Evaluation
Online learning
links with optimization and games
Université Paris–Saclay

solving games: extrapolation vs optimism

Let $d \geqslant 1$ and let h_{ent} be the entropic regularizer on the simplex Δ_d.

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

$$\partial h_{\mathrm{ent}}(x) = \left\{ (\log x_i + \lambda)_{1 \leqslant i \leqslant d} \right\}_{\lambda \in \mathbb{R}}.$$

2) Let $(u_t)_{t \geqslant 0}$ be a sequence in $\mathbb{R}^d$ and $((x_t, y_t))_{t \geqslant 0}$ be UMD iterates associated with h_{ent} and $(u_t)_{t \geqslant 0}$. Prove that $(x_t)_{t \geqslant 0}$ is uniquely determined.

3) Let $m, n \geqslant 1$ and let $h : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be defined as

$$h(a, b) = h_{\mathrm{ent}}(a) + h_{\mathrm{ent}}(b), \qquad (a, b) \in \mathbb{R}^m \times \mathbb{R}^n,$$

where h_{ent} denotes both the entropic regularizer on Δ_m and on Δ_n. Prove that h is a regularizer on $\Delta_m \times \Delta_n$.

4) Let $G : \Delta_m \times \Delta_n \to \mathbb{R}^m \times \mathbb{R}^n$ be a monotone operator, $\gamma > 0$, and $((x_t, w_t, y_t, z_t))_{t \geqslant 0}$ a sequence of UMP iterates (see definition in Section 8.3) associated with regularizer h, operator G and step-size γ. Prove that $(x_t)_{t \geqslant 0}$ and $(w_t)_{t \geqslant 0}$ are uniquely determined. Derive a guarantee (in the case of this regularizer h).

5) Write the iterates from the last question in the special case of solving a two-player zero-sum game, and write the corresponding guarantee. Compare these iterates and guarantee with the optimistic exponential weights algorithm.

6) Perform numerical experiments in the context of solving two-player zero-sum games to compare the performance of the above algorithms with the (regular) exponential weights algorithm, RM and RM+.