AGT 7th tutorial*

7 December 2023

1 Regret minimization (external, internal, swap)

There are N available actions $X = \{1, \dots, N\}$ and at each time step t the online algorithm A selects a probability distribution $p^t = (p_1^t, \dots, p_N^t)$ over X. After the distribution p^t is chosen at time step t, the adversary chooses a loss vector $\ell^t = (\ell_1^t, \dots, \ell_N^t) \in [-1, 1]^N$, 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 $L_A^T = \sum_{t=1}^T \ell_A^t$. The *external regret* of A is $R_A^T = \max_{i \in X}\{L_A^T - L_i^T\}$.

Exercise 1 (*). *Prove the following statements about lower bounds for external regret.*

(a) *Let N a T be natural numbers such that N is power of 2 and $T < \log_2 N$, then there exists choice of random vectors of losses from $\{0, 1\}$ such that every online algorithm A satisfy $\mathbb{E}[L_A^T] \geq T/2$ and $L_{min}^T = 0$*

(b) *For $N = 2$, there exists a choice of random vectors of losses from $\{0, 1\}$ such that every online algorithm A satisfy $\mathbb{E}[L_A^T - L_{min}^T] \geq \Omega(\sqrt{T})$.*

Exercise 2. *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 $(p^t)_{t=1}^T$ 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 $f^t = (f_1^t, \dots, f_N^t)$ 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 3. *Show that the swap regret is at most N times larger than the internal regret.*

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 $a^1, \dots, a^T$, $a^i \in X = \{1, 2, 3\}$, and a loss sequence $\ell_a^1, \dots, \ell_a^T$ 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 $a_i' \in A_i$.

*Informace o cvičení naleznete na http://kam.mff.cuni.cz/˜ryzak/