Algorithmic game theory – Tutorial 6*

November 18th, 2024

1 Regret minimization

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 1. *Show formally that every correlated equilibrium is a coarse correlated equilibrium.*

Exercise 2. *Compute all coarse correlated equilibria in the Prisoner's dilemma game.*

	T	S
T	(2,2)	(0,3)
S	(3,0)	(1,1)

Table 1: The game from Exercise 2

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 3. *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.

Exercise 4 (*). *Prove the following statements about lower bounds on the external regret.*

(a) *For integers N and T with $T < \lfloor \log_2 N \rfloor$, there exists a stochastic generation of losses such that, for every online algorithm A, we have $\mathbb{E}[L^T_A] \geq T/2$ and yet $L^T_{min} = 0$.*

(b) *In the case of $N = 2$ actions, there exists a stochastic generation of losses such that, for every online algorithm A, we have $\mathbb{E}[L^T_A - L^T_{min}] \geq \Omega(\sqrt{T})$.*

Exercise 5. *Consider the following setting in which the agent A tries to learn the setup in an adversary environment while using information given to him by a set S_0 of N experts. The setting proceeds in a sequence of steps $t = 1, \dots, T$. In every step t, the environment picks $y_t \in \{0, 1\}$, which is unknown to A and to the experts, and each expert i gives a recommendation $f_{i,t} \in \{0, 1\}$ to A. The agent A then makes prediction $z_t \in \{0, 1\}$ based on the experts' advice and then sees y_t. The goal of A is to minimize the number $M^T(A)$ of steps t in which $z_t \neq y_t$.*

(a) *Assume that, in each step t, the agent A selects z_t to be the majority vote of all experts from S_{t-1} and, after seeing y_t, he lets S_t be the number of agents $i \in S_{t-1}$ with the right guess $f_{i,t} = z_t$. Also, assume that there is a perfect expert who always guesses right. Prove that then $M^T(A) \leq \log_2 N$.*

(b) *Modify the above algorithm of the agent so that $M^T(A) \leq O((m + 1) \log_2 N)$ when the best expert makes $m \geq 0$ mistakes.*

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