No-regret dinamika

Biskopics Boglárka

2022. május 20.

Bevezetés

- egy játékos eset
- regret definíció, no-regret algoritmus
- multiplicative weights algoritmus
- több játékos eset, no-regret dinamika

Modell

Legyen A a játékos lehetséges cselekedeteinek halmaza, $|A| = n$, $n \geq 2$
Minden $t = 1, 2, \ldots, T$ lépésben

- a játékos választ egy p^t kevert stratégiát, azaz egy valószínűségi eloszlást a lehetséges cselekedetein
- az ellenfél választ egy $c^t : A \longrightarrow [-1, 1]$ költségvektort (ez a játékos választása után történik)
- a^t cselekedet kiválasztása p^t alapján. A játékos megtudja c^t költségvektort, a költsége $c^t(a^t)$

- cél: összköltség minimalizálása
- probléma: a költségvektor a játékos stratégia-választása után kerül kiválasztásra

- célunk: egy jó algoritmus megkeresése
- próbáljunk a $\sum_{t=1}^{T} \min_{a \in A} c^t(a)$ értékhez, azaz a legolcsóbb cselekedet-sorozathoz viszonyítani

Példa

- legyen $n = 2$, nézzünk egy tetszőleges algoritmust
- az ellenfél c^t-t $(0, 1)$-nek vagy $(1, 0)$-nak választja úgy, hogy 1 legyen a nagyobb valószínűséggel választott cselekedetünk költsége
- a várható költségünk legalább $\frac{T}{2}$, a legolcsóbb cselekedet-sorozat költsége 0

- ez nem eredményes, túl erős
- inkább $\min_{a \in A} \sum_{t=1}^{T} c^t(a)$ értékhez viszonyítunk, azaz a legolcsóbb rögzített cselekedethez

Definíció

Rögzített $c^1, c^2, \ldots, c^T$ költségek mellett az $a^1, a^2, \ldots, a^T$ cselekedetek sorozatának *regret*-je az

$$\frac{1}{T}\left[\sum_{t=1}^{T} c^t(a^t) - \min_{a\in A}\sum_{t=1}^{T} c^t(a)\right]$$

érték.

Definíció

*No-regret algoritmus*ról beszélünk, ha $\forall \varepsilon > 0$-hoz létezik egy elég nagy $T = T(\varepsilon)$, hogy minden ellenfélre a regret legfeljebb ε.

Megfigyelés

Az algoritmusnak randomizáltnak kell lennie.

Ha determinisztikus lenne:

- minden lépésben egy rögzített a^t cselekedetet hajtunk végre
- $c^t(a^t) = 1$, minden más cselekedet költsége 0
- az algoritmus költsége T, a lehető legjobb rögzített cselekedet költsége legfeljebb $\frac{T}{n}$
- így a regret nem tartana 0-hoz

Alsó korlát a regret-re

Ha a lehetséges cselekedetek száma n, a várható regret legalább $b\sqrt{\ln n/T}$, ahol $b > 0$ egy T-től és n-től független konstans.

Multiplicative weights algoritmus

Tétel

Ha n a lehetséges cselekedetek száma és $T \geq 4 \ln n$, akkor létezik online algoritmus, amire a várható regret legfeljebb $2\sqrt{\ln n / T}$ bármilyen ellenfél esetén.

Multiplicative weights algoritmus

1. Kezdetben állítsuk minden cselekedet súlyát 1-re, azaz legyen $w^1(a) = 1\ \forall a \in A$-ra.
2. $t = 1, 2, \dots, T$-re:
   - Legyen $p^t := \frac{w^t}{\Gamma^t}$, ahol $\Gamma^t = \sum_{a \in A} w^t(a)$, cselekedjünk p^t szerint.
   - A c^t költségek megismerése után változtassuk a súlyokat a következőképpen: $w^{t+1}(a) := w^t(a)(1 - \varepsilon c^t(a))\ \forall a \in A$-ra.

- az elemzésből kiderül, hogy $\varepsilon = \sqrt{\ln n / T}$ választással teljesülnek a tétel feltételei

No-regret dinamika

- több játékos; véges, költség-minimalizáló játék

No-regret dinamika

Minden $t = 1, 2, \ldots, T$ lépésben:

1. Minden játékos választ egy kevert stratégiát valamilyen no-regret algoritmus alapján. Az i. játékos stratégiája: p_i^t.
2. Minden játékos kap egy c_i^t költségvektort, ahol $c_i^t(s_i)$ az s_i tiszta stratégia várható költsége a többi játékos választott kevert stratégiáira nézve, azaz:

$$c_i^t(s_i) = \mathbf{E}_{\mathbf{s}_{-i}^t \sim \sigma_{-i}^t}[C_i(s_i, \mathbf{s}_{-i}^t)],$$

ahol $\sigma_{-i}^t = \prod_{j \neq i} p_j^t$.

- a játékosok a c_i^t költségvektorokat használják a no-regret algoritmusukban a súlyok módosítására.

Definíció

Egy σ eloszlás az $S_1 \times \cdots \times S_k$ lehetséges kimenetelek halmazán egy költség-minimalizáló játékban *gyenge korrelált egyensúly*, ha $\forall i \in \{1, 2, \ldots, k\}$ játékosra és $\forall s'_i \in S_i$-re

$$\mathbf{E}_{\mathbf{s}\sim\sigma}[C_i(\mathbf{s})] \leq \mathbf{E}_{\mathbf{s}\sim\sigma}[C_i(s'_i, \mathbf{s}_{-i})].$$

Tétel

Tegyük fel, hogy egy költség-minimalizáló játékban k játékos esetén T lépés után minden játékos várható regret-je legfeljebb ε. Legyen $\sigma^t = \prod_{i=1}^{k} p_i^t$ az eloszlás a lehetséges kimeneteleken a t-edik lépésben és $\sigma = \frac{1}{T}\sum_{t=1}^{T} \sigma^t$. Ekkor σ gyenge korrelált egyensúlyhoz tart olyan értelemben, hogy

$$\mathbf{E}_{\mathbf{s}\sim\sigma}[C_i(\mathbf{s})] \leq \mathbf{E}_{\mathbf{s}\sim\sigma}[C_i(s'_i, \mathbf{s}_{-i})] + \varepsilon$$

minden i játékosra és $s'_i \in S_i$-re.

Köszönöm a figyelmet!