Algoritmická teorie her – příklady na 7. cvičení*

5. prosince 2023

1 Regret minimizace

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 $R_{A,\mathcal{F}}^T = \max_{F\in\mathcal{F}} \left\{L_A^T - L_{A,F}^T\right\}$. 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 $F_i = (F_i^t)_{t=1}^T$ 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 1. *Ukažte, že swap regret je vždy nanejvýš N-krát tak velký jako interní regret.*

Exercise 2. *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 $\sum_{a_{-i}\in A_{-i}} C_i(a_i; a_{-i})p(a_i; a_{-i}) \le \sum_{a_{-i}\in A_{-i}} C_i(a_i'; a_{-i})p(a_i; a_{-i})$ 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 $\sum_{a\in A} C_i(a)p(a) \le \sum_{a\in A} C_i(a_i'; a_{-i})p(a)$ pro každého hráče $i \in P$ a každé $a_i' \in A_i$.

2 Hry v rošířené formě

Sekvenční forma hry G s neúplnou informací je čtveřice (P, S, u, C) kde P je množinou n hráčů, $S = (S_1, \dots, S_n)$, kde S_i je množinou posloupností hráče i, $u = (u_1, \dots, u_n)$, kde $u_i\colon S \to \mathbb{R}$ je výplatní funkcí hráče i a $C = (C_1, \dots, C_n)$ je množinou lineárních podmínek na realizační pravděpodobnosti hráče i.

Exercise 3. *Zkonstruujte rozšířenou formu Hry na kuře z Tabulky 1 a určete její sekvenční formu a problém komplementarity pro nalezení Nashových ekvilibrií v této hře.*

	Zatoč	Rovně
Zatoč	(0,0)	(-1,1)
Rovně	(1,-1)	(-10,-10)

Tabulka 1: Hra na kuře v normální formě.

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

Exercise 1. *Ukažte, že swap regret je vždy nanejvýš N-krát tak velký jako interní regret.*

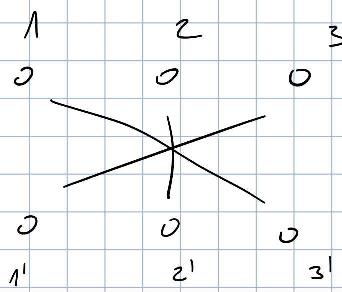
Note done...

Exercise 2. *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 $a \in X$.

– když dám všem akcím stejnou pravděpodobnost, bude externí regret 0.

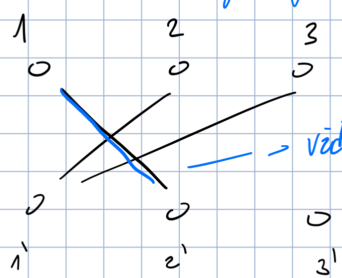
	l_1	l_2	l_3
$t=1$	1	1	0
$t=2$	1	0	1
$t=3$	0	1	1



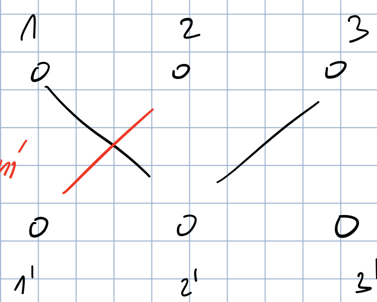
Swap regret za každé 3 kroky $= \frac{2}{3}T$

tohle je externí regret 0, protože jsem všude vygeneroval 2

2 ↳ kdybych přepl na jinou akci (2 nebo 3), měl bych furt loss 2.



→ vždycky kdybych došel na 1 akci 2, tak mě vůbec nezajímají ty jiné přepisy u 2 a 3.



tohle externí regret není

Teď řekl důležitou věc!

A teď další.

Algoritmická teorie her – příklady na 8. cvičení*

12. prosince 2023

1 Konvergence no-regret učení

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 $R_{A,\mathcal{F}}^T = \max_{F \in \mathcal{F}} \left\{ L_A^T - L_{A,F}^T \right\}$. 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 $F_i = (F_i^t)_{t=1}^T$ 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 1. *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.

Exercise 2. *Buď $G = (P, A, C)$ hra v normálním tvaru pro n hráčů, $\varepsilon > 0$ a $T = T(\varepsilon) \in \mathbb{N}$. Nechť po T krocích No-regret dynamics má každý hráč $i \in P$ průměrný externí regret nanejvýš ε. Definujme $p^t = \prod_{i=1}^n p_i^t$ součin smíšených strategií hráčů v kroce t a buď $p = \frac{1}{T}\sum_{t=1}^T p^t$. Ukažte, že p is ε-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. *Buď $G = (P, A, C)$ hra v normálním tvaru pro n hráčů, $\varepsilon > 0$ a $T = T(\varepsilon) \in \mathbb{N}$. Nechť po T krocích No-swap-regret dynamics má každý hráč $i \in P$ průměrný swap regret nanejvýš ε. Definujme $p^t = \prod_{i=1}^n p_i^t$ součin smíšených strategií hráčů v kroce t a buď $p = \frac{1}{T}\sum_{t=1}^T p^t$. Ukažte, že p is ε-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 $F\colon A_i \to A_i$.

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

Exercise 1. *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.

$\Leftarrow$ Kdyby neplatilo, že všechny členy uvnitř sum budou mít stejnou nerovnost

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

$\Rightarrow$ $$\mathbb{E}_{a\sim p}[C_i(a)] = \sum_{a_i\in A_i} p(a_i)\cdot \mathbb{E}_{a\sim p}[C_i(a) \mid a_i] \leq \sum_{a_i\in A_i} p(a_i)\, \mathbb{E}[C_i(a_i', a_{-i}) \mid a_i] =$$

a_i' lze chápat jako $F(a_i)$

$$= \sum_{a_i\in A_i} p(a_i) \mathbb{E}_{a\sim p}[C_i(F(a_i), a_{-i}) \mid a_i] =$$

$$= \mathbb{E}_{a\sim p}[C_i(F(a_i), a_{-i})] \quad \checkmark$$

Exercise 2. *Buď $G = (P, A, C)$ hra v normálním tvaru pro n hráčů, $\varepsilon > 0$ a $T = T(\varepsilon) \in \mathbb{N}$. Nechť po T krocích No-regret dynamics má každý hráč $i \in P$ průměrný externí regret nanejvýš ε. Definujme $p^t = \prod_{i=1}^n p_i^t$ součin smíšených strategií hráčů v kroce t a buď $p = \frac{1}{T}\sum_{t=1}^T p^t$. Ukažte, že p is ε-hrubým korelované ekvilibrium, neboli*

násobím po složkách

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

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

$$p = \frac{1}{T}\sum_{t=1}^{T}\prod_{i=1}^{n} p_i^t$$

každý hráč má po T krocích průměrnou cost $\leq \varepsilon = \dfrac{R^{ext,T}}{T}$

využijeme toho, že costy nezávisí na čase.

$$L_{A_i}^T = \sum_{t=1}^T \ell_{A_i}^t = \sum_{t=1}^T \sum_{a\in A_i} p^t(a)\cdot C_i(a) = \mathbb{E}_{a\sim p^t}[C_i(a)]$$

$= T\cdot p$

$$T\cdot \mathbb{E}_{a\sim p}[C_i(a)] = \sum_{t=1}^T \mathbb{E}_{a\sim p^t}[C_i(a)]$$

Teď uvažme, že budu místo toho hrát jinou akci: $\sum_{t=1}^T \mathbb{E}_{a\sim p^t}[C_i(a_i', a_{-i})]$

tahle je furt stejná

$$\varepsilon = \frac{R^{ext,T}}{\cancel{T}} \geq \frac{-\sum_t \mathbb{E}_{a\sim p^t}[u_i(a_i', a_{-i})] + \sum_t \mathbb{E}_{a\sim p^t}[u_i(a)]}{\cancel{T}}$$

tohle chci minimalizovat

↳ protože $R^{ext,T} = MIN(\ \ \ )$

jelikož $R^{ext,T} \in \mathcal{O}(\sqrt{T})$

takže $\varepsilon = \frac{R^{ext,T}}{T} \in \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$

$$\underset{a_i' \in A_i}{MAX} \left\{ \sum_t \mathbb{E}_{a\sim p^t}[u_i(a_i', a_{-i})] + \cancel{\sum_{t} \mathbb{E}_{a\sim p^t}[u_i(a)]} \right\} \neq R^{ext,T}$$

$$\underset{a_i' \in A_i}{MAX} \left\{ \sum_t \mathbb{E}_{a\sim p^t}[u_i(a_i', a_i)] \right\}$$

2 Hry v rošířené formě

Sekvenční forma hry G s neúplnou informací je čtveřice (P, S, u, C) kde P je množinou n hráčů, $S = (S_1, \ldots, S_n)$, kde S_i je množinou posloupností hráče i, $u = (u_1, \ldots, u_n)$, kde $u_i \colon S \to \mathbb{R}$ je výplatní funkcí hráče i a $C = (C_1, \ldots, C_n)$ je množinou lineárních podmínek na realizační pravděpodobnosti hráče i.

Exercise 3. *Zkonstruujte rozšířenou formu Hry na kuře z Tabulky 1 a určete její sekvenční formu a problém komplementarity pro nalezení Nashových ekvilibrií v této hře.*

	Zatoč	Rovně
Zatoč	(0,0)	(-1,1)
Rovně	(1,-1)	(-10,-10)

Tabulka 1: Hra na kuře v normální formě.

$$X(\sigma_G) = \prod_{c \in \sigma_G} \beta_c$$

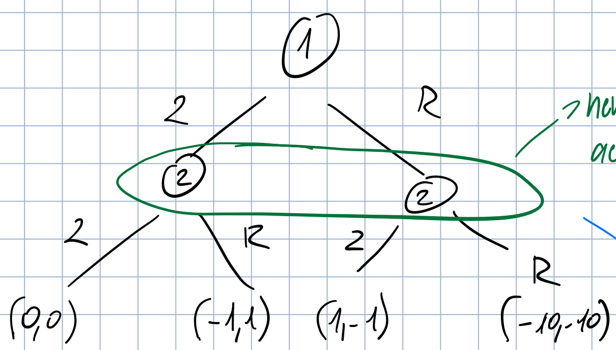
$$S_1 = \{\emptyset, 2, R\} = S_2$$

$$A = \begin{array}{c} \emptyset \\ 2 \\ r \end{array} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & -10 \end{pmatrix} = B^T$$

tohle nevede do terminálu

payoff matice

díky symetrii



nevíš, ve kterém stavu actually seš

já musím mít strategii, která bere v potaz všechny možnosti

$$X(\emptyset) = 1$$

pust, že se dostanu do daného stavu

$$X(\sigma_h) = \sum_{a \in C_h} X(\sigma_h a)$$

pravděpodobnost dané sekvence

všechny akce, které můžu z h provést

přepsané maticově

normalizace

$$\begin{array}{c} \\ \emptyset \\ 2/R \end{array} \begin{array}{c} \begin{array}{ccc} \emptyset & 2 & R \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & -1 \end{pmatrix} \end{array} \cdot \begin{pmatrix} X(\emptyset) \\ X(2) \\ X(R) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

protože nevím kam jsem se dostal takže to reprezentuji info-stav