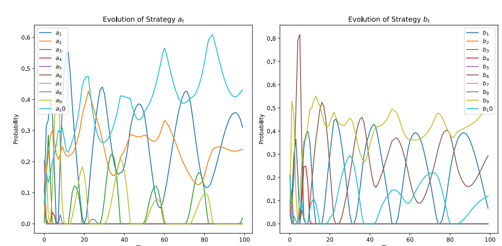

Figure 8: Evolution of strategies in RM+ (10 dimensions)