Algoritmická teorie her – příklady na 6. cvičení*

29. listopadu 2023

1 Regret minimalizace

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

Exercise 2. *Spočítejte všechna hrubá korelovaná ekvilibria ve hře Vězňovo dilema.*

	T	S
T	(2,2)	(0,3)
S	(3,0)	(1,1)

Tabulka 1: Hra z příkladu 2

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 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 3. *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 $\{1, \ldots, T\}$ na vhodně intervaly I_m pro $m = 0, 1, 2, \ldots$ a pusťte A s vhodným parametrem η_m na všech krocích z I_m.

Exercise 4 (*). *Dokažte následující tvrzení o dolních odhadech na externí regret.*

(a) *Jsou-li N a T přirozená čísla taková, že N je mocnina dvojky a $T < \log_2 N$, pak existuje volba náhodných vektorů ztrát z $\{0, 1\}$ takových, že každý online aloritmus A splňuje $\mathbb{E}[L_A^T] \geq T/2$ a $L_{min}^T = 0$.*

(b) *Pro $N = 2$, existuje volba náhodných vektorů ztrát z $\{0, 1\}$ takových, že každý online aloritmus A splňuje $\mathbb{E}[L_A^T - L_{min}^T] \geq \Omega(\sqrt{T})$.*

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

1)

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

dělám jinou akci než mi korelátor říká

$$\sum_{a \in A} \sum_{a_{-i} \in A_{-i}} C_i(a_i, a_{-i}) \cdot p(a_i, a_{-i}) \leq \sum_{a_{-i} \in A_{-i}} C_i(a_i', a_{-i}) \, p(a_i, a_{-i})$$ — korelovaný ekv.

$$= \sum_{a \in A} C_i(a) \cdot p(a) \leq \sum_{a \in A} C_i(a_i', a_{-i}) \, p(a)$$ — hrubé korel. ekv.

<- vyhlásím se na to, co mi říká korelátor, budu vždycky hrát jen jednu akci.

Uvažme pro spor:

$$\sum_{a \in A} C_i(a) \cdot p(a) \boxed{>} \sum_{a \in A} C_i(a_i', a_{-i}) \cdot p(a)$$

Pak v té vnější sumě nahoře bych musel mít alespoň jedno číslo na pravé straně tahy s obrácenou nerovností, čímž bych to ale rozbil. $\boxed{\times}$

2)

tohle je vlastně jedno číslo

$$\sum_{a \in A} \overline{C_i(a) \cdot p(a)} \leq \sum_{a \in A} C_i(a_i', a_{-i}) \, p(a)$$

Exercise 2. *Spočítejte všechna hrubá korelovaná ekvilibria ve hře Vězňovo dilema.*

	T	S
T	(2,2)	(0,3)
S	(3,0)	(1,1)

-> tohle je cost matice!

Tabulka 1: Hra z příkladu 2

za hráče 1:

$$\sum_{i,j \in \{T,S\}} C_1(i,j) \cdot p(i,j) = \underline{2p(t,t) + 3p(s,t) + p(s,s)}$$

$$\sum_{a \in A} C_1(T, a_{-i}) \cdot p(a) = \underline{2p(t,t) + 2p(s,t)}$$

$$\sum_{a \in A} C_1(S, a_{-i}) \cdot p(a) = \underline{3p(t,t) + p(t,s) + 3p(s,t) + p(s,s)}$$

zkombinujme tyto nerovnice:

$$p(s,t) + p(s,s) \leq 0$$
$$0 \leq p(t,t) + p(t,s)$$

Analogicky zde

analogicky díky symetrii costů:

$$p(t,s) + p(s,s) \leq 0$$
$$0 \leq p(t,t) + p(s,t)$$

-> $p(t,s), p(s,s) = 0$

$p(s,t) = 0$

$\Rightarrow p(t,t) = \underline{1}$

3)

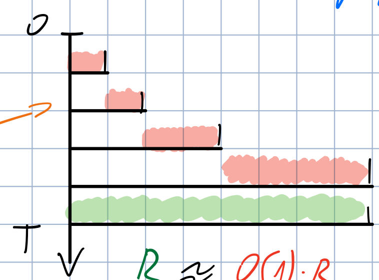
Exercise 3. *Nechť je A algoritmus s parametrem $\eta \in (0, 1/2]$ a s externím regretem nanejvýš $\alpha/\eta + \beta\eta T$ pro nějaké konstanty α, β (které mohou záviset na počtu akcí N). Ukázali jsme, že volba $\eta = \sqrt{\alpha/(T\beta)}$ minimalizuje odhad na regret. Modifikujte tento algoritmus 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 $\{1, \ldots, T\}$ na vhodně intervaly I_m pro $m = 0, 1, 2, \ldots$ a pusťte A s vhodným parametrem η_m na všech krocích z I_m.

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 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 $R_A^T = \max_{i \in X}\{L_A^T - L_i^T\}$.

Poly. w. alg. má nevýhodu, že optimální volba $\mu = \sqrt{\frac{\log(N)}{T}}$

jenže já nevím, kolik kroků je optimální



$R \approx O(1) \cdot R$

tohle chceme ukázat

Maximální ztráta je faktický pro $\mu = \sqrt{\alpha / T\beta}$

$R = \frac{\alpha}{\mu} + \beta \mu T = 2\sqrt{\alpha\beta T} =$

$= C\sqrt{T}$ $\to T = 4^n \Rightarrow$

sečtu všechny regrety

$R' = \sum_{m=0}^{n} C \cdot \sqrt{4^m}$

$= C \cdot \sum_{m=0}^{n} 2^m \le C \cdot 2^{n+1}$

$R' \le C \cdot 2^{n+1}$
$R = C \cdot 2^n$ } Takže to bude nanejvýš 2x větší.

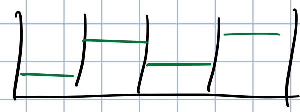
Pak kdyby alg. věděl to T, tak: $R = C\sqrt{T} = C 2^n$

4)

Exercise 4 (*). *Dokažte následující tvrzení o dolních odhadech na externí regret.*

(a) *Jsou-li N a T přirozená čísla taková, že N je mocnina dvojky a $T < \log_2 N$, pak existuje volba náhodných vektorů ztrát z $\{0,1\}$ takových, že každý online aloritmus A splňuje $\mathbb{E}[L_A^T] \geq T/2$ a $L_{min}^T = 0$.*

(b) *Pro $N = 2$, existuje volba náhodných vektorů ztrát z $\{0,1\}$ takových, že každý online aloritmus A splňuje $\mathbb{E}[L_A^T - L_{min}^T] \geq \Omega(\sqrt{T})$.*



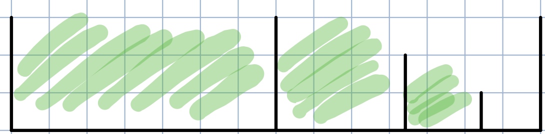
v každým kolu dostane jen alespoň 1/2 loss

a) $T=1$:

Chci $\sum p_i c_i = \frac{1}{2}$

$T=n$:

S každým tím kolem rozpůlím půlku zbylého intervalu a dám mu loss 1.



po $\log_2 N > T$ budu mít alespoň jednu akci, které jsem ještě nedal loss.

b) Budu loss určovat uniformně náhodně, tedy si hodím mincí a jedna akce bude mít loss 1, druhá 0.

$\to$ po T kolech bude ten loss $T/2$.

to kolik padlo jedniček

agenti: A_1, $L^t_{A_1} = \sum_{n=0}^{t} n \cdot \binom{t}{n} \cdot \left(\frac{1}{2}\right)^t$ šance, že jich tolik padne

A_2, $L^t_{A_2} = \sum_{n=0}^{t} (t-n) \cdot \binom{t}{n} \cdot \left(\frac{1}{2}\right)^t$

My chceme:

$$\mathbb{E}[L_A - L_{A_i}] = \mathbb{E}[T/2 - L_{A_i}] \geq \Omega(\sqrt{T})$$

stačí odhadnout takle

$\binom{T}{T/2-k} \cdot \left(\frac{1}{2}\right)^T$ šance že se tak stalo.

↳ o kolik k méně mu mě padl loss.