Algorithmic game theory – Tutorial 7*

November 29th, 2024

1 Swap regret and internal regret

There are N available actions $X = \{1, \ldots, N\}$ and at each time step t the online algorithm A selects a probability distribution $p^t = (p_1^t, \ldots, p_N^t)$ over X. After the distribution p^t is chosen at time step t, the adversary chooses a loss vector $\ell^t = (\ell_1^t, \ldots, \ell_N^t) \in [-1, 1]^N$, where the number ℓ_i^t is the loss of action i in time t. The algorithm A then experiences loss $\ell_A^t = \sum_{i=1}^N p_i^t \ell_i^t$. After T steps, the loss of action i is $L_i^T = \sum_{t=1}^T \ell_i^t$ and the loss of A is $L_A^T = \sum_{t=1}^T \ell_A^t$.

Given the sequence $(p^t)_{t=1}^T$ of the probability distributions used by A and a modification rule F, we define a *modified sequence* $(f^t)_{t=1}^T = (F^t(p^t))_{t=1}^T$, where $f^t = (f_1^t, \ldots, f_N^t)$ and $f_i^t = \sum_{j\colon F^t(j)=i} p_j^t$. The *loss of the modified sequence* is $L_{A,F}^T = \sum_{t=1}^T \sum_{i=1}^N f_i^t \ell_i^t$. Given a sequence ℓ^t of loss vectors, the *regret of A with respect to $\mathcal{F}$* is $R_{A,\mathcal{F}}^T = \max_{F\in\mathcal{F}} \left\{L_A^T - L_{A,F}^T\right\}$. The external regret of A is then $R_{A,\mathcal{F}^{ex}}^T$ for $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of N modification rules $F_i = (F_i^t)_{t=1}^T$, where each F_i^t always outputs action i. The *internal regret* of A is $R_{A,\mathcal{F}^{in}}^T$ for the set $\mathcal{F}^{in} = \{F_{i,j} \colon (i,j) \in X \times X, i \neq j\}$ of $N(N-1)$ modification rules $F_{i,j} = (F_{i,j}^t)_{t=1}^T$, where, for every time step t, $F_{i,j}^t(i) = j$ and $F_{i,j}^t(i') = i'$ for each $i' \neq i$. The *swap regret* of A is $R_{A,\mathcal{F}^{sw}}^T$ for the set $\mathcal{F}^{sw}$ of all modification rules $F\colon X \to X$.

Exercise 1. *Show that the swap regret is at most N times larger than the internal regret.*

Exercise 2. *Show that a probability distribution p is a correlated equilibrium, that is,*

$$\mathbb{E}_{a\sim p}[C_i(a) \mid a_i] \leq \mathbb{E}_{a\sim p}[C_i(a_i'; a_{-i}) \mid a_i]$$

for every player $i \in P$ and all $a_i, a_i' \in A_i$ if and only if

$$\mathbb{E}_{a\sim p}[C_i(a)] \leq \mathbb{E}_{a\sim p}[C_i(F(a_i); a_{-i})]$$

for each player $i \in P$ and each modification rule $F\colon A_i \to A_i$. What if we consider modification rules only from $\mathcal{F}^{in}$?

Hint: It might be useful to see that $\mathbb{E}_{a\sim p}[C_i(a)] = \sum_{a_i\in A_i} P(a_i) \cdot \mathbb{E}_{a\sim p}[C_i(a) \mid a_i]$, where $P(a_i) = \sum_{a_{-i}\in A_{-i}} p(a_i; a_{-i})$ is the probability that player i is recommended to play a_i.

Exercise 3. *Let $G = (P, A, C)$ be a normal-form game of n players, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$. Assume that after T steps of the No-internal-regret dynamics every player $i \in P$ time-averaged internal regret at most ε. Let $p^t = \prod_{i=1}^n p_i^t$ be a product probability distribution of the mixed strategies used by the players in step t and set $p = \frac{1}{T}\sum_{t=1}^T p^t$. Show that p is ε-correlated equilibrium.*

Hint: It might help to solve Exercise 2 first.

Exercise 4. *Show an example with $N = 3$ where the external regret is zero and the swap regret goes to infinity with T.*

Clarification: you need to choose only a sequence of actions $a^1, \ldots, a^T$, $a^i \in X = \{1, 2, 3\}$, and a loss sequence $\ell_a^1, \ldots, \ell_a^T$ for every $a \in X$.

*Information about the course can be found at http://kam.mff.cuni.cz/~balko/