Exercices
Regret learning in games
Université Paris–Saclay

Let $m, n \geqslant 1$ be integers, $A \in \mathbb{R}^{m\times n}$, $\|\cdot\|_{(m)}$ and $\|\cdot\|_{(n)}$ norms on $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively and denote $\|\cdot\|_{(m*)}$ and $\|\cdot\|_{(n*)}$ their respective dual norms, $h^{(m)}$ and $h^{(n)}$ regularizers on Δ_m and Δ_n, $((a_t, b_t, y_t, z_t))_{t\geqslant 1}$ a sequence in $\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ and $K^{(m)}, K^{(n)}, \eta, \eta' > 0$. We assume that

- $((a_t, y_t))_{t\geqslant 1}$ is a sequence of strict UMD iterates associated with regularizer $h^{(m)}$ and dual increments $(\eta A(2b_t - b_{t-1}))_{t\geqslant 1}$ (with convention $b_0 = 0$),
- $((b_t, z_t))_{t\geqslant 1}$ is a sequence of strict UMD iterates associated with regularizer $h^{(n)}$ and dual increments $(\eta' A^\top a_{t+1})_{t\geqslant 1}$,
- $h^{(m)}$ is $K^{(m)}$-strongly convex for $\|\cdot\|_{(m)}$,
- $h^{(n)}$ is $K^{(n)}$-strongly convex for $\|\cdot\|_{(n)}$.

Let $T \geqslant 1$, $a \in \Delta_m$ and $b \in \Delta_n$.

1) Informally compare the above iterates with optimistic regret learning in two-player zero-sum games.

2) Recall the bound on

$$\sum_{t=1}^{T} \langle \eta A b_t, a - a_t \rangle .$$

3) Prove that

$$\sum_{t=1}^{T} \langle \eta' A^\top a_{t+1}, b - b_{t+1} \rangle \leqslant D_{h^{(n)}}(b, b_1;\ z_1) - \frac{K^{(n)}}{2} \sum_{t=1}^{T} \|b_{t+1} - b_t\|_{(m)}^2 .$$

4) Deduce a convergence guarantee in the context of solving the two-player zero-sum game associated with matrix A.

5) Express $(a_t)_{t\geqslant 1}$ and $(b_t)_{t\geqslant 1}$ in the special case of $h^{(m)}$ and $h^{(n)}$ being the entropic regularizers on Δ_m and Δ_n respectively. Write the corresponding convergence guarantee.

6) Conduct numerical experiments to compare the performance of the algorithm from the last question with the optimistic exponential weights algorithm, the regular exponential weights algorithm, RM, RM+.