Algorithmic game theory – Tutorial 4*

December 3, 2020

1 Regret minimization

There are N available actions $X = \{1, \dots, N\}$ and at each time step t the online algorithm A selects a probability distribution $p^t = (p^t_1, \dots, p^t_N)$ over X. After the distribution p^t is chosen at time step t, the adversary chooses a loss vector $\ell^t = (\ell^t_1, \dots, \ell^t_N) \in [-1, 1]^N$, where the number ℓ^t_i is the loss of action i in time t. The algorithm A then experiences loss $\ell^t_A = \sum_{i=1}^{N} p^t_i \ell^t_i$. After T steps, the loss of action i is $L^T_i = \sum_{t=1}^{T} \ell^t_i$ and the loss of A is $L^T_A = \sum_{t=1}^{T} \ell^t_A$. The *external regret* of A is $R^T_A = \max_{i \in X}\{L^T_A - L^T_i\}$.

Exercise 1. *Let A be an algorithm with parameter $\eta \in (0, 1/2]$ and with external regret at most $\alpha/\eta + \beta\eta T$ for some constants α, β (that may depend on the number N of actions). We showed that choosing $\eta = \sqrt{\alpha/(T\beta)}$ minimizes the bound. Modify this algorithm so that we obtain an external regret bound that is at most $O(1)$-times larger that the original bound for any T. In particular, you cannot run A with a parameter η that depends on T.*

Hint: Partition the set $\{1, \dots, T\}$ into suitable intervals I_m for $m = 0, 1, 2, \dots$ and run A with a suitable parameter η_m in every step from I_m.

Given the sequence $(p^t)^T_{t=1}$ of the probability distributions used by A and a modification rule F, we define a *modified sequence* $(f^t)^T_{t=1} = (F^t(p^t))^T_{t=1}$, where $f^t = (f^t_1, \dots, f^t_N)$ and $f^t_i = \sum_{j\colon F^t(j)=i} p^t_j$. The *loss of the modified sequence* is $L^T_{A,F} = \sum_{t=1}^{T} \sum_{i=1}^{N} f^t_i \ell^t_i$. Given a sequence ℓ^t of loss vectors, the *regret of A with respect to $\mathcal{F}$* is $R^T_{A,\mathcal{F}} = \max_{F \in \mathcal{F}} \left\{L^T_A - L^T_{A,F}\right\}$. The external regret of A is then $R^T_{A,\mathcal{F}^{ex}}$ for $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of N modification rules $F_i = (F^t_i)^T_{t=1}$, where each F^t_i always outputs action i. The *internal regret* of A is $R^T_{A,\mathcal{F}^{in}}$ 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 $F_{i,j} = (F^t_{i,j})^T_{t=1}$, where, for every time step t, $F^t_{i,j}(i) = j$ and $F^t_{i,j}(i') = i'$ for each $i' \neq i$. The *swap regret* of A is $R^T_{A,\mathcal{F}^{sw}}$ for the set $\mathcal{F}^{sw}$ of all modification rules $F \colon X \to X$.

Exercise 2. *Show that the swap regret is at most N times larger than the internal regret.*

Exercise 3. *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 $a^1, \dots, a^T$, $a^i \in X = \{1, 2, 3\}$, and a loss sequence $\ell^1_a, \dots, \ell^T_a$ for every $a \in X$.

For a normal-form game $G = (P, A, C)$ of n players, a probability distribution $p(a)$ on A is a *correlated equilibrium* in G if $\sum_{a_{-i} \in A_{-i}} C_i(a_i; a_{-i})p(a_i; a_{-i}) \leq \sum_{a_{-i} \in A_{-i}} C_i(a'_i; a_{-i})p(a_i; a_{-i})$ for every player $i \in P$ and all $a_i, a'_i \in A_i$. A probability distribution $p(a)$ on A is a *coarse correlated equilibrium* in G if $\sum_{a \in A} C_i(a)p(a) \leq \sum_{a \in A} C_i(a'_i; a_{-i})p(a)$ for every player $i \in P$ and every $a'_i \in A_i$.

Exercise 4. *Show formally that every correlated equilibrium is a coarse correlated equilibrium.*

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