IfI Summer School 2018 on Machine Learning:	**June 28, 2018**
Machine Learnign and Market Design	Lecturer: Jason D. Hartline

Topics and References

- Auction Theory, Price of Anarchy:
  - Hartline "Mechanism Design and Approximation", Ch 2, 3, 6.
- Online Learning:
  - "Algorithmic Game Theory", Ch 4 (Blum and Mansour)
- Econometrics:
  - Chawla, Hartline, Nekipelov (2017) "Mechanism Redesign"
  - Nekipelov, Syrgkanis, Tardos (2015) "Econometrics for Learning Agents"

Not covered

- Sample complexity.

Overview

"online learning and online markets"

- auction theory
  - first-price auction
  - Bayes-Nash equilibrium
  - price of anarchy
  - econometric inference
  - revenue maximization
- online learning
  - external and internal regret
  - expert learning
  - internal regret learning
  - multi-armed bandit learning
- markets and learning
  - optimal pricing via learning
  - learning and equilibria
  - econometric inference for learning agents
  - price of anarchy for learning agents

Part I: Auction Theory

Equilibrium

"given a game, what is outcome when players behave selfishly?"

Mechanism Design

"design the game so that selfish behavior leads to desired outcome"

Two objectives:

- welfare
- profit

Single-item auctions

"sell a single **item** to one of several **bidders**, each with **private value** for item."

Solution 1: first-price auction (FPA)

- accept sealed bids.
- winner is highest bidder.
- charge winner their bid.

Question 1: what's a good bidder strategy?

Question 2: what is auction outcome?

Solution 2: "English auction"

- raise price from zero.
- bidders drop out until one bidder remaining.
- remaining bidder is winner.
- charge winner the current price.

Question 1: what's a good bidder strategy?

Answer: stop when price > value."

Question 2: what is auction outcome?

Answer: winner has highest value, pays second highest value.

Note: English auction maximizes welfare.

Challenge: takes a long time to run.

Idea: simulate English auction with sealed bids. [Vickrey, '61; Nobel prize]

Solution 3: second-price auction (SPA)

- accept sealed bids.
- simulate English auction:
  - winner is highest bidder.
  - charge winner second highest bid.

Question 1: what's a good bidder strategy?

Answer: bid your value.

Question 2: what is auction outcome?

Answer: winner has highest value, pays second highest value.

Question: how can the seller maximize their profit?

Example 1:

- second-price auction,
- two bidders, and
- values uniformly at random between 0 and 1 (i.e., $U[0, 1]$)

Question: what is second-price's profit?

Review of probability

- Random variable, e.g., $X \sim U[0,1]$
- cumulative distribution function, $F_X(z) = \mathbf{Pr}[X < z]$, e.g., $F_X(z) = z$.
- density function $f_X(z) = \frac{dF_X(z)}{dz}$, e.g., $f_X(z) = 1$.
- expectation, $\mathbf{E}[X] = \int_{-\infty}^{\infty} z f_X(z) dz$, e.g., $\mathbf{E}[X] = \int_0^1 z \cdot 1 dz = 1/2$.

Answer:

- $\mathbf{E}[\text{profit}] = \mathbf{E}[\text{2nd highest bid}] = \mathbf{E}[\text{2nd highest value}]$
- what is $\mathbf{E}[\text{2nd highest value}]$?
- Picture: 0——$\frac{1}{3}$——$\frac{2}{3}$——1

Question: can we get more profit?

Def: second-price auction w. reserve price r

- accept sealed bids.
- add "seller bid" r
- winner is highest bidder.
  (if seller wins, keep item)
- charge winner second highest bid.

Example 2:

- second-price auction with reserve price 1/2,
- two bidders, and
- values uniformly at random between 0 and 1 (i.e., $U[0,1]$)

Question: what is profit of second-price with reserve 1/2?

Answer:

- sort $v_{(1)} > v_{(2)}$
- consider cases:

	Case	profit	probability
A	$v_{(1)} > 1/2 > v_{(2)}$	1/2	1/2
B	$1/2 > v_{(1)} > v_{(2)}$	0	1/4
C	$v_{(1)} > v_{(2)} > 1/2$	$\mathbf{E}\big[v_{(2)} \mid C\big]$	1/4

- Calculate $\mathbf{E}\big[v_{(2)} \mid C\big] = 4/6$
  Picture: 0——-1/2–4/6–5/6–1
- calculate total: $\mathbf{E}\big[\text{SPA}_{1/2}\big] = 1/2 \cdot 1/2 + 0 \cdot 1/4 + 4/6 \cdot 1/4 = 5/12$.

Note: $\mathbf{E}\big[\text{SPA}_{1/2}\big] = 5/12 > \mathbf{E}[\text{SPA}] = 1/3$

Question: what is best reserve price?

Question: what is best auction?

Equilibrium

"given a game, what is outcome when players behave selfishly?"

Incomplete information games (i.e., auctions)

"players have private information that specifies their payoff"

Notation

- vectors $\boldsymbol{v} = (v_1, \ldots, v_n)$
- hiding coordinates:
  $\boldsymbol{v}_{-i} = (v_1, \ldots, v_{i-1}, ?, v_{i+1}, \ldots, v_n)$.
- filling in coordinates:
  $(\boldsymbol{v}_{-i}, z) = (v_1, \ldots, v_{i-1}, z, v_{i+1}, \ldots, v_n)$.

Def: a **strategy** is a function from private info to an action.

Example: strategy for second-price auction: "bid your value"

Note: "bid your value" is is a dominant strategy equilibrium (DSE) for second-price auction.

Bayes-Nash Equilibrium

"how do agents play, when no DSE?"

Recall: first-price auction has no DSE.

Example:

- first price auction
- two bidders, values $U[0, 1]$.

Question: what is equilibrium?

Answer: (guess and verify)

- if player 2 bids $b_2 \sim U[0, 1/2]$, how should player 1 bid?
- what is 1's expected utility with bid b_1?

$$\begin{aligned}
\mathbf{E}[u_1] &= (v_1 - b_1) \times \mathbf{Pr}[1 \text{ wins}] \\
&= (v_1 - b_1)\mathbf{Pr}[b_1 > b_2] \\
&= (v_1 - b_1)\mathbf{Pr}[b_1 > v_2/2] \\
&= (v_1 - b_1)\mathbf{Pr}[2b_1 > v_2] \\
&= (v_1 - b_1)F(2b_1) \\
&= (v_1 - b_1)2b_1 \\
&= 2v_1 b_1 - 2b_1^2
\end{aligned}$$

- to maximize, take derivative and set to zero, solve
- $b_1 = v_1/2$.
- conclusion: equilibrium!

Def: players with a common prior know the distribution of the private info, $\boldsymbol{v} \sim \boldsymbol{F}$.

Def: a strategy profile of $\boldsymbol{s} = (s_1, \ldots, s_n)$ (s_i maps value v_i to bid b_i) is a Bayes-Nash equilibrium (BNE) if for all i $s_i(v_i)$ is a best response when other agents play $\boldsymbol{s}_{-i}(\boldsymbol{v}_{-i})$ with $\boldsymbol{v}_{-i} \sim \boldsymbol{F}_{-i|v_i}$ (conditioned on v_i).

Single-dimensional Games

"the agent's private information is single-dimensional"

Def:

- Value: v_i = value of agent i for "service"
- outcome of game is $\boldsymbol{x}$ and $\boldsymbol{p}$
- game outcome for i:
  - $x_i = \begin{cases} 1 & i \text{ is served} \\ 0 & \text{otherwise} \end{cases}$
  - p_i = payment i makes.
- Utility: $u_i = v_i x_i - p_i$
- agents are risk neutral.

Game rules $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$ map $\boldsymbol{b}$ to outcomes and payments.

- $\tilde{x}_i(\boldsymbol{b})$ = outcome to i when bids are $\boldsymbol{b}$.
- $\tilde{p}_i(\boldsymbol{b})$ = outcome to i when bids are $\boldsymbol{b}$.

Compose game $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$ with $\boldsymbol{s}$ to map $\boldsymbol{v}$ to outcomes and payments:

- $x_i(\boldsymbol{v}) = \tilde{x}_i(\boldsymbol{s}(\boldsymbol{v}))$ = outcome when bidder values are $\boldsymbol{v}$
- $p_i(\boldsymbol{v}) = \tilde{p}_i(\boldsymbol{s}(\boldsymbol{v}))$ = payment for i when values are $\boldsymbol{v}$.

For values $\boldsymbol{v} \sim \boldsymbol{F}$:

- $x_i(v_i) = \mathbf{E}[x_i(\boldsymbol{v}) \mid v_i]$

  $= \mathbf{E}_{\boldsymbol{v}_{-i}}[x_i(\boldsymbol{v}_{-i}, v_i)]$.
- $p_i(v_i) = \mathbf{E}[p_i(\boldsymbol{v}) \mid v_i]$.
- $u_i(v_i) = v_i x_i(v_i) - p_i(v_i)$.

Note: in notation $x_i(v_i)$: $\tilde{x}_i$, $\boldsymbol{s}$, $\boldsymbol{F}$ are implicit.

Characterization of BNE

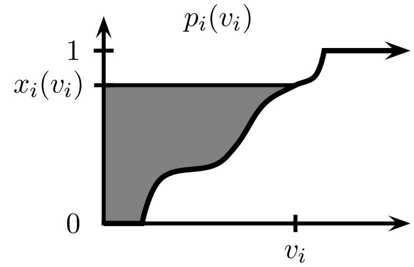
"can we tell if an outcome can be a BNE?"

Def: BNE (for onto strategies):
$\forall i, v_i, z: \quad v_i x_i(v_i) - p_i(v_i) \geq v_i x_i(z) - p_i(z)$.

Theorem: G, onto $\boldsymbol{s}$, and $\boldsymbol{F}$ are in BNE iff

1. $x_i(v_i)$ is monotone non-decreasing,
2. $p_i(v_i) = v_i x_i(v_i) - \int_0^{v_i} x_i(z)dz + p_i(0)$,

and often $p_i(0) = 0$.

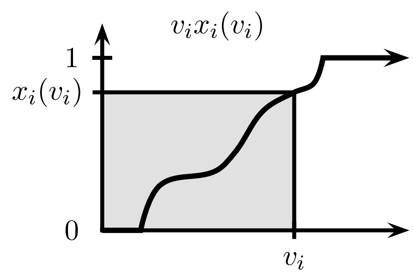


Proof: (BNE $\Longleftarrow$ char) "by picture"

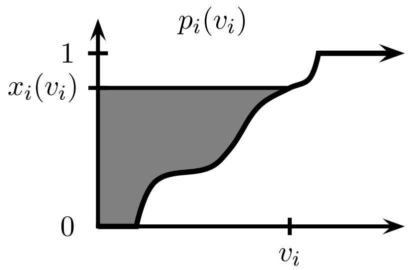
Show that i prefers $s_i(v_i)$ over $s_i(z)$

(Case 1: $z < v_i$; opposite case analogous)

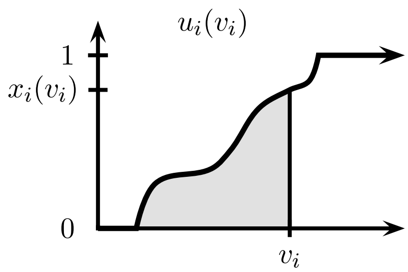
- $u_i(v_i, z)$ = utility witch value v_i playing $s_i(z)$.
- calculate $u_i(v_i, v_i) = v_i x_i(v_i) - p_i(v_i)$
  - plot $v_i x_i(v_i)$



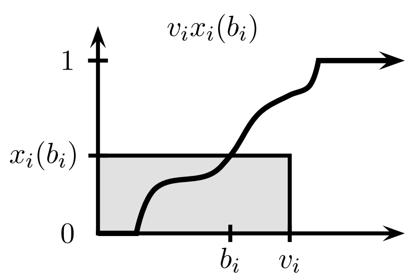
- plot $p_i(v_i)$



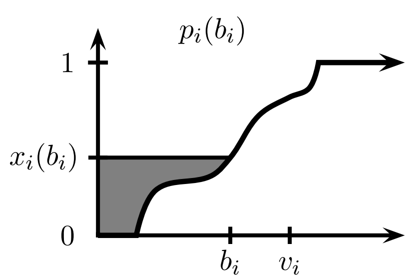
- subtract to get $u_i(v_i)$



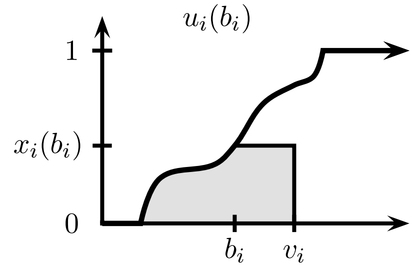
- calculate $u_i(v_i, z) = v_i x_i(z) - p_i(z)$
  - plot $v_i x_i(z)$



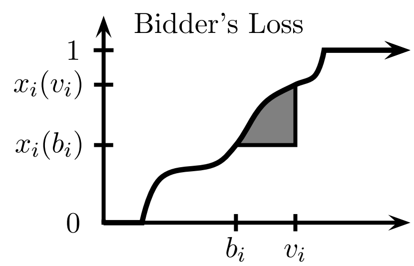
  - plot $p_i(z)$



  - subtract to get $u_i(v_i, z)$



- agent loss is: $u_i(v_i, v_i) - u_i(v_i, z)$



QED

Proof: (BNE $\Rightarrow$ char)

Monotonicity:

- Recall: $u_i(v_i, z) = v_i x_i(z) - p_i(z)$
- BNE $\Rightarrow u_i(v_i, v_i) \geq u_i(v_i, z)$

$$z_2 x_i(z_2) - p_i(z_2) \geq z_2 x_i(z_1) - p_i(z_1)$$
$$z_1 x_i(z_1) - p_i(z_1) \geq z_1 x_i(z_2) - p_i(z_2)$$

  add and cancel

$$z_2 x_i(z_2) + z_1 x_i(z_1) \geq z_2 x_i(z_1) + z_1 x_i(z_2)$$

- Regroup:

$$(z_2 - z_1) x_i(z_2) - (z_2 - z_1) x_i(z_1) \geq 0$$
$$(z_2 - z_1)(x_i(z_2) - x_i(z_1)) \geq 0$$

  then

$$z_2 - z_1 > 0 \Rightarrow x(z_2) \geq x(z_1)$$

$\Rightarrow$ $x_i(\cdot)$ is monotone!

Payment identity (Proof 1):

- solve for $\xi = p_i(z_2) - p_i(z_1)$

  $\Rightarrow$
  $z_2(x_i(z_2) - x_i(z_1)) \geq \xi \geq z_1(x_i(z_2) - x_i(z_1))$

- draw picture.
- draw $p_i(\cdot)$ that satisfies bounds.
- plug in $z_2 = v$ and $z_1 = 0$ for identity.

Payment identity (Proof 2):

- Recall: $u_i(v_i, z) = v_i x_i(z) - p_i(z)$
- BNE implies $u_i(v_i, z)$ maximized $z = v_i$

  $\Rightarrow$ derivative is zero at $z = v_i$.

  - Differentiate with respect to z

$$\frac{d}{dz} u_i(v_i, z) = v_i x_i'(z) - p_i'(z)$$

$$v_i x_i'(v_i) - p_i'(v_i) = 0$$

- holds for all v_i, thus identity:

$$p_i'(z) = z x_i'(z)$$

- integrate both sizes from 0 to v_i

$$\int_0^{v_i} p_i'(z) dz = \int_0^{v_i} z x_i'(z) dz$$

  $p_i(v_i) - p_i(0) = [z x_i(z)]_0^{v_i} - \int_0^{v_i} x_i(z) dz$

- regroup:

  $p_i(v_i) = v_i x_i(v_i) - \int_0^{v_i} x_i(z) dz + p_i(0)$.

QED

Revenue Equivalence

"auctions with same BNE outcome have same profit"

Question: what is outcome of second-price auction?

Answer: bidder with highest value.

Question: who wins in BNE of first-price auction?

Answer: bidder with highest value.

Result: second- and first-price have same expected profit.

Welfare Analysis in Equilibrium

"price of anarchy: bound welfare of complex BNE"

Recall: iid, single-item auction

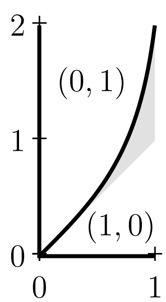
- first-price auction is efficient.
- first-price with monopoly reserve is revenue-optimal.

Question: what about asymmetry: non-iid?

Example: two agents, $v_1 \sim U[0,1]$, $v_2 \sim U[0,2]$

- $s_1(v) = \frac{2}{3v}(2 - \sqrt{4-3v^2})$
- $s_2(v) = \frac{2}{3v}(\sqrt{4-3v^2}) - 2)$

Note: highest-valued agent does not always win.



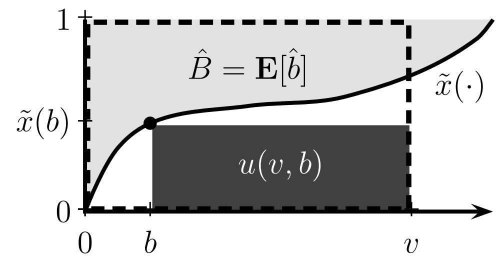
Goal: show that equilibria are still pretty good.

Winner-pays-bid best-response geometry

- bid allocation rule: $\tilde{x}(b)$

  $= \mathbf{Pr}[\text{win with bid } b]$.

  $= \mathbf{Pr}[\text{crit. bid} < b]$.
- utility: $u(v,b) = (v-b)\tilde{x}(b)$
- expected critial bid: $\hat{B} = \int_0^\infty (1-\tilde{x}(b))db$.

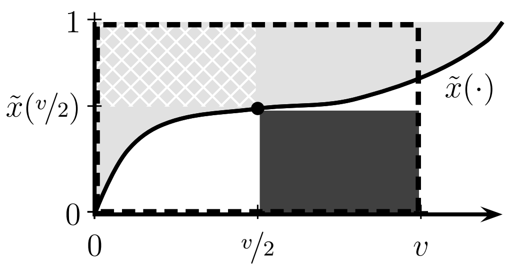


Welfare Analysis

Lemma (Utility Value Covering):
in BNE, $u(v) + \hat{B} \geq \frac{e-1}{e} v$

Proof: (for 1/2)

- BNE $u(v) + \hat{B} \geq u(v, v/2) + \hat{B} \geq v/2$.
- picture:



Def: auction is $\mu \geq 1$ revenue covered if, any bid dists, and feasible alloc $\boldsymbol{y}$,

$$\mu \mathbf{E}[\text{Rev}] \geq \textstyle\sum_i \hat{B}_i y_i$$

Lemma: first-price auction is $\mu = 1$ revenue covered.

Proof: $\sum_i \hat{B}_i y_i$

$\leq \sum_i \mathbf{E}[\text{Rev}]\, y_i$

$\leq \mathbf{E}[\text{Rev}] \sum_i y_i$

$\leq \mathbf{E}[\text{Rev}]$

Theorem: μ revenue-covered auction,
BNE welfare $\geq \frac{e-1}{e\mu}$ optimal welfare.

- fix $\boldsymbol{v}$
- value covering

  $\Rightarrow u_i(v_i) + \hat{B}_i \geq \frac{e-1}{e} v_i$
- for optimal $x_i^\star(\boldsymbol{v})$:

  $\Rightarrow u_i(v_i) + \hat{B}_i x_i^\star(\boldsymbol{v}) \geq \frac{e-1}{e} v_i x_i^\star(\boldsymbol{v})$
- sum over i & revenue covering

  $\Rightarrow \sum_i u_i(v_i) + \mu \mathbf{E}[\text{Rev}] \geq \frac{e-1}{e} \text{OPT}(\boldsymbol{v})$
- expectation over $\boldsymbol{v}$

  $\Rightarrow \underbrace{\mathbf{E}[\text{Util}] + \mu \mathbf{E}[\text{Rev}]}_{\mu\text{Welfare} \geq} \geq \frac{e-1}{e} \text{OPT}$

Def: single-minded combinatorial auction

- m items,
- agent i: value v_i for items S_i.
- $x_i = 1$ if i gets all items in S_i.

Example: single-minded CA not revenue covered.

- agent 0: $S_0 = \{1, \ldots, m\}$, $b_0 = 1$.
- agent i: $S_i = \{i\}$, $b_i = 0$.

$\Rightarrow$ Rev $= 1$, $\hat{B}_i = 1$, $\sum_i \hat{B}_i = m$,

Conclusion: revenue covering

- implies auction good
- no dependence on BNE.
- simple to check

Econometric Inference

"for observed bids in mechanism, infer values"

Observation: in equilibrium and limit with bids from the distribution

- $\tilde{x}$ and $\tilde{p}$ are continuous and observable.
- bidder's first-order condition is satisfied:
$$\frac{d}{db}[v\tilde{x}(b) - \tilde{p}(b)] = 0$$
$\Rightarrow \; v = \tilde{p}'(b)/\tilde{x}'(b)$.

Inference for first-price auction

- $\tilde{p}(b) = b\tilde{x}(b)$
- $\tilde{p}'(b) = b\tilde{x}'(b) + \tilde{x}(b)$

$\Rightarrow \; v = b + \tilde{x}(b)/\tilde{x}'(b)$.

Example: first-price auction, bids $U[0, 1/2]$.

- $\tilde{x}(b) = 2b$; $\tilde{x}'(b) = 2$.
- $v = b + (2b)/(2) = 2b$.

Note: with finite data, need estimator for $\tilde{x}$ and $\tilde{x}'$.

- $\tilde{x}$ is estimated at rate $\sqrt{N}$.
- $\tilde{x}'$ is estimated at rate $N^{1/3}$.
(e.g., by smoothing)

Profit Maximization

"among all auctions, which has highest profit?"

Recall: in BNE

- allocation monotonicity:
$x_i(v_i)$ is non-decreasing.
- payment identity:
$p_i(v_i) = v_i x_i(v_i) - \int_0^{v_i} x_i(z)dz$

Claim: in BNE
$$\mathbf{E}_{v_i}[p_i(v_i)] = \mathbf{E}_{v_i}\left[\left(v_i - \tfrac{1-F(v_i)}{f(v_i)}\right) x_i(v_i)\right]$$

Proof: (sketch)
$$\begin{aligned}\mathbf{E}_{v_i}[p_i(v_i)] &= \int_0^\infty p_i(z)f(z)dz \\ &= \int_0^\infty z x_i(z) f(z) dz - \int_0^\infty \int_0^z x_i(w)dwdz\end{aligned}$$
swap order of integration, simplify
$$\begin{aligned}&= \int_0^\infty \left(z - \tfrac{1-F(z)}{f(z)}\right) x_i(z)f(z)dz \\ &= \mathbf{E}_{v_i}\left[\left(v_i - \tfrac{1-F(v_i)}{f(v_i)}\right) x_i(v_i)\right]\end{aligned}$$
QED

Def: virtual valuation:
$$\phi(v_i) = v_i - \tfrac{1-F(v_i)}{f(v_i)}.$$

Example: for uniform distribution
$$\phi(v_i) = v_i - \tfrac{1-v_i}{1} = 2v_i - 1.$$

Def: virtual surplus for allocation $\boldsymbol{x}$ is
$$\sum\nolimits_i \phi(v_i)x_i$$

Note: $\mathbf{E}$[profit] = $\mathbf{E}$[virtual surplus]

Goal: maximize virtual surplus, subject to monotone allocation

Approach:

- relax monotonicity constraint.
- solve.
- check monotonicity constraint.

Idea: to optimize virtual surplus: choose $\boldsymbol{x} = \text{argmax}_{\boldsymbol{x}'} \sum_i \phi(v_i)x'_i$.

Question: is this monotone? **Answer:** yes, when virtual values are monotone.

Theorem: for monotone v.v.'s, optimal auction allocates to bidder with highest positive virtual valuation.

Note: winner i has

$$\begin{aligned} \phi(v_i) &> \max(0, \phi(v_j)) \\ \phi^{-1}(\phi(v_i)) &> \phi^{-1}(\max(0, \phi(v_j))) \\ v_i &> \max(\phi^{-1}(0), v_j) \end{aligned}$$

Corollary: for monotone virtual values, optimal auction is second-price with reserve price $\phi^{-1}(0)$

Question: optimal auction for two bidders $U[0,1]$?

Answer: second-price with reserve price $\phi^{-1}(0) = 1/2$

(End of Part I)

Part II: Online Learning

Expert Learning

"learn to do as well as best expert in hindsight, payoffs observed"

Model

- k actions (a.k.a., "experts")
- T rounds
- payoff $v_j^t \in [0, h]$ (action j, round t)
- in round t:
  (a) choose an action j^t
  (b) learn payoffs $v_1^t, \ldots, v_k^t$
  (c) obtain payoff $v_{j^t}^t$.

Goal: profit close to best action in hindsight

$\text{OPT} = \max_j \sum_{t=1}^{T} v_j^t$

Algorithm 0: follow the leader (FTL)

- let $V_j^t = \sum_{r=1}^{t} v_j^r$.
- in round t choose

  $j^t = \text{argmax}_j V_j^{t-1}$

Example: (2 actions)

	1	2	3	4	5	6	...
Action 1	1/2	0	1	0	1	0	
Action 2	0	1	0	1	0	1	

- OPT $\approx n/2$
- FTL ≈ 0.

Lemma: all deterministic expert algorithm are $\Omega(n)$-approx.

Conclusion: must randomize!

Learning Algorithms

Algorithm 1: exponential weights (EW$_\epsilon$; a.k.a., Hedge)

- let $V_j^t = \sum_{r=1}^{t} v_j^r$.
- in round t choose j^t with probability proportional to $(1+\epsilon)^{V_j^{t-1}/h}$

Theorem: for expert payoffs in $[0, h]$,

$$\mathbf{E}[\text{EW}] \geq (1-\epsilon)\,\text{OPT} - \tfrac{h}{\epsilon} \ln k$$

Corollary: For T steps and payoffs in $[0, h]$,

$$\text{regret(EW)} \leq 2h\sqrt{\tfrac{\ln k}{T}}$$

Proof of corollary:

- OPT $\leq hT$.
- set $\epsilon h T = \frac{h}{\epsilon} \log k$

$\Rightarrow \epsilon = \sqrt{\frac{1}{T} \ln k}$

$\Rightarrow \mathbf{E}[\text{EW}] \geq \text{OPT} - 2h\sqrt{T \ln n}$

Algorithm 2: follow the perturbed leader (FTPL)

1. hallucinate:

   $V_j^0 = h \times \#$ heads in a row

2. follow the hallucinating leader.

   in round t choose

   $j^t = \text{argmax}_j \{V_j^0 + V_j^{t-1}\}$

Theorem: for expert payoffs in $[0, h]$,

$$\mathbf{E}[\text{FTPL}] \geq \text{OPT}/2 - O(h \ln k)$$

Lemma 1: $\mathbf{E}[\overbrace{\text{BTPL}}^{\text{be the perturbed leader}}] \geq \text{OPT} - O(h \ln k)$

Lemma 2: $\mathbf{E}[\text{FTPL}] \geq \mathbf{E}[\text{BTPL}]/2$.

Proof: (of Lemma 1)

- H_t = perturbed leader's score at t

  $= \max_j (V_j^0 + V_j^t)$
- $h_t = H_t - H_{t-1}$
- BTPL_t = BTPL's payoff from round t.

1. $\text{BTPL} \geq H_T - H_0$
   (a) $\text{BTPL}_t \geq h_t$
      - best expert after t has score H_t
      - best expert before t has score H_{t-1}
      - $BTPL_t$ = best experts payoff from t

        $\geq h_t$

   (b) $\text{BTPL} = \sum_t \text{BTPL}_t \geq \sum_t h_t = H_T - T_0$.
2. $H_T \geq \text{OPT}$

   $H_T = \max_j (V_j^0) + V_j^T) \geq \max_j V_j^T$
3. $\mathbf{E}[H_0] = \Theta(h \log k)$

   H_0 = max of k geometric r.v.s

   (a) flip coins in rounds

      (1 for each expert)

   (b) discard tails expert

      (about half survive)

   (c) how many rounds until none left?

   $\Rightarrow$ $\Theta(\log k)$ rounds

   $\Rightarrow$ $\Theta(h \log k)$ maximum hallucination.

$\Rightarrow$ $\mathbf{E}[\text{BTPL}] \geq \mathbf{E}[H_T - H_0] \geq \text{OPT} - \Theta(h \log k)$.

QED

Proof: (of lemma 2)

Approach:

- $q_j^t = \mathbf{Pr}[\text{FTPL chooses } j \text{ in round } t]$
- $p_j^t = \mathbf{Pr}[\text{BTPL chooses } j \text{ in round } t]$
- show $q_j^t \geq \frac{1}{2} p_j^t$
- apply linearity of expectation:

  BTPL $= \sum_{jt} p_j^t v_j^t$;
  FTPL $= \sum_{jt} q_j^t v_j^t$

Analysis of coupled process

(a) start with raw scores including round t:

$V_1^t, \ldots, V_k^t$

(b) add geometric noise as:

(iv) pick expert j with lowest score

(v) flip coin:

heads: add h to j's score.

tails: discard expert j.

(vi) repeat until one expert j^* left

(c) flip j^*'s coin.

heads (H):

best score $\geq h$ + 2nd best score.

$\Rightarrow$ BTPL and FTPL pick j^*

$\Rightarrow \mathbf{Pr}[\text{FTPL picks } j \text{ at } t \mid j^* \wedge H]$

$= \mathbf{Pr}[\text{BTPL picks } j \text{ at } t \mid j^* \wedge H]$

$= 1$

tails (T):

- $\mathbf{Pr}[\text{FTPL picks } j^* \text{ at } t \mid j^* \wedge T] \geq 0$
- $\mathbf{Pr}[\text{BTPL picks } j \text{ at } t \mid j^* \wedge T] = 1$

$\Rightarrow q_j^t \geq p_j^t/2$

$\Rightarrow$ FTPL $\geq$ BTPL $/2$.

QED

From External to Internal Regret

"reduce internal regret to external regret"

Def:

- alg chooses $\boldsymbol{q}^1, \ldots, \boldsymbol{q}^T$
  $q_j^t = \mathbf{Pr}[\text{alg picks } j \text{ in round } t]$
- external regret:

  $\sum_t \boldsymbol{q}^t \cdot \boldsymbol{v}^t \geq \sum_t v_j^t - R$
- internal regret: for deviation $f : [k] \to [k]$

  $\sum_t \boldsymbol{q}^t \cdot \boldsymbol{v}^t \geq \sum_t \sum_j q_j^t v_{f(j)}^t - R$

Idea:

- use external regret alg for each action.
- mix over external regret algs so as: $\mathbf{Pr}[\text{pick alg } j] = \mathbf{Pr}[\text{pick expert j}]$

```
        q1      ---
[A1]   <==>    |   |  p
 .     p1 v    |   |---->
 .             | H |
 .             |   |  v
 .      qk     |   |<----
[Ak]   <==>    |   |
       pk v     ---
```

Linear Algebra Review

Fact 1: any $n \times n$ square matrix Q has eigenvectors $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$

$$Q\boldsymbol{q}_j = \lambda_j \boldsymbol{q}_j.$$

Def: a stochastic matrix Q has rows (or columns) summing to 1.

Fact 2: a stochastic matrix Q has principle eigenvector $\boldsymbol{q}$ with eigenvalue 1, i.e,

$$Q\boldsymbol{q} = \boldsymbol{q}.$$

Def: principle eigenvector of stochastic transition matrix is fixed point.

Algorithm: External to Internal Regret Reduction

1. instantiate k external regret algorithms $(A_1, \ldots, A_k)$
2. in round t, algs recommend $Q^t = [\boldsymbol{q}_1^t, \ldots, \boldsymbol{q}_2^t]$ (transposed)
3. let $\boldsymbol{p}^t$ be fixed point for Q^t

   i.e., $Q^t\boldsymbol{p}^t = \boldsymbol{p}^t$.
4. choose expert with prob. from $\boldsymbol{p}$.

   (same as choosing algorithm j with prob p_j and then choosing expert with prob. from $\boldsymbol{q}_j$)
5. each alg j's payoff is $p_j^t \boldsymbol{v}^t$.

Theorem: If algs have external regret at most R, then reduction has internal regret at most kR, i.e., for all $f : [k] \to [k]$,

$$\sum_t \boldsymbol{p}^t \cdot \boldsymbol{v}^t \geq \sum_t \sum_j p_j^t \cdot v_{f(j)}^t - kR.$$

Proof:

- for any j, j', A_j satisfies:

  $$\sum_t p_j^t (\boldsymbol{q}_j^t \cdot \boldsymbol{v}^t) \geq \sum_t p_j^t v_{j'}^t - R$$

  (because A_j has external regret $\leq R$)
- consider sum over j of LHS:

  $\sum_t \sum_j p_j^t (\boldsymbol{q}_j^t \cdot \boldsymbol{v}^t)$

  $= \sum_t (\boldsymbol{p}^t \cdot Q^t) \cdot \boldsymbol{v}^t$

  $= \sum_t \boldsymbol{p}^t \cdot \boldsymbol{v}^t$
- sum both sides over j letting $j' = f(j)$:

  $\sum_t \boldsymbol{p}^t \cdot \boldsymbol{v}^t \geq \sum_t \sum_j p_j^t \cdot v_{f(j)}^t - kR.$

QED

Corollary: exists an algorithm H with average internal regret for all $f : [k] \to [k]$,

$$\text{regret}(H, f) \leq 2kh\sqrt{\frac{\ln k}{T}}.$$

From Full to Partial Information

"partial information: only learn payoff of chosen expert"

a.k.a. the multi-armed bandit problem

Model

In round t

- choose expert j^t
- learn payoff $v^t_{j^t}$
- **goal:** approximate $\mathrm{OPT} = \max_j \sum_t v^t_j$

Challenge: must tradeoff explore versus exploit.

Approach: reduce full information to partial information

Idea 1: instead of actual payoff, give alg unbiased estimator of payoff.

- if alg suggests $\boldsymbol{q}^t$
- and samples $j^t \sim \boldsymbol{q}^t$
- (real payoffs are $\boldsymbol{v}$, learn v_{j^t})
- report payoff $\tilde{\boldsymbol{v}}^t = (0, \ldots, v_{j^t}/q_{j^t}, \ldots, 0)$

Note:

- reported payoffs in $[0, \tilde{h}]$ with $\tilde{h} = \max_{j,t} v^t_j / q^t_j$.
- if q^t_j is small, then v^t_j/q^t_j can be very big.

Idea 2: pick a random bandit with some minimal probability ϵ/k.

$\Rightarrow \tilde{h} \leq hk/\epsilon$.

```
  ---            ---  (1-eps)q
|    |    q     |    |  +eps/k
|    | ------> |    | -------->
| A |           |MAB|
|    | (0,vj/  |    |
|    |   ~qj,0)|    |     vj
|    | <------ |    | <-------
  ---            ---
```

Algorithm: Partial to Full Info Algorithm

1. full info alg recommends $\boldsymbol{q}^t$
2. use $\tilde{q}^t_j = (1-\epsilon)q^t_j + \epsilon/k$
   (bandit $j^t \sim \tilde{\boldsymbol{q}}^t$ chosen)
3. report to full info alg $\tilde{\boldsymbol{v}}$ as

$$\tilde{v}^t_j = \begin{cases} v^t_j/\tilde{q}^t_j & \text{if } j = j^t \\ 0 & \text{otherwise.} \end{cases}$$

Lemma: For expert learning (EL) alg with regret $R(h)$ for payoffs in $[0,h]$, the multi-armed bandit (MAB) alg satisfies

$$\mathbf{E}[\mathrm{MAB}] \geq (1-\epsilon)\,\mathrm{OPT} - R(hk/\epsilon)$$

Theorem: for payoff in $[0,h]$ and MAB-EW satisfies

$$\mathbf{E}[\text{MAB-EW}] \geq (1-2\epsilon)\,\mathrm{OPT} - \tfrac{kh}{\epsilon^2}\ln k$$

Proof: combine lemma + EW regret bound

Corollary: for payoff in $[0,h]$ and MAB-EW total regret satisfies

$$\text{regret(MAB-EW)} \geq 3h[(kT^2 \log k)]^{1/3}.$$

Proof: similar to before.

Proof of Lemma

1. what does EL guarantee?

   for any $\tilde{\boldsymbol{v}}^1, \ldots, \tilde{\boldsymbol{v}}^T$, and $j^\star = \operatorname{argmax}_j \sum_t v_j^t$

   $$\mathrm{EL} = \sum_t \boldsymbol{q}^t \cdot \tilde{\boldsymbol{v}}^t \geq \sum_t \tilde{v}_{j^\star}^t - R.$$

   $$\begin{array}{c} \mathbf{E}[\mathrm{EL}] = \sum_t \mathbf{E}\big[\boldsymbol{q}^t \cdot \tilde{\boldsymbol{v}}^t\big] \geq \sum_t \mathbf{E}\big[\tilde{v}_{j^\star}^t\big] - R. \\ = \qquad\qquad\quad = \\ \sum_t \mathbf{E}\big[\boldsymbol{q}^t \cdot \boldsymbol{v}^t\big] \geq \sum_t v_{j^\star}^t - R = \mathrm{OPT} - R \end{array}$$

   For left-hand side equality:

   $$\begin{aligned} \mathbf{E}\big[\boldsymbol{q}^t \cdot \tilde{\boldsymbol{v}}^t\big] &= \int_{\boldsymbol{q}^t} \mathbf{E}\big[\boldsymbol{q}^t \cdot \tilde{\boldsymbol{v}}^t \mid \boldsymbol{q}^t\big] \, \mathbf{Pr}\big[\boldsymbol{q}^t\big] \\ &= \int_{\boldsymbol{q}^t} \mathbf{E}\big[\boldsymbol{q}^t \cdot \boldsymbol{v}^t \mid \boldsymbol{q}^t\big] \, \mathbf{Pr}\big[\boldsymbol{q}^t\big] \\ &= \mathbf{E}\big[\boldsymbol{q}^t \cdot \boldsymbol{v}^t\big] . \end{aligned}$$

2. what is MAB's performance?

   $$\begin{aligned} \mathrm{MAB} &= \sum_t \tilde{\boldsymbol{q}}^t \cdot \boldsymbol{v}^t \\ &= \sum_t (1-\epsilon)\boldsymbol{q}^t \cdot \boldsymbol{v}^t + \tfrac{\epsilon}{k} \sum_j v_j^t \\ &\geq (1-\epsilon) \sum_t \boldsymbol{q}^t \cdot \boldsymbol{v}^t \end{aligned}$$

3. combine:

   $$\mathbf{E}[\mathrm{MAB}] \geq (1-\epsilon)\,\mathrm{OPT} - R$$

Comment: reduction from partial-information and internal-regret to full-information external regret made possible by <u>worst-case</u> expert learning algorithm!

(End of Part II)

Part III: Markets and Learning

Online Pricing

"optimize a posted price to online buyers"

Model

- n agents arrive in sequence.
- in round i:
  - post price $\hat{v}_i$ to agent i
  - agent i has value $v_i \in [1, h]$, buys if $v_i \geq \hat{v}_i$

Goal: optimize revenue: $\sum_{i:v_i \geq \hat{v}_i} \hat{v}_i$

Reduction to Multi-armed Bandit

Recall: k bandits, T rounds, payoffs in $[0, h]$: $\text{MAB} \geq (1-2\epsilon)\,\text{OPT} - \frac{kh}{\epsilon^2} \ln k$.

- discretize prices: $\{(1+\epsilon)^k : k \in \{0, \ldots, \lceil \log_{1+\epsilon} h \rceil\}\}$
- plug in bound: $\text{MAB} \geq (1-3\epsilon)\,\text{OPT} + \frac{h}{\epsilon^3} \ln h \ln \ln h$

Note: can improve using non-uniform MAB analysis

- payoff from posting price $\hat{v}$ is $\hat{v}$ or 0.
- explore price $\hat{v}$ with probability proportional to $\hat{v}$

Note: can learn optimal reserve in multi-agent auction similarly.

Learning Equilibria

"what happens when agents play learning algorithms"

Example: first-price auction

In round t:

- if you (agent 1) bid b_1
- win if $b_1 \geq \hat{b}_1 = \max_{i \neq 1} b_i$
- payoff if win: $v_1 - b_1$.

Reduction to online pricing

- optimize shaded amount: $v_1 - b_1$ (cf. price to post).
- win if shaded amount $\leq v_1 - \hat{b}_1$. (cf. buyer's value)
- payoff is shaded amount if win. (cf. price if buyer buys)

Correlated equilibrium

"story: a mediator chooses strategies, makes suggestion, do agents want to follow suggestion?"

Def: Actions $\boldsymbol{b} \sim \boldsymbol{G}$ (correlated) is CE if:
$\mathbf{E}_{\boldsymbol{b}\sim\boldsymbol{G}}[u_i(\boldsymbol{b}) \mid b_i] \geq \mathbf{E}_{\boldsymbol{b}\sim\boldsymbol{G}}[u_i(b_i^*, \boldsymbol{b}_{-i}) \mid b_i]$

Note: CE are convex,
i.e., $\boldsymbol{G} = \alpha \boldsymbol{G}' + (1-\alpha)\boldsymbol{G}''$

Coarse correlated equilibrium

"story: if players don't get to see recommendation, only get to choose to accept or play another action"

Def: Actions $\boldsymbol{b} \sim \boldsymbol{G}$ (correlated) is CCE if:
$$\mathbf{E}_{\boldsymbol{b}\sim\boldsymbol{G}}[u_i(\boldsymbol{b})] \geq \mathbf{E}_{\boldsymbol{b}\sim\boldsymbol{G}}[u_i(b_i^*, \boldsymbol{b}_{-i})]$$

Fact: set of CE $\subseteq$ set of CCE.

Theorem: No regret dynamics converges to coarse correlated equilibrium.

$$\text{“sequence } (\boldsymbol{b}^1, \ldots, \boldsymbol{b}^T) \text{ is no regret”} \iff \text{“}\boldsymbol{b} \sim \boldsymbol{G} = U\{\boldsymbol{b}^1, \ldots, \boldsymbol{b}^T\} \text{ is CCE”}$$

Theorem: No internal regret dynamics converges to correlated equilibrium.

$$\text{“sequence } (\boldsymbol{b}^1, \ldots, \boldsymbol{b}^T) \text{ is no internal regret”} \iff \text{“}\boldsymbol{b} \sim \boldsymbol{G} = U\{\boldsymbol{b}^1, \ldots, \boldsymbol{b}^T\} \text{ is CE”}$$

Inference for Learning Agents

“infer fundamentals from bids of learning agents”

Recall: from BNE bid distribution can infer value of bidder from bid.

Goal: generalize to learning agents

Given bid profiles $\mathcal{B} = (\boldsymbol{b}^1, \ldots, \boldsymbol{b}^T)$, infer rationalizable sets $\boldsymbol{R}$ as

$$R_i = \{(\epsilon_i, v_i) : \mathcal{B} \text{ is } \epsilon_i \text{ regret for } i \text{ with value } v_i\}$$

Construction:

- ϵ_i regret $\Rightarrow$ for all z:

  $\sum_t u_i(\boldsymbol{b}^t) \geq \sum_t u_i(\boldsymbol{b}^t_{-i}, z) - \epsilon_i$
- recall: $u_i(\boldsymbol{b}) = v_i \tilde{x}_i(\boldsymbol{b}) - \tilde{p}_i(\boldsymbol{b})$
- swapping to bid z:
  - $\Delta\tilde{x}_i(z) = \sum_t \left[\tilde{x}_i(\boldsymbol{b}^t_{-i}, z) - \tilde{x}_i(\boldsymbol{b}^t)\right]$
  - $\Delta\tilde{p}_i(z) = \sum_t \left[\tilde{p}_i(\boldsymbol{b}^t_{-i}, z) - \tilde{p}_i(\boldsymbol{b}^t)\right]$
- ϵ_i regret $\Rightarrow$ for all z:

$$v_i \Delta\tilde{x}(z) - \Delta\tilde{p}_i(z) \leq \epsilon_i$$

- each z gives a linear constraint on $(\epsilon_i, v_i) \in R_i$.

Note: The region defined is convex

Welfare Analysis of Learning Agents

"bound price of anarchy for learning agents"

Def: price of anarchy: optimal welfare / welfare under learning.

Recall: $\hat{B} = \mathbf{E}[\hat{v}]$ = expected critical value.

Recall: in BNE, $u(v) + \hat{B} \geq \frac{e-1}{e} v$

Lemma: in CCE, same holds.

Proof: exercise. (same argument)

Recall: auction is $\mu \geq 1$ revenue covered if, any bid dists, and feasible alloc $\boldsymbol{y}$,

$$\mu \mathbf{E}[\text{Rev}] \geq \textstyle\sum_i \hat{B}_i y_i$$

Recall: μ revenue-covered auction,
BNE welfare $\geq \frac{e-1}{e\mu}$ optimal welfare.

Theorem: same holds for CCE.

Proof: exercise. (same argument)

Welfare Analysis from Data

"can improve welfare analysis with data"

Note: μ is observed in bid data.

(End of Part III)