CS 396: Online Markets

Lecture 10: Learning to Bid

Last Time:

- auction theory
- second-price auction
- first-price auction
- complete information analysis (Nash equilibrium)
- incomplete information analysis (Bayes-Nash equilbrium)

Today:

- learning to bid
- full feedback
- partial feedback
- equilibrium of no-regret learning (coarse correlated equilibrium)

Exercise: Optimal Bid

Recall:

- cumulative distribution function: $F_X(\mathsf{z}) = \mathbf{Pr}[X < \mathsf{z}]$
- uniform distribution on $[0, 1]$: $F_X(\mathsf{z}) = \mathsf{z}$
- first-price auction: highest bidder wins, winner pays bid.

Setup:

- you are bidding in a first-price auction
- other bidders with i.i.d. uniform bids on $[0, 1]$

Questions: your value is $\mathsf{v} = 1/2$

- What should you bid against one other bidder?
- What should you bid against two other bidders?

Online Bidding

"bidding in a repeated auction"

Model:

- repeated first-price auction
  - highest bidder wins (random tie-breaking)
  - winner pays bid
- static value $\mathsf{v} \in [1, h]$
- in round i:
  - bid b^i
  - competing bid $\hat{\mathsf{b}}^i \in [1, h]$
  - win if $\mathsf{b}^i > \hat{\mathsf{b}}^i$
    - payoff: $\mathsf{v} - \mathsf{b}^i$
  - lose otherwise
    - payoff: 0
- feedback model:
  - full: learn $\hat{\mathsf{b}}_i^i$
  - partial: learn "win" or "lose"

Discretization

"discretize bid space, run learning algorithm"

Examples: for bidder with value v

- linear discretization: $\mathsf{b}_j = \mathsf{v} - \epsilon\, j$ for $k = h/\epsilon$
- geometric discretization: $\mathsf{b}_j = \mathsf{v} - (1 + \epsilon)^j$ for $k = \log_{1+\epsilon} h \approx \frac{1}{\epsilon} \ln h$

Note: e.g.,

- with $\mathsf{b}_j = \mathsf{v} - (1 + \epsilon)^j$,
- on winning utility is $\mathsf{v} - \mathsf{b}_j = (1 + \epsilon)^j$.

Recall Thm: n rounds, k actions, payoffs in $[0, h]$:
$\mathrm{EW} \geq (1 - \epsilon)\,\mathrm{OPT} - h/\epsilon \ln k$

Cor: EW and linear discretization is:
$\mathrm{EW} \geq (1 - 2\epsilon)\,\mathrm{OPT} - \frac{h}{\epsilon^2} \ln h/\epsilon$

Cor: EW and geometric discritization is:
$\mathrm{EW} \geq (1 - 2\epsilon)\,\mathrm{OPT} - \frac{h}{\epsilon^2} \ln \frac{1}{\epsilon} \ln h$

Can solve for ϵ to optimize per-round regret

Q: can these bounds be improved?

A: yes

Idea: non-uniform bounds on payoffs

- action j is bid b_j, payoff is either 0 or b_j
- upper bound of h is lose.

Full Feedback

"faster learning for full feedback"

Recall: Follow the Perturbed Leader (FTPL)

- learning rate ϵ
- hallucinate: $\mathsf{v}_j^0 = h\times$ "num tails of ϵ-bias coin flipped in a row"
- let $\mathsf{V}_j^i = \sum_{r=1}^{i} \mathsf{v}_j^r$
- in round i choose: $j^i = \operatorname{argmax}_j \mathsf{v}_j^0 + \mathsf{V}_j^{i-1}$

Recall Analysis:

- stability lemma: FTPL and BTPL do the same thing w.p. $(1-\epsilon)$
- small perturbation:
  BTPL $\geq$ OPT $-\mathbf{E}$[MAXPTRB]

Idea: non-uniform hallucination:

- payoffs of action j are in [0,h_j]$
- hallucinate: $\mathsf{v}_j^0 = h_j\times$ "num tails of ϵ-bias coin flipped in a row"

Lemma: hallucination for geometric discritizing
$\mathbf{E}$[MAXPTRB] $\leq h/\epsilon^2$

Proof:

- $\mathbf{E}$[geometric with rate ϵ] $= {}^{1-\epsilon}/_{\epsilon}$
- $\mathbf{E}\left[\max_j \mathsf{v}_j^0\right] \leq \mathbf{E}\left[\sum_j \mathsf{v}_j^0\right] = \sum_j h_j(1-\epsilon)/\epsilon$
- geometric $h_j = (1+\epsilon)^j$
- $\sum_{j=0}^{k}(1+\epsilon)^j \leq h\sum_{j=0}^{\infty}(1+\epsilon)^{-j} \leq h\frac{1+\epsilon}{\epsilon}$
- $\mathbf{E}$[MAXPTRB] $\leq h/\epsilon^2$

Thm: FTPL and geometric discritization is
ALG $\geq (1-2\epsilon)\,\mathrm{OPT} - \frac{h}{\epsilon}^2$

Partial Feedback

"faster learning for partial feedback"

Recall: MAB Reduction to OLA

In round i:

1. $\boldsymbol{\pi} \leftarrow$ OLA
2. draw $j^i \sim \tilde{\boldsymbol{\pi}}$ with

$$\tilde{\pi}_j^i = (1-\epsilon)\,\pi_j^i + {}^{\epsilon}/_{k}$$

3. take action j^i
4. report $\tilde{\mathbf{v}}$ to OLA with

$$\tilde{\mathrm{v}}_j^i = \begin{cases} {}^{\mathsf{v}_j^i}/_{\pi_j^i} & \text{if } j = j^i \\ 0 & \text{otherwise.} \end{cases}$$

Recall Thm:
$\mathbf{E}[\mathrm{MAB}] \geq (1-2\epsilon)\,\mathrm{OPT} - {}^{h\,k}/_{\epsilon^2} \ln k$

Recall Analysis:

Challenge 2: keep $\tilde{h}$ small

Idea 2: pick random action with some minimal probability ${}^{\epsilon}/_{k}$

Lemma 2: if $\pi_j^i \geq {}^{\epsilon}/_{k}$ then $\tilde{\mathrm{v}}_j^i \leq \tilde{h} = {}^{kh}/_{\epsilon}$

Q: improve this with $h_j = (1+\epsilon)^j$?

A: geometric exploration

- explore bid $\mathsf{b}_j = (1+\epsilon)^j$ with prob. $\propto (1+ epsilon)^j$
- $H = \sum_{j=0}^{k}(1+\epsilon)^j \approx h/\epsilon$
- $\tilde{h}_j = (1+\epsilon)^j\, H/\epsilon(1+\epsilon)^j \approx h/\epsilon^2$

Thm: MAB with geometric exploration $\mathbf{E}[\mathrm{MAB}] \geq (1-3\epsilon)\,\mathrm{OPT} - {}^{h}/_{\epsilon^3} \ln h/\epsilon^2$

Equilibrium of No-regret Learning

"outcomes of games under learning"

Recall: mixed Nash: players indepently randomize

Q: would learning in repeated game converge to independent randomization?

A: not generally.

Defn: coarse correlated equilibrium (CCE)

- mediator offers joint distributions of actions (a_R, a_C)
- players either
  - follow mediator
  - pick any fixed outcome
- CCE if best response is to follow mediator

Example: rock-paper-sissors

	Rock	Paper	Sissors
Rock	**-6**,-6	**-1**,1	**1**,-1
Paper	**1**,-1	**-6**,-6	**-1**,1
Sissors	**-1**,1	**1**,-1	**-6**,-6

Q: Nash?

A: uniform mixing

Q: other CCE?

A: uniform mixing over $\{(R,P),(R,S),(P,R),(P,S),(S,R),(S,P)\}$

- payoff from following mediator: 0
- payoff from any fixed action: $0 \times 2/3 - 6 \times 1/3 = -2$

Thm: play is no-regret iff distribution of play is CCE.

- suppose $((a_R^0, a_C^0), \ldots, (a_R^n, a_C^n))$ is no-regret
- no-regret for player R, for all a_R^*: $\sum_i R_{a_R^i, a_C^i} \geq \sum_i R_{a_R^*, a_C^i}$
- consider mediator:
  - pick i uniformly from round $\{1, \ldots, n\}$
  - recommend a_R^i to R and a_C^i to C.
- no-regret $\Leftrightarrow$ CCE.