Algoritmická teorie her – příklady na 8. cvičení*

1 No-regret dynamics

Máme množinu $X = \{1, \ldots, N\}$ s N akcemi a v každém kroce t online algoritmus A vybere pravděpodobnostní rozdělení $p^t = (p_1^t, \ldots, p_N^t)$ na X. Poté, co je rozdělení p^t vybráno v kroce t, nepřátelské prostředí zvolí ztráty $\ell^t = (\ell_1^t, \ldots, \ell_N^t) \in [-1, 1]^N$, 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 kumulovaná ztráta akce i rovna $L_i^T = \sum_{t=1}^T \ell_i^t$ a kumulovaná ztráta algoritmu A je $L_A^T = \sum_{t=1}^T \ell_A^t$. Budeme také používat $L_{min}^T = \min_{i \in X} L_i^T$. *Externí regret* algoritmu A je $R_A^T = \max_{i \in X}\{L_A^T - L_i^T\} = L_A^T - L_{min}^T$.

Algorithm 1.1: Polynomial weights algorithm(X, T, η)

Vstup : Množina akcí $X = \{1, \ldots, N\}$, $T \in \mathbb{N}$ a $\eta \in (0, 1/2]$.
Výstup : Pravděpodobnostní rozdělení p^t pro každý krok $t \in \{1, \ldots, T\}$.
$w_i^1 \leftarrow 1$ pro každé $i \in X$,
$p^1 \leftarrow (1/N, \ldots, 1/N)$,
for $t = 2, \ldots, T$
do $\begin{cases} w_i^t \leftarrow w_i^{t-1}(1 - \eta \ell_i^{t-1}), \\ W^t \leftarrow \sum_{i \in X} w_i^t, \\ p_i^t \leftarrow w_i^t / W^t \text{ pro každé } i \in X. \end{cases}$
Vystup $\{p^t : t \in \{1, \ldots, T\}\}$.

Algorithm 1.2: No-regret dynamics(G, T, ε)

Vstup : Hra $G = (P, A, C)$ v normálním tvaru pro n hráčů, $T \in \mathbb{N}$ a $\varepsilon > 0$.
Výstup : Pravděpodobnostní rozdělení p_i^t na A_i pro každé $i \in P$ a $t \in \{1, \ldots, T\}$.
for $t = 1, \ldots, T$
do $\begin{cases} \text{Každý hráč } i \in P \text{ zvolí nezávisle strategii } p_i^t \text{ za použití} \\ \text{algoritmu s průměrným regretem } \varepsilon\text{, kde akce} \\ \text{odpovídají čistým strategiím.} \\ \text{Každý hráč } i \in P \text{ obdrží ztrátový vektor } \ell_i^t = (\ell_i^t(a_i))_{a_i \in A_i}\text{, kde} \\ \ell_i^t(a_i) \leftarrow \mathbb{E}_{a_{-i}^t \sim p_{-i}^t}[C_i(a_i; a_{-i}^t)] \text{ přes součinové rozdělení} \\ p_{-i}^t = \prod_{j \neq i} p_j^t. \end{cases}$
Vystup $\{p^t : t \in \{1, \ldots, T\}\}$.

Příklad 1. *Předpokládejme, že oba hráči v Bitvě o duši Gothamu používají Polynomial weights algorithm s $\eta = 1/2$ během No-regret dynamics. Simulujte první tři kroky algoritmu No-regret dynamics, neboli spočítejte strategie p_1^3 and p_2^3.*

	Spolupracovat	Detonovat
Spolupracovat	$(0, 0)$	$(0, -1)$
Detonovat	$(-1, 0)$	$(0, 0)$

Tabulka 1: Matice ztrát ve Hře o duši Gothamu.

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

2 Swap regret a interní regret

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

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

Příklad 3. *Dokažte, že pravděpodobnostní rozdělení 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 $F\colon A_i \to A_i$.

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.

Příklad 4. *Nalezněte příklad s $N = 3$, ve kterém je externí 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 $a \in X$.