Algoritmická teorie her – příklady na 5. cvičení*

28. listopadu 2022

1 Regret minimalizace

Máme množinu $X = \{1, \dots, N\}$ s N akcemi a v každém kroce t online algoritmus A vybere pravděpodobnostní rozdělení $p^t = (p_1^t, \dots, p_N^t)$ na X. Poté, co je rozdělení p^t vybráno v kroce t, nepřátelské prostředí zvolí ztráty $\ell^t = (\ell_1^t, \dots, \ell_N^t) \in [-1,1]^N$, kde ℓ_i^t je ztrátou za akci i v čase t. Algoritmus A poté prodělá ztrátu $\ell_A^t = \sum_{i=1}^N p_i^t \ell_i^t$. Po T krocích je ztráta akce i rovna $L_i^T = \sum_{t=1}^T \ell_i^t$ a ztráta algoritmu A je $L_A^T = \sum_{t=1}^T \ell_A^t$. *Externí regret* algoritmu A je $R_A^T = \max_{i \in X}\{L_A^T - L_i^T\}$.

Exercise 1. *Nechť je A algoritmus s parametrem $\eta \in (0, 1/2]$ a s externím regretem nanejvýš $\alpha/\eta + \beta\eta T$ pro nějaké konstanty α, β (které mohou záviset na počtu akcí N). Ukázali jsme, že volba $\eta = \sqrt{\alpha/(T\beta)}$ minimalizuje odhad na regret. Modifikujte tento algoritmus tak, abychom dostali odhad na regret, který je nanejvýš $O(1)$-krát větší než původní odhad pro každé T. Tedy nechceme, aby parametr η závisel na T.*

Nápověda: Rozdělte množinu $\{1, \dots, T\}$ na vhodně intervaly I_m pro $m = 0, 1, 2, \dots$ a pusťte A s vhodným parametrem η_m na všech krocích z I_m.

Pro posloupnost $(p^t)_{t=1}^T$ pravděpodobnostních rozdělení vystoupených algoritmem A a pro modifikační pravidlo $F \colon X \to X$, definujeme *modifikovanou posloupnost* $(f^t)_{t=1}^T = (F^t(p^t))_{t=1}^T$, kde $f^t = (f_1^t, \dots, f_N^t)$ a $f_i^t = \sum_{j\colon F^t(j)=i} p_j^t$. *Ztráta modifikované posloupnosti* je pak $L_{A,F}^T = \sum_{t=1}^T \sum_{i=1}^N f_i^t \ell_i^t$. Pro posloupnost ℓ^t ztrátových vektorů je *regret algoritmu A vzhledem k $\mathcal{F}$* rovný $R_{A,\mathcal{F}}^T = \max_{F \in \mathcal{F}} \{L_A^T - L_{A,F}^T\}$. Externí regret algoritmu A je potom $R_{A,\mathcal{F}^{ex}}^T$, kde $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ obsahuje modifikační pravidla $F_i = (F_i^t)_{t=1}^T$ taková, že F_i^t vždy vrací akci i. *Interní regret* algoritmu A je $R_{A,\mathcal{F}^{in}}^T$ pro množinu $\mathcal{F}^{in} = \{F_{i,j} \colon (i,j) \in X \times X, i \neq j\}$ of $N(N-1)$ modifikačních pravidel $F_{i,j} = (F_{i,j}^t)_{t=1}^T$, kde v každém kroce t, $F_{i,j}^t(i) = j$ a $F_{i,j}^t(i') = i'$ pro každé $i' \neq i$. *Swap regret* algoritmu A je $R_{A,\mathcal{F}^{sw}}^T$ pro množinu $\mathcal{F}^{sw}$ všch modifikačních pravidel $F \colon X \to X$.

Exercise 2. *Ukažte, že swap regret je vždy nanejvýš N-krát tak velký jako interní regret.*

Exercise 3. *Nalezněte příklad s $N = 3$, ve kterém je externím regret nulový, ale swap regret je neomezený jako funkce T.*

Upřesnění: stačí pouze vybrat příslušnou posloupnost akcí $a^1, \dots, a^T$, $a^i \in X = \{1,2,3\}$ a posloupnost ztrát $\ell_a^1, \dots, \ell_a^T$ pro každé $a \in X$.

Pro hru $G = (P, A, C)$ v normálním tvaru pro n hráčů je pravděpodobnostní rozdělení $p(a)$ na A *korelovaným ekvilibriem* v G, pokud $\sum_{a_{-i} \in A_{-i}} C_i(a_i; a_{-i}) p(a_i; a_{-i}) \leq \sum_{a_{-i} \in A_{-i}} C_i(a_i'; a_{-i}) p(a_i; a_{-i})$ pro každého hráče $i \in P$ a všechna $a_i, a_i' \in A_i$. Pravděpodobnostní rozdělení $p(a)$ na A je *hrubým korelovaným ekvilibriem* v G, pokud $\sum_{a \in A} C_i(a) p(a) \leq \sum_{a \in A} C_i(a_i'; a_{-i}) p(a)$ pro každého hráče $i \in P$ a každé $a_i' \in A_i$.

Exercise 4. *Formálně dokažte, že každé korelované ekvilibrium je hrubým korelovaným ekvilibriem.*

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