ONLINE LEARNING, LINKS WITH OPTIMIZATION AND GAMES: FINAL PROJECT

Solving games: extrapolation vs optimism

Alex Pierron
https://github.com/AlexPierron/Online-Learning-Links-with-Optimization-and-Games
Paris-Saclay University / Mathematical Institute of Orsay
M2 Mathematics & AI

Online learning, links with optimization and games
2023-2024

Table of contents

I Initial theoretical properties

We set $d \geq 1$ and $\hbar_{\text{ent}}$ the entropic regularizer on the simplex Δ_d. We remind both these definitions:

Definition

$$\Delta_d = \left\{ x \in \mathbb{R}_+^d, \sum_{i=1}^{d} x_i = 1 \right\}$$

is the simplex on $\mathbb{R}^d$

Definition

$$\hbar_{\text{ent}}(x) = \begin{cases} \sum_{i=1}^{d} x_i \log(x_i) \text{ if } x \in \Delta_d \\ +\infty \text{ else} \end{cases}$$

is the entropic regularizer on Δ_d

I.1 Expression of the subdifferential of $\hbar_{\text{ent}}$

Let $x \in \Delta_d$ such that $x_i > 0$ for all $1 \leq i \leq d$.

Goal

$$\text{Prove that } \partial \hbar_{\text{ent}} = \left\{ (\log(x_i) + \lambda)_{1 \leq i \leq d} \right\}_{\lambda \in \mathbb{R}}$$

Proof:

① **First we show that** $\left\{ (\log(x_i) + \lambda)_{1 \leq i \leq d} \right\}_{\lambda \in \mathbb{R}} \subset \delta \hbar_{\text{ent}}$:

Let $y \in \left\{ (\log(x_i) + \lambda)_{1 \leq i \leq d} \right\}_{\lambda \in \mathbb{R}}$. We have therefore

$$y = (\log(x_i) + \lambda)_{1 \leq i \leq d} \text{ for } \lambda \in \mathbb{R}.$$

We want to show that

$$\forall x \in \Delta_d, \forall x' \in \mathbb{R}^d, \hbar_{\text{ent}}(x') \geq \hbar_{\text{ent}}(x) + \langle x | y - x' \rangle \tag{1}$$

We have two cases:

1. If $x' \notin \Delta_d$, it is trivial to see that (1) is valid since $\hbar_{\text{ent}}(x') = +\infty$ and the right term is finite.

2. If $x' \in \Delta_d$:
   We then have $\hbar_{\text{ent}}(x') = \sum_{i=1}^{d} {x_i}' \log({x_i}')$ and we have

$$\hbar_{\text{ent}}(x) + \langle x | y - x' \rangle = \sum_{i=1}^{d} x_i \log(x_i) + {x_i}' \log(x_i) - x_i \log(x_i) + \lambda({x_i}' - x_i)$$

   by using the definition of y. Since $\sum_{i=1}^{d} x_i = 1 = \sum_{i=1}^{d} {x_i}'$ we have the following:

$$\hbar_{\text{ent}}(x) + \langle x | y - x' \rangle = \sum_{i=1}^{d} {x_i}' \log(x_i)$$

   We notice that

$$\hbar_{\text{ent}}(x') - \hbar_{\text{ent}}(x) + \langle x | y - x' \rangle = \sum_{i=1}^{d} {x_i}' \log\left(\frac{{x_i}'}{x_i}\right)$$

   this sum can be viewed as a KL-divergence beteween the probability measure associated to x, x' since $x, x' \in \Delta_d$ and therefore we have

$$\hbar_{\text{ent}}(x') - \hbar_{\text{ent}}(x) + \langle x | y - x' \rangle = \sum_{i=1}^{d} {x_i}' \log\left(\frac{{x_i}'}{x_i}\right) \geq 0$$

   Wich allow us to conclude that $y \in \partial \hbar_{\text{ent}}$

We have shown that $\left\{ (\log(x_i) + \lambda)_{1 \leq i \leq d} \right\}_{\lambda \in \mathbb{R}} \subset \partial \hbar_{\text{ent}}$

② **We show that**

$$\partial \hbar_{\text{ent}} \subset \left\{ (\log(x_i) + \lambda)_{1 \leq i \leq d} \right\}_{\lambda \in \mathbb{R}}$$

let $y \in \partial \hbar_{\text{ent}}$, we know that $\hbar_{\text{ent}}$ is a regularizer on Δ_d by the course, so $\hbar_{\text{ent}}$ is a proper, lower semicontinuous and convex function.
Therefore by proposition 1.4.6 from the course, we have the following equivalent statements:

1. $$x \in \partial \hbar^*_{\text{ent}}(y)$$
2. $$y \in \partial \hbar_{\text{ent}}(x)$$
3. $$\langle y | x \rangle = \hbar_{\text{ent}}(x) + \hbar^*_{\text{ent}}(y)$$

   with $\hbar^*_{\text{ent}}$ the Legendre-Fenchel transform of $\hbar_{\text{ent}}$
4. $$x \in \arg\max_{x' \in \mathbb{R}^d} \{ \langle y | x' \rangle - \hbar_{\text{ent}}(x') \}$$
5. $$y \in \arg\max_{y' \in \mathbb{R}^d} \{ \langle y' | x \rangle - \hbar^*_{\text{ent}}(y') \}$$

Here we use points 2 and 4.Let $y \in \partial \hbar_{\text{ent}}$, by definition of $\hbar_{\text{ent}}$ we have :

$$x \in \arg\max_{x' \in \mathbb{R}^d} \{ \langle y | x' \rangle - \hbar_{\text{ent}}(x') \} \leftrightarrow x \in \arg\max_{x' \in \Delta^d} \{ \langle y | x' \rangle - \hbar_{\text{ent}}(x') \}$$

We use the condition $\sum_{i=1}^{d} x_i = 1$ with Lagrange multiplier to find the $\arg\max_{x' \in \Delta^d} \{\langle y | x' \rangle - \hbar_{\text{ent}}(x')\}$:

$$\begin{aligned}\mathcal{L}(x, \lambda) &= \langle y | x \rangle - \hbar_{\text{ent}}(x) + \lambda\left(-1 + \sum_{i=1}^{d} x_i\right) \\ &= \sum_{i=1}^{d} y_i x_i - x_i \log(x_i) + \lambda\left(-1 + \sum_{i=1}^{d} x_i\right)\end{aligned}$$

This function is clearly derivable, we then obtain the following partial derivatives and second partial derivatives:

$$\frac{\partial \mathcal{L}}{\partial x_i} = y_i - \log(x_i) - 1 + \lambda$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \sum_{i=1}^{d} x_i - 1 = 0$$

$$\frac{\partial^2 \mathcal{L}}{\partial x_i^2} = -\frac{1}{x_i}, \frac{\partial^2 \mathcal{L}}{\partial \lambda^2} = 0 \text{ and } \frac{\partial^2 \mathcal{L}}{\partial \lambda \partial x_i} = 0$$

Thefore by the hessian we know that the point $x \in \mathbb{R}^d$ where

$$\frac{\partial \mathcal{L}}{\partial x_i} = y_i - \log(x_i) - 1 + \lambda = 0$$

is the global maximum. We have therefore

$$\forall i \in \mathbb{R}^d, y_i = \log(x_i) + 1 - \lambda, \forall \lambda \in \mathbb{R}$$

Since this result is valid for all $\lambda \in \mathbb{R}$, we can directly conclude by proposition 1.4.6 that

$$\begin{aligned} y &\in \left\{(\log(x_i) + \lambda)_{1 \leq i \leq d}\right\}_{\lambda \in \mathbb{R}} \\ \leftrightarrow \partial \hbar_{\text{ent}} &\subset \left\{(\log(x_i) + \lambda)_{1 \leq i \leq d}\right\}_{\lambda \in \mathbb{R}}\end{aligned}$$

③ Since

$$\partial \hbar_{\text{ent}} \subset \left\{(\log(x_i) + \lambda)_{1 \leq i \leq d}\right\}_{\lambda \in \mathbb{R}} \text{ and } \left\{(\log(x_i) + \lambda)_{1 \leq i \leq d}\right\}_{\lambda \in \mathbb{R}} \subset \partial \hbar_{\text{ent}}$$

We eventually conclude that:

$$\partial \hbar_{\text{ent}} = \left\{(\log(x_i) + \lambda)_{1 \leq i \leq d}\right\}_{\lambda \in \mathbb{R}}$$

■

I.2 UMD iterates with $\hbar_{\text{ent}}$

Let $(u_t)_{t\geq 0}$ be a sequence in $\mathbb{R}^d$ and $((x_t, y_t))_{t\geq 0}$ be UMD iterates associated with $\hbar_{\text{ent}}$ and $(u_t)_{t\geq 0}$.

Goal

Prove that $(x_t)_{t\geq 0}$ is uniquely determined

Proof:
We know that with the previous notations, we have the following relations for UMD iterates in this case:

1. $\forall t \geq 0, y_t \in \partial\hbar_{\text{ent}}(x_t)$ with $\partial\hbar_{\text{ent}}(x_t) = \left\{\left(\log\left(x_{t_i}\right) + \lambda\right)_{1\leq i\leq d}\right\}_{\lambda\in\mathbb{R}}$
2. $\forall t \geq 0, x_{t+1} = \nabla h^*(y_t + u_t)$ and we know by the course that for h_{ent} we have

$$\nabla h^*(y) = \left(\frac{\exp(y_i)}{\sum_{j=1}^d \exp(y_j)}\right)_{1\leq i\leq d} \qquad \forall y \in \mathbb{R}^d$$

We will prove the desired property "$(x_t)_{t\geq 0}$ is uniquely determined" by induction:

① **<u>We show that x_1 is uniquely determined:</u>**
we have $y_1 = \left(\log\left(x_{0_i}\right) + \lambda\right)_{1\leq i\leq d}$ with $\lambda \in \mathbb{R}$.

$$\begin{aligned}
x_1 = \nabla h^*(y_0 + u_0) &= \left(\frac{\exp\left(y_{0_i} + u_{0_i}\right)}{\sum_{j=1}^d \exp\left(y_{0_j} + u_{0_j}\right)}\right)_{1\leq i\leq d} \\
&= \left(\frac{\exp\left(\log x_{0_i} + \lambda + u_{0_i}\right)}{\sum_{j=1}^d \exp\left(\log x_{0_j} + \lambda + u_{0_j}\right)}\right)_{1\leq i\leq d} \\
&= \left(\frac{\exp(\lambda)\exp\left(\log x_{0_i}\right)\exp\left(u_{0_i}\right)}{\exp(\lambda)\sum_{j=1}^d \exp\left(\log x_{0_j}\right)\exp\left(u_{0_j}\right)}\right)_{1\leq i\leq d} \\
&= \left(\frac{\exp\left(\log x_{0_i}\right)\exp\left(u_{0_i}\right)}{\sum_{j=1}^d \exp\left(\log x_{0_j}\right)\exp\left(u_{0_j}\right)}\right)_{1\leq i\leq d}
\end{aligned}$$

We note that this term is entirely determined, so x_1 is entirely determined. We do not have to check for x_0 value since it is a known value (initialization value). u_0 is also know.

② **<u>We suppose the property true at step $t \in \mathbb{N}$, we want to show that it is true at step $t+1$.</u>**
We use the exact same calculation but we change the indices and we use

that x_t is uniquely determined by hypothesis, and that we have access to u_t at time $t+1$.

$$\begin{aligned} x_{t+1} &= \nabla h^*(y_t + u_t) \\ &= \left(\frac{\exp\left(y_{t_i} + u_{t_i}\right)}{\sum_{j=1}^d \exp\left(y_{t_j} + u_{t_j}\right)} \right)_{1 \leq i \leq d} \\ &= \left(\frac{\exp\left(\log x_{t_i} + \lambda + u_{t_i}\right)}{\sum_{j=1}^d \exp\left(\log x_{t_j} + \lambda + u_{t_j}\right)} \right)_{1 \leq i \leq d} \\ &= \left(\frac{\exp\left(\log x_{t_i}\right) \exp\left(u_{t_i}\right)}{\sum_{j=1}^d \exp\left(\log x_{t_j}\right) \exp\left(u_{t_j}\right)} \right)_{1 \leq i \leq d} \end{aligned}$$

This term is uniquely determined since we know by hypothesis that x_t is uniquely determined and that u_t is also known.

The property is true for $t = 1$ (because x_0 can be set up manually) and is verified by induction for all $t \in \mathbb{N}$, therefore we can conclude that $(x_t)_{t \geq 0}$ is uniquely determined.

■

II Theoretical analysis

II.1 Introducing a new regularizer $\hbar$ based on $\hbar_{\text{ent}}$

Let $m, n \geq 1$. We have the following new definition:

Definition

$$\hbar : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$$
$$\hbar(a, b) = \hbar_{\text{ent}}(a) + \hbar_{\text{ent}}(b), (a, b) \in \mathbb{R}^m \times \mathbb{R}^n$$

where $\hbar_{\text{ent}}$ denotes both the entropic regularizer on Δ_m and Δ_n.

Goal

Prove that $\hbar$ is a regularizer on $\Delta_m \times \Delta_n$

Proof:
We have to prove the following points to ensure that $\hbar$ is a regularizer on $\Delta_m \times \Delta_n$:

1. prove that $\mathrm{cl}\ \mathrm{dom}\ \mathrm{h} = \Delta_m \times \Delta_n$:
   By definition of $\hbar_{\mathrm{ent}}$ on $\mathbb{R}^m$ and $\mathbb{R}^n$ (depending on which space we are operating), it is straighforward that $\mathrm{dom}\ \mathrm{h} = \Delta_m \times \Delta_n$.
   We also know that both Δ_m and Δ_n are closed convex set in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively so it follows that
   $$\mathrm{dom}\ h = \Delta_m \times \Delta_n = \mathrm{cl}\ \mathrm{dom}\ h$$
2. Prove that $\hbar$ is strictly convex on $\Delta_m \times \Delta_n$:
   We know that $\hbar_{\mathrm{ent}}$ is a regularizer on both Δ_m and Δ_n (depending on which space we are operating). Therefore since
   $$\hbar(a,b) = \hbar_{\mathrm{ent}}(a) + \hbar_{\mathrm{ent}}(b), \forall (a,b) \in \mathbb{R}^m \times \mathbb{R}^n$$
   and $\hbar_{\mathrm{ent}}$ is strictly convex on both Δ_m and Δ_n by definition of a regularizer, we have:
   Let $t \in [0,1], (a,b) \in \Delta_m \times \Delta_n, (c,d) \in \Delta_m \times \Delta_n$ with $(a,b) \neq (c,d)$
   $$\begin{aligned}\hbar(t*(a,b)+(1-t)*(c,d)) &= \hbar((ta+(1-t)c, tb+(1-t)d)) \\ &= \hbar_{\mathrm{ent}}(ta+(1-t)c) + \hbar_{\mathrm{ent}}(tb+(1-t)d) \\ < t\hbar_{\mathrm{ent}}(a) + (1-t)\hbar_{\mathrm{ent}}(c) + t\hbar_{\mathrm{ent}}(b) + (1-t)\hbar_{\mathrm{ent}}(d) &= t\hbar(a,b) + (1-t)\hbar(c,d)\end{aligned}$$
   So $\hbar$ is strictly convex on $\Delta_m \times \Delta_n$.
3. Prove that $\hbar$ is lower semicontinuous:
   Since $\hbar_{\mathrm{ent}}$ is a regularizer on Δ_m and Δ_n, it is also lower semicontinuous i.e $\forall a \in \mathbb{R}$, the set $\{x \in \mathbb{R}^m \text{ (or } \mathbb{R}^n), \hbar_{\mathrm{ent}}(x) < a\}$ is closed.
   By sum of two function lower semicontinuous, it is easy to see that $\hbar$ is also lower semicontinuous.
4. To have a regularizer on $\Delta_m \times \Delta_n$, we finally conclude by using the fact that $\Delta_m \times \Delta_n$ is a compact set since both Δ_m and Δ_n are closed and bounded convex set, therefore by Tychonoff's theorem $\Delta_m \times \Delta_n$ is compact.

We have proved that $\hbar$ is a regularizer on $\Delta_m \times \Delta_n$. ■

II.2 UMP iterates with $\hbar$

Let $G : \Delta_m \times \Delta_n \to \mathbb{R}^m \times \mathbb{R}^n$ be a monotone operator, $\gamma > 0$ and $((x_t, w_t, y_t, z_t))_{t \geq 0}$ a sequence of UMP iterates associated with regularizer $\hbar$, operator G and stepsize γ. G is furthermore supposed to be L-Lipschitz with respect to $\|\cdot\|_1$ and $\|\cdot\|_\infty$ i.e

$$\| G(x) - G(x') \|_\infty \leq L \cdot \| x - x' \|_1, \quad (x, x') \in (\Delta_m \times \Delta_n)^2$$

.

Goal

Prove that $(x_t)_{t\geq 0}$ and $(w_t)_{t\geq 0}$ are uniquely determined and derive a guarantee.

Proof:

① **We first prove that $(x_t)_{t\geq 0}$ and $(w_t)_{t\geq 0}$ are uniquely determined.**

By the definition of UMP iterates, we have notably the following properties:

1. $\forall t \geq 0, y_t \in \partial \hbar(x_t)$
2. $x_{t+1} = \nabla h^*(y_t - \gamma G(w_t))$
3. $\forall t \geq 0, z_t \in \partial \hbar(x_t)$
4. $w_t = \nabla h^*(z_t - \gamma G(x_t))$

For the following we write

$$(x_t) = (a_t, b_t) \in \mathbb{R}^m \times \mathbb{R}^n \forall t \geq 0$$
$$(w_t) = (c_t, d_t) \in \mathbb{R}^m \times \mathbb{R}^n \forall t \geq 0$$

We remind that

$$\hbar : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$$
$$\hbar(a, b) = \hbar_{\text{ent}}(a) + \hbar_{\text{ent}}(b), (a, b) \in \mathbb{R}^m \times \mathbb{R}^n$$

so we have

$$\partial \hbar(x) = \partial h(a, b) = \partial \hbar_{\text{ent}}(a) \times \partial \hbar_{\text{ent}}(b)$$
$$= \left\{(\log(a_i) + \lambda_1)_{1\leq i\leq m}\right\}_{\lambda_1 \in \mathbb{R}} \times \left\{(\log(b_i) + \lambda_2)_{1\leq i\leq n}\right\}_{\lambda_2 \in \mathbb{R}} \forall (a, b) \in \Delta^m \times \Delta^n$$

By using the links between $\hbar$ and $\hbar_{\text{ent}}$, we obtain also :

$$\forall (a, b) \in \mathbb{R}^m \times \mathbb{R}^n, \nabla \hbar^*(a, b) = (\nabla \hbar^*_{\text{ent}}(a), \nabla \hbar^*_{\text{ent}}(b))$$
$$= \left(\left(\frac{\exp(a_i)}{\sum_{j=1}^m \exp(a_j)}\right)_{1\leq i<m}, \left(\frac{\exp(b_i)}{\sum_{j=1}^n \exp(b_j)}\right)_{1\leq i<n}\right)$$

We can therefore prove that $(x_t)_{t\geq 0}$ and $(w_t)_{t\geq 0}$ are uniquely determined by double induction and using a similar reasoning that for II.2 for both terms.

Firstly w_0 and x_1 are uniquely determined (x_0 is the initialization value and therefore known) because each component of both z_0 and y_0 is uniquely determined (same calculations as for II.2)

Let the property "$(x_{t+1})_{t\geq 0}$ and $(w_t)_{t\geq 0}$ are uniquely determined" true up to rank t, we want to show that it is true for rank $t+1$:
We have

$$w_{t+1} = \nabla h^*(z_{t+1} - \gamma G(x_{t+1})$$

and we know that $z_{t+1} \in \partial \hbar(x_{t+1})$ so

$$z_{t+1} = \left(\left(\log\left(a_{(t+1)_i}\right) + \lambda_1\right)_{1\leq i\leq m}, \left(\log\left(b_{(t+1)_i}\right) + \lambda_2\right)_{1\leq i\leq n}\right), (\lambda_1, \lambda_2) \in \mathbb{R}^2$$

Since $G : \Delta_m \times \Delta_n \to \mathbb{R}^m \times \mathbb{R}^n$ is a monotonous operator, we can write

$$G(x_{t+1}) = G(a_{t+1}, b_{t+1}) = ({a'}_{t+1}, {b'}_{t+1}) \text{ with } ({a'}_{t+1}, {b'}_{t+1}) \in \mathbb{R}^m \times \mathbb{R}^n$$

Since x_t is uniquely determined by induction hypothesis and G is a known function, we have:

$$w_{t+1} = \nabla h^*(z_{t+1} - \gamma G(x_{t+1}))$$

$$= \left(\left(\frac{\exp\left(\log\left(a_{(t+1)_i}\right) + \lambda_1 - \gamma {a'}_{(t+1)_i}\right)}{\sum_{j=1}^m \exp\left(\log\left(a_{(t+1)_j}\right) + \lambda_1 - \gamma {a'}_{(t+1)_j}\right)}\right)_{1\leq i<m}, \left(\frac{\exp\left(\log\left(b_{(t+1)_i}\right) + \lambda_2 - \gamma {b'}_{(t+1)_i}\right)}{\sum_{j=1}^n \exp\left(\log\left(b_{(t+1)_j}\right) + \lambda_2 - \gamma {b'}_{(t+1)_j}\right)}\right)_{1\leq i<n}\right)$$

$$= \left(\left(\frac{\exp\left(\log\left(a_{(t+1)_i}\right) - \gamma {a'}_{(t+1)_i}\right)}{\sum_{j=1}^m \exp\left(\log\left(a_{(t+1)_j}\right) - \gamma {a'}_{(t+1)_j}\right)}\right)_{1\leq i<m}, \left(\frac{\exp\left(\log\left(b_{(t+1)_i}\right) - \gamma {b'}_{(t+1)_i}\right)}{\sum_{j=1}^n \exp\left(\log\left(b_{(t+1)_j}\right) - \gamma {b'}_{(t+1)_j}\right)}\right)_{1\leq i<n}\right)$$

We recognized only fully and uniquely determined terms.

The same approach to prove that x_{t+2} is uniquely determined, based on the fact that w_{t+1} is uniquely determined and with the same type of intermediate arguments and calculations.

By double induction we have shown that $x_{(t)_{t\geq 0}}$ and $(w_t)_{t\geq 0}$ are uniquely determined.

② **<u>We want to derive a guarantee for these iterates.</u>**

First, we know that G is L-lipschitz continuous for $\|\cdot\|_1$ by assumption.
Now we show that $\hbar$ is strongly convex for $\|\cdot\|_1$ We know that the entropic regularizer $\hbar_{\text{ent}}$ is 1-strongly convex with respect to the l_1 norm on the probability simplex.

Let $(a_1, b_1), (a_2, b_2) \in \Delta_m \times \Delta_n$ and $\alpha \in [0, 1]$. Then,

$$\hbar(\alpha(a_1, b_1) + (1-\alpha)(a_2, b_2)) = \hbar_{\text{ent}}(\alpha a_1 + (1-\alpha)a_2) + \hbar_{\text{ent}}(\alpha b_1 + (1-\alpha)b_2)$$

$$\leq \alpha\hbar_{\text{ent}}(a_1) + (1-\alpha)\hbar_{\text{ent}}(a_2) - \frac{1}{2}\alpha(1-\alpha)\ \|a_1 - a_2\|_1^2$$

$$+\alpha\hbar_{\text{ent}}(b_1) + (1-\alpha)\hbar_{\text{ent}}(b_2) - \frac{1}{2}\alpha(1-\alpha)\ \|b_1 - b_2\|_1^2$$

$$= \alpha\hbar(a_1, b_1) + (1-\alpha)\hbar(a_2, b_2) - \frac{1}{2}\alpha(1-\alpha)(\|a_1 - a_2\|_1^2 + \|b_1 - b_2\|_1^2).$$

Since the l_1 norm on $\mathbb{R}^m \times \mathbb{R}^n$ is defined as $\|(a,b)\|_1 = \|a\|_1 + \|b\|_1$, we have

$$\|(a_1,b_1) - (a_2,b_2)\|_1 = \|a_1 - a_2\|_1 + \|b_1 - b_2\|_1,$$

and thuss

$$\|a_1 - a_2\|_1^2 + \|b_1 - b_2\|_1^2 \geq \frac{1}{2} \|(a_1,b_1) - (a_2,b_2)\|_1^2.$$

because $\forall (c,d) \in \mathbb{R}^2, c^2 + d^2 - \frac{1}{2}(c+d)^2 = \left(\frac{1}{\sqrt{2}}c - \frac{1}{\sqrt{2}}d\right)^2 > 0.$

Combining these inequalities, we obtain

$$\hbar(\alpha(a_1,b_1) + (1-\alpha)(a_2,b_2)) \leq \alpha\hbar(a_1,b_1) + (1-\alpha)\hbar(a_2,b_2) - \frac{1}{4}\alpha(1-\alpha)\|(a_1,b_1) - (a_2,b_2)\|_1^2,$$

which shows that $\hbar$ is $\frac{1}{2}$-strongly convex with respect to the l_1 norm on its domain $\Delta_m \times \Delta_n$.

We can now apply proposition 8.3.3 with $\hbar$ which give us the following guarantee for step size $\gamma = \frac{K}{L}$:

$$\forall T \geq 0, \forall x \in \Delta_m \times \Delta_n, \langle G(x) | \overline{w}_T - x \rangle \leq 2 \cdot \frac{L \cdot D_\hbar(x, x_0, y_0)}{T+1}$$

$$\text{where } \overline{w}_T = \frac{1}{T+1} \sum_{t=0}^{T} w_t \text{ and with}$$

$$D_\hbar(x, x_0, y_0) = \hbar(x) - \hbar(x_0) - \langle y_0 | x_0 - x \rangle \text{ the Bregman divergence.}$$

■

II.3 solving a two player zero-sum game

Goal

1°/ Write the previous iterates in the case of solving a two-player zero-sum game and write the corresponding guarantee.
2°/ Compare these iterates and guarantee with the optimistic exponential weights algorithm.

Proof:
① We write the previous iterates in the case of solving a two-player zero-sum game:
In such a situation, we have the following expression for the Lipschitz continuous on $\Delta_m \times \Delta_n$ monotone operator G:

$$G : \Delta^m \times \Delta^n \to \mathbb{R}^m \times \mathbb{R}^n$$
$$a, b \to (-Ab, A^T a) \text{ with } A \in \mathbb{R}^{m \times n}$$

We then use the previous relations between the component of the UMP iterates to obtain the following irates in the case of solving a two-player zero-sum game with writing $x_t = (a_t, b_t) \in \Delta_m \times \Delta_n$:

1. $\forall t \geq 0, y_t = \left(y_{a_t}, y_{b_t}\right) = \left(\log(a_i)_{1 \leq i \leq m}, \log(b_i)_{1 \leq i \leq n}\right)$
2. $$G_t = \gamma G(x_t) = (-\gamma A b_t, \gamma A^T a_t) = (p_t, q_t)$$
3. $$a_t' = \left(\frac{\exp\left(\log(a_i) - p_{t_i}\right)}{\sum_{j=1}^m \left(\exp\left(\log(a_j) - p_{t_j}\right)\right)}\right)_{1 \leq i \leq m}$$
$$b_t' = \left(\frac{\exp\left(\log(b_i) - q_{t_i}\right)}{\sum_{j=1}^m \left(\exp\left(\log(b_j) - q_{t_j}\right)\right)}\right)_{1 \leq i \leq n}$$
$$w_t = (a_t', b_t') \in \Delta_m \times \Delta_n$$
4. $$\mathcal{G}_t = \gamma G(w_t) = (-\gamma A b_t', \gamma A^T a_t') = (p_t', q_t')$$
5. $$a_{t+1} = \left(\frac{\exp\left(\log(a_i) - a_{t_i}'\right)}{\sum_{j=1}^m \left(\exp\left(\log(a_j) - a_{t_j}'\right)\right)}\right)_{1 \leq i \leq m}$$
$$b_{t+1} = \left(\frac{\exp\left(\log(b_i) - b_{t_i}'\right)}{\sum_{j=1}^m \left(\exp\left(\log(b_j) - b_{t_j}'\right)\right)}\right)_{1 \leq i \leq n}$$
$$x_{t+1} = (a_{t+1}, b_{t+1}) \in \Delta_m \times \Delta_n$$

For $y_0 = (0_{\mathbb{R}^m}, 0_{\mathbb{R}^n})$ we have

$$D_\hbar(x, x_0, y_0) = \hbar(x) - \hbar(x_0) = \hbar_{\text{ent}}(a) + \hbar_{\text{ent}}(b) - \hbar_{\text{ent}}(a_0) - \hbar_{\text{ent}}(b_0)$$
$$\leq \log(m) + \log(n) \text{ by property of } \hbar_{\text{ent}}$$

We know that G is $L = \max(\|A\|_\infty, \|A^T\|_\infty)$ Lispchitz continuous and $\hbar$ is $K = \frac{1}{2}$ strongly convex for $\|\cdot\|_1$.

We use the guarantee found in the previous question and obtain the following guarantee by using similar tools to the ones used to prove Proposition 9.4.2:

$$\forall T \geq 0, \forall x_t \in \Delta_m \times \Delta_n, \begin{cases} \left\langle -Ab_t \middle| \overline{a'}_T - a_t \right\rangle \leq \frac{2L(\log(m) + \log(n))}{T+1} \\ \left\langle Aa_t \middle| \overline{b'}_T - b_t \right\rangle \leq \frac{2 \cdot L \cdot (\log(m) + \log(n))}{T+1} \end{cases}$$

with $\overline{a'}_T = \frac{1}{T+1}\sum_{t=0}^{T} a'_t$, $\overline{b'}_T = \frac{1}{T+1}\sum_{t=0}^{T} b'_t$ and for step-size $\gamma = \frac{K}{L}$

It is therefore possible to have approximate weak solution for a game at the rate $\frac{1}{T}$.

In the case of two-player zero-sum games, we have the following guarantee for the optimistic exponential weight algorithm:

$$\delta_A\left(\overline{a}_T, \overline{b}_T\right) \leq \frac{\|A\|_\infty\ (\log(m) + \log(n) + 2)}{T+1}$$

with $\overline{a}_T = \frac{1}{T+1}\sum_{t=0}^{T} a_t$ and $\overline{b}_T = \frac{1}{T+1}\sum_{t=0}^{T} b_t$

This guarantee is at the same rate $\frac{1}{T}$ than the one for the studied algorithm but it is directly on the gap, which is the quantity that interest us the most. Therefore this bound is more interesting to justify that we reach the Nash equilibrium. ■

III Numerical Experiments and comparison

Numerical experiments are proposed in order to empirically compare this algorithm with the other seen during the course.
These experiments are conducted exclusivly on two-player zero-sum games. The proposed algorithm that we call here "extrapolation" is compared to the exponential weights, RM and RM+ algorithm.

The python script and the affiliated notebook are available respectivly in the Pierron_algorithms.py script and Pierron_experiments.ipynb notebook.
All functions are properly documentated.

The following python modules are needed :

1. **matplotlib**
2. **numpy**
3. **itertools**
4. **tqdm**

The code is entirely reproductible but can take up to 2 hour to fully run due to the high number of extended experiments. If the goal is to only see the results and the plots, please do not run the notebook when opening it.

The conducted experiments show that the extrapolation algorithm requires a small grid search in order to find the correct factor γ enabling a fast convergence.
We do this to compare the performance between the proposed γ value and some other step size that might be relevant.

Once a good candidate for γ is found, we observe that this algorithm is very efficient and among the two fastest algorithms to converge with RM+ among the proposed panel.
We also compare the extrapolatio algorithm with $\gamma = \frac{K}{L}$. We found a similar result

as when we use the best gamma found with a grid search. This confirms that this step-size is a good choice and allow to not use a computationnally grid-search.

We see that for a reasonable amount of time (around 4000-5000 iterations, which takes between 10 and 25 sec), all the algorithms tends to very small regret values (less than 0.01 for most games).

nevertheless, RM and RM+ seems to be more robust to complex games where the payoff matrix become larger and with more abstract values.

This fact tends to say that it is better to use RM+ or RM for complex payoff matrix and the extrapolation algorithms for simpler payoff matrix.

As a bonus, the theoretical upper bounds for RM, RM+ and EW are represented for different games, illustrating the good behavior of these algorithms.