Data-Driven Optimal Auction Theory

Tim Roughgarden
(Columbia University)

Single-Item Auctions

The Setup:

- 1 seller with 1 item
- n bidders, bidder i has private valuation v_i

Single-Item Auctions

The Setup:

- 1 seller with 1 item
- n bidders, bidder i has private valuation v_i

Question: which auction maximizes seller revenue?

Issue: different auctions do better on different valuations.

- e.g., Vickrey (second-price) auction with/without a reserve price

Single-Item Auctions

The Setup:

- 1 seller with 1 item
- n bidders, bidder i has private valuation v_i

Single-Item Auctions

The Setup:

- 1 seller with 1 item
- n bidders, bidder i has private valuation v_i

Distributional assumption: bidders' valuations $v_1,\ldots,v_n$ drawn independently from distributions $F_1,\ldots,F_n$.

- F_i's known to seller, v_i's unknown

Goal: identify auction that maximizes expected revenue.

Optimal Single-Item Auctions

[Myerson 81]: characterized the optimal auction, as a function of the prior distributions $F_1,\ldots,F_n$.

- Step 1: transform bids to virtual bids: $b_i \rightarrow \varphi_i(b_i)$
  - formula depends on distribution: $\varphi_i(b_i) = b_i - [1 - F_i(b_i)] / f_i(b_i)$
- Step 2: winner = highest positive virtual bid (if any)
- Step 3: price = lowest bid that still would have won

Optimal Single-Item Auctions

[Myerson 81]: characterized the optimal auction, as a function of the prior distributions $F_1,\ldots,F_n$.

- Step 1: transform bids to virtual bids: $b_i \to \varphi_i(b_i)$
  - formula depends on distribution: $\varphi_i(b_i) = b_i - [1 - F_i(b_i)] / f_i(b_i)$
- Step 2: winner = highest positive virtual bid (if any)
- Step 3: price = lowest bid that still would have won

I.i.d. case: 2nd-price auction with monopoly reserve price.

Optimal Single-Item Auctions

[Myerson 81]: characterized the optimal auction, as a function of the prior distributions $F_1,\ldots,F_n$.

- Step 1: transform bids to virtual bids: $b_i \rightarrow \varphi_i(b_i)$
  - formula depends on distribution: $\varphi_i(b_i) = b_i - [1 - F_i(b_i)] / f_i(b_i)$
- Step 2: winner = highest positive virtual bid (if any)
- Step 3: price = lowest bid that still would have won

I.i.d. case: 2nd-price auction with monopoly reserve price.

General case: requires full knowledge of $F_1,\ldots,F_n$

Key Question

Issue: where does this prior come from?

Key Question

Issue: where does this prior come from?

Modern answer: from data (e.g., past bids).

- e.g., [Ostrovsky/Schwarz 09] fitted distributions to past bids, applied optimal auction theory (at Yahoo!)

Key Question

Issue: where does this prior come from?

Modern answer: from data (e.g., past bids).

Question: *How much data is necessary and sufficient to apply optimal auction theory?*

- "data" = samples from unknown distributions $F_1,\ldots,F_n$ (e.g., inferred from bids in previous auctions)
- goal = near-optimal revenue [(1-ε)-approximation]
- formalism inspired by "PAC" learning theory [Vapnik/Chervonenkis 71, Valiant 84]

Some Related Work

Asymptotic regime: [Neeman 03], [Segal 03], [Baliga/Vohra 03], [Goldberg/Hartline/Karlin/Saks/Wright 06]

- for every distribution, expected revenue approaches optimal as number of samples tends to infinity

Some Related Work

Asymptotic regime: [Neeman 03], [Segal 03], [Baliga/Vohra 03], [Goldberg/Hartline/Karlin/Saks/Wright 06]

- for every distribution, expected revenue approaches optimal as number of samples tends to infinity

Uniform bounds for finite-sample regime: [Elkind 07], [Dhangwatnotai/Roughgarden/Yan 10]

Some Related Work

Asymptotic regime: [Neeman 03], [Segal 03], [Baliga/Vohra 03], [Goldberg/Hartline/Karlin/Saks/Wright 06]

- for every distribution, expected revenue approaches optimal as number of samples tends to infinity

Uniform bounds for finite-sample regime: [Elkind 07], [Dhangwatnotai/Roughgarden/Yan 10], [Cole/Roughgarden 14], [Chawla/Hartline/Nekipelov 14], [Medina/Mohri 14], [Cesa-Bianchi/Gentile/Mansour 15], [Dughmi/Han/Nisan 15]

Some Related Work

Asymptotic regime: [Neeman 03], [Segal 03], [Baliga/Vohra 03], [Goldberg/Hartline/Karlin/Saks/Wright 06]

- for every distribution, expected revenue approaches optimal as number of samples tends to infinity

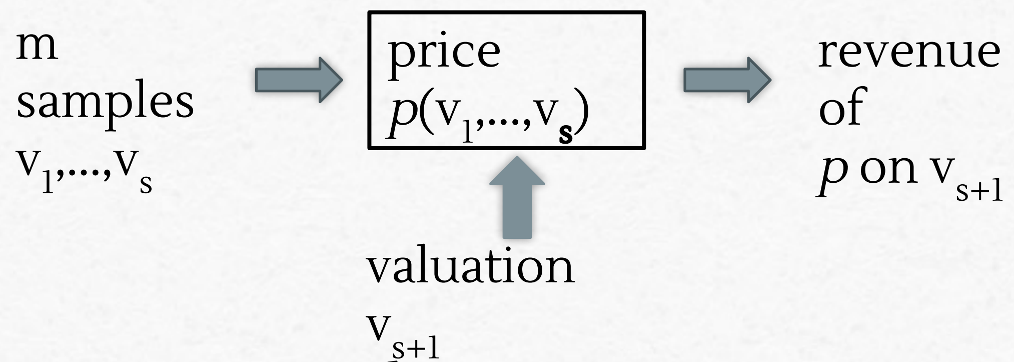
Uniform bounds for finite-sample regime: [Elkind 07], [Dhangwatnotai/Roughgarden/Yan 10], [Cole/Roughgarden 14], [Chawla/Hartline/Nekipelov 14], [Medina/Mohri 14], [Cesa-Bianchi/Gentile/Mansour 15], [Dughmi/Han/Nisan 15], [Huang/Mansour/Roughgarden 15], [Morgenstern/Roughgarden 15,16], [Devanur/Huang/Psomas 16], [Roughgarden/Schrijvers 16], [Hartline/Taggart 17], [Gonczarowski/Nisan 17], [Syrgkanis 17], [Cai/Daskalakis 17], [Balcan/Sandholm/Vitercik 16,18], [Gonczarowski/Weinberg 18], [Hartline/Taggart 19], [Guo/Huang/Zhang 19]

Formalism: Single Buyer

Step 1: seller gets s samples $v_1,\ldots,v_s$ from unknown F

Step 2: seller picks a price $p = p(v_1,\ldots,v_s)$

Step 3: price p applied to a fresh sample v_{s+1} from F



Goal: design *p()* so that $E_{v_1,\ldots,v_s}[p(v_1,\ldots,v_s)\cdot(1-F(p(v_1,\ldots,v_s)))]$ is close to $\max_p[p\cdot(1-F(p))]$ (no matter what F is)

Results for a Single Buyer

1. no assumption on F: no finite number of samples yields non-trivial revenue guarantee (uniformly over F)

Results for a Single Buyer

1. no assumption on *F*: no finite number of samples yields non-trivial revenue guarantee (uniformly over F)
2. if *F* is "regular": with s=1...

Results for a Single Buyer

1. no assumption on *F*: no finite number of samples yields non-trivial revenue guarantee (uniformly over F)
2. if *F* is "regular": with s=1, setting $p(v_1) = v_1$ yields a ½-approximation (consequence of [Bulow/Klemperer 96])

Results for a Single Buyer

1. no assumption on *F*: no finite number of samples yields non-trivial revenue guarantee (uniformly over F)
2. if *F* is "regular": with s=1, setting $p(v_1) = v_1$ yields a ½-approximation (consequence of [Bulow/Klemperer 96])
3. for regular *F*, arbitrary ε: ≈ $(1/\varepsilon)^3$ samples necessary and sufficient for (1-ε)-approximation [Dhangwatnotai/Roughgarden/Yan 10], [Huang/Mansour/Roughgarden 15]

Results for a Single Buyer

1. no assumption on *F*: no finite number of samples yields non-trivial revenue guarantee (uniformly over F)
2. if *F* is “regular”: with s=1, setting $p(v_1) = v_1$ yields a ½-approximation (consequence of [Bulow/Klemperer 96])
3. for regular *F*, arbitrary ε: ≈ $(1/\varepsilon)^3$ samples necessary and sufficient for (1-ε)-approximation [Dhangwatnotai/Roughgarden/Yan 10], [Huang/Mansour/Roughgarden 15]
4. for *F* with a monotone hazard rate, arbitrary ε: ≈ $(1/\varepsilon)^{3/2}$ samples necessary and sufficient for (1-ε)-approximation [Huang/Mansour/Roughgarden 15]

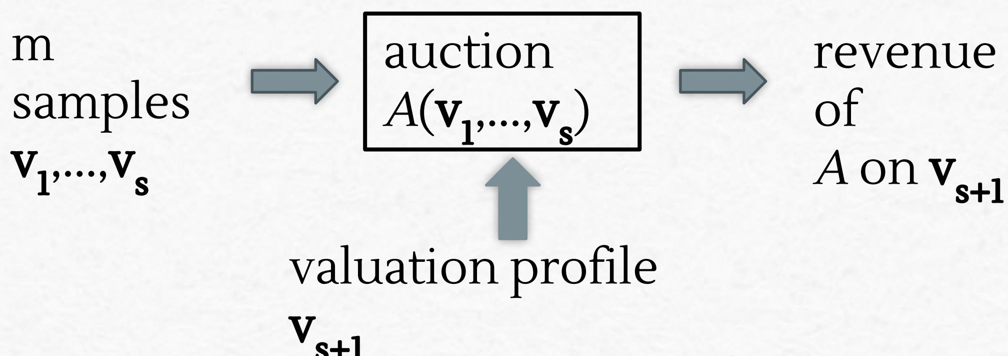
Formalism: Multiple Buyers

Step 1: seller gets s samples $\mathbf{v_1},\ldots,\mathbf{v_s}$ from $\boldsymbol{F} = F_1 \times \cdots \times F_n$

- each $\mathbf{v_i}$ an n-vector (one valuation per bidder)

Step 2: seller picks single-item auction $A = A(\mathbf{v_1},\ldots,\mathbf{v_s})$

Step 3: auction A is run on a fresh sample $\mathbf{v_{s+1}}$ from $\boldsymbol{F}$



Goal: design A so $E_{v_1,\ldots,v_s}[E_{v_{s+1}}[\text{Rev}(A(v_1,\ldots,v_s)(v_{s+1}))]]$ close to OPT

Results: Single-Item Auctions

Theorem: [Cole/Roughgarden 14] The sample complexity of learning a (1-ε)-approximation on an optimal single-item auction is polynomial in n and ε^{-1}.

- n bidders, independent but non-identical regular valuation distributions

Results: Single-Item Auctions

Theorem: [Cole/Roughgarden 14] The sample complexity of learning a (1-ε)-approximation on an optimal single-item auction is polynomial in n and ε^{-1}.

- n bidders, independent but non-identical regular valuation distributions

Optimal bound: [Guo/Huang/Zhang 19] $O(n/\varepsilon^{-3})$ samples.

- $O(n/\varepsilon^{-2})$ for MHR distributions
- tight up to logarithmic factors

A General Approach

Goal: [Morgenstern/Roughgarden 15,16] seek meta-theorem: for "simple" classes of mechanisms, can learn a near-optimal mechanism from few samples.

But what makes a mechanism "simple" or "complex"?

What Is...Simple?

Simple vs. Optimal Theorem [Hartline/Roughgarden 09] (extending [Chawla/Hartline/Kleinberg 07]): in single-parameter settings, independent but not identical private valuations:

expected revenue of VCG with monopoly reserves $\geq$ ½ •(OPT expected revenue)

Pseudodimension: Examples

Proposed simplicity measure of a class C of mechanisms: *pseudodimension* of the real valued functions (from valuation profiles to revenue) induced by C.

Pseudodimension: Examples

Proposed simplicity measure of a class C of mechanisms: *pseudodimension* of the real valued functions (from valuation profiles to revenue) induced by C.

Examples:

- Vickrey auction, anonymous reserve O(1)
- Vickrey auction, bidder-specific reserves O(n log n)
- 1 buyer, selling k items separately O(k log k)
- virtual welfare maximizers unbounded

Pseudodimension: Implications

Theorem: [Haussler 92], [Anthony/Bartlett 99] if C has low pseudodimension, then it is easy to learn from data the best mechanism in C.

Pseudodimension: Implications

Theorem: [Haussler 92], [Anthony/Bartlett 99] if C has low pseudodimension, then it is easy to learn from data the best mechanism in C.

- obtain $s = \tilde{\Omega}(\varepsilon^{-2} d)$ samples $\mathbf{v_1},\ldots,\mathbf{v_s}$ from F, where d = pseudodimension of C, valuations in [0,1]
- let M^* = mechanism of C with maximum total revenue on the samples

Guarantee: with high probability, expected revenue of M^* (w.r.t. F) within ε of optimal mechanism in C.

Consequences

Meta-theorem: simple vs. optimal results automatically extend from known distributions to unknown distributions with a polynomial number of samples.

Examples:

- Vickrey auction, anonymous reserve O(1)
- Vickrey auction, bidder-specific reserves O(n log n)
- grand bundling/selling items separately O(k log k)

$$s = \tilde{\Omega}(\varepsilon^{-2} d)$$

Guarantee: with , with high probability, expected revenue of M^* (w.r.t. F) within ε of optimal mechanism in C

Simplicity-Optimality Trade-Offs

Simple vs. Optimal Theorem: in single-parameter settings, independent but not identical private valuations:

expected revenue of VCG with monopoly reserves $\geq$ ½ •(OPT expected revenue)

Simplicity-Optimality Trade-Offs

Simple vs. Optimal Theorem: in single-parameter settings, independent but not identical private valuations:

expected revenue of VCG with monopoly reserves $\geq$ ½ •(OPT expected revenue)

t-Level Auctions: can use t reserves per bidder.

- winner = bidder clearing max # of reserves, tiebreak by value

Simplicity-Optimality Trade-Offs

Simple vs. Optimal Theorem: in single-parameter settings, independent but not identical private valuations:

expected revenue of VCG with monopoly reserves $\geq$ ½ •(OPT expected revenue)

t-Level Auctions: can use t reserves per bidder.

- winner = bidder clearing max # of reserves, tiebreak by value

Theorem: (i) pseudodimension = $O(nt \log nt)$;

(ii) to get a $(1-\varepsilon)$-approximation, enough to take

$t \approx 1/\varepsilon$

Summary

- key idea: weaken knowledge assumption from known valuation distribution to sample access
- learning theory offers useful framework for reasoning about how to use data to learn a near-optimal auction
  - and a formal definition of "simple" auctions --- polynomial sample complexity (or polynomial pseudo-dimension)
- analytically tractable in many cases
- future directions: (i) incentive issues in data collection; (ii) censored data; (iii) computational complexity issues; (iv) online version of problem

FIN

Benefits of Approach

- relatively faithful to current practices
  - data from recent past used to predict near future
- quantify value of data
  - e.g., how much more data needed to improve revenue guarantee from 90% to 95%?
- suggests how to optimally use past data
- optimizing from samples a potential “sweet spot” between worst-case and average-case analysis
  - inherit robustness from former, strong guarantees from latter

Related Work

- menu complexity [Hart/Nisan 13]
  - measures complexity of a single deterministic mechanism
  - maximum number of distinct options (allocations/prices) available to a player (ranging over others' bids)
  - selling items separately = maximum-possible menu complexity (exponential in the number of items)
- mechanism design via machine learning [Balcan/Blum/Hartline/Mansour 08]
  - covering number measures complexity of a family of auctions
  - prior-free setting (benchmarks instead of unknown distributions)
  - near-optimal mechanisms for unlimited-supply settings

Pseudodimension: Definition

[Pollard 84] Let F = set of real-valued functions on X.

(for us, X = valuation profiles, F = mechanisms, range = revenue)

F *shatters* a finite subset S={$\mathbf{v_1}$,...,$\mathbf{v_s}$} of X if:

- there exist real-valued thresholds $t_1,\ldots,t_s$ such that:
- for every subset T of S

$f(\mathbf{v_i}) \geq t_i \Leftrightarrow \mathbf{v_i}$ in T

- there exists a function f in F such that:

Pseudodimension: Example

Let C = second-price single-item auctions with bidder-specific reserves.

Claim: C can only shatter a subset $S=\{\mathbf{v_1},\ldots,\mathbf{v_s}\}$ if $s = O(n \log n)$. (hence pseudodimension $O(n \log n)$)

Pseudodimension: Example

Let C = second-price single-item auctions with bidder-specific reserves.

Claim: C can only shatter a subset S={$\mathbf{v_1,\ldots,v_s}$} if s = O(n log n). (hence pseudodimension O(n log n))

Proof sketch: Fix S.

- Bucket auctions of C according to relative ordering of the n reserve prices with the ns numbers in S. (#buckets ≈ $(ns)^n$)
- Within a bucket, allocation is constant, revenue varies in simple way => at most s^n distinct "labelings" of S.
- Since need 2^s labelings to shatter S, s = O(n log n).