

# Algoritmická teorie her – příklady na 8. cvičení\*

12. prosince 2023

## 1 Konvergence no-regret učení

Pro posloupnost  $(p^t)_{t=1}^T$  pravděpodobnostních rozdělání vystoupených algoritmem  $A$  a pro modifikační pravidlo  $F: X \rightarrow X$ , 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  $f_i^t = \sum_{j: F^t(j)=i} p_j^t$ . *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  $\ell^t$  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  $\mathcal{F}^{ex} = \{F_i: i \in X\}$  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  $\mathcal{F}^{in} = \{F_{i,j}: (i,j) \in X \times X, i \neq j\}$  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  $\mathcal{F}^{sw}$  všech modifikačních pravidel  $F: X \rightarrow X$ .

**Exercise 1.** Dokažte, že pravděpodobnostní rozdělání  $p$  je korelovaným ekvilibriem, neboli

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

pro každého hráče  $i \in P$  a všechna  $a_i, a'_i \in A_i$ , právě tehdy, když

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

pro každého hráče  $i \in P$  a každé modifikační pravidlo  $F: A_i \rightarrow A_i$ .

*Hint:* Může se hodit nahlédnout, že platí  $\mathbb{E}_{a \sim p}[C_i(a)] = \sum_{a_i \in A_i} P(a_i) \cdot \mathbb{E}_{a \sim p}[C_i(a) \mid a_i]$ , kde  $P(a_i) = \sum_{a_{-i} \in A_{-i}} p(a_i; a_{-i})$  je pravděpodobnost, že hráči  $i$  je doporučeno  $a_i$ .

**Exercise 2.** Bud'  $G = (P, A, C)$  hra v normálním tvaru pro  $n$  hráčů,  $\varepsilon > 0$  a  $T = T(\varepsilon) \in \mathbb{N}$ . Necht' po  $T$  krocích No-regret dynamics má každý hráč  $i \in P$  průměrný externí regret nanejvýš  $\varepsilon$ . Definujme  $p^t = \prod_{i=1}^n p_i^t$  součin smíšených strategií hráčů v kroce  $t$  a bud'  $p = \frac{1}{T} \sum_{t=1}^T p^t$ . Ukažte, že  $p$  is  $\varepsilon$ -hrubým korelované ekvilibirium, neboli

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

pro každého hráče  $i \in P$  a každou akci  $a'_i \in A_i$ .

**Exercise 3.** Bud'  $G = (P, A, C)$  hra v normálním tvaru pro  $n$  hráčů,  $\varepsilon > 0$  a  $T = T(\varepsilon) \in \mathbb{N}$ . Necht' po  $T$  krocích No-swap-regret dynamics má každý hráč  $i \in P$  průměrný swap regret nanejvýš  $\varepsilon$ . Definujme  $p^t = \prod_{i=1}^n p_i^t$  součin smíšených strategií hráčů v kroce  $t$  a bud'  $p = \frac{1}{T} \sum_{t=1}^T p^t$ . Ukažte, že  $p$  is  $\varepsilon$ -korelované ekvilibirium, neboli

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

pro každého hráče  $i \in P$  a každé modifikační pravidlo  $F: A_i \rightarrow A_i$ .

---

\*Informace o cvičení naleznete na <http://kam.mff.cuni.cz/~balko/>