Learning for Revenue Optimization

Andrés Muñoz Medina
Renato Paes Leme
How to succeed in business with basic ML?
Complications

- What if the seller only sees a sample of the population?
- What if the seller doesn’t know every buyer’s valuation?
- Can buyers lie and don’t provide their true valuation?
- What if valuations change as a function of features?
Outline

- Online revenue optimization
- Batch revenue optimization
Various flavors of this problem

- One buyer (pricing) vs multiple buyers (auctions)
- Fixed valuations (realizable), random valuations (stochastic) and worst-case valuations (adversarial)
- Contextual vs non-contextual
- Strategic vs myopic buyers
Definitions

- **Valuation** \((v)\): What a buyer is willing to pay for a good
- **Bid**: How much a buyer claims she is willing to pay
- **Reserve price** \((p)\): Minimum price acceptable to the seller
- **Revenue** \((Rev)\): How much the seller gets from selling
- **Interactions** \((T)\): Number of times buyer and seller interact
Single buyer

- Valuation \( v_t = \text{maximum willingness to pay} \)
- Reserve price \( p_t \)
- Myopic (price taking buyer): buys whenever \( v_t \geq p_t \)
  - i.e. doesn’t reason about consequences of purchasing decision
  - revenue function is \( \text{Rev}(p_t, v_t) = p_t \mathbb{1}_{v_t \geq p_t} \)

- Strategic buyer: reasons about how purchasing decisions affect future prices
Single myopic buyer

- Realizable setting: valuation is fixed but unknown
  \[ v_t = v \in [0, 1] \]

- Stochastic setting: valuations are sampled from an unknown distribution
  \[ v_t \sim D \]

- Adversarial setting: no assumption made on valuations

- Seller’s goal: Minimize regret
Single myopic buyer
Fixed valuation

- \( v_t = v \in [0, 1] \)

- Regret: \( \mathcal{R} =Tv - \sum_{t=1}^{T} Rev(p_t, v_t) \)
Binary Search

- At round $k$, $S_k = [a_k, a_k + \Delta_k]$, $s = 0$ and $\Delta_{k+1} = \Delta_k / 2$
- While price accepted $p_t = a_k + s\Delta_{k+1}$; $s = s + 1$
- Rejection: Start new round $a_{k+1}$ is last accepted price
- Stop $\Delta_k < \frac{1}{T}$, offer $p_t = a_k$ for all $t$
Fast Search

- Kleinberg and Leighton 2007
- At round $k$, $S_k = [a_k, a_k + \Delta_k]$, $s = 0$ and $\Delta_{k+1} = \frac{\Delta_k^2}{k}$
- While price accepted $p_t = a_k + s\Delta_{k+1}$; $s = s + 1$
- Rejection: Start new round $a_{k+1}$ is last accepted price
- Stop $\Delta_k < \frac{1}{T}$, offer $p_t = a_k$ for all $t$
Kleinberg and Leighton search

Analysis:

- In each round there is at most one no-sale
- For each sale, the regret is at most $\Delta_k$
- There are at most $\Delta_k / \Delta_{k+1} = 1 / \Delta_k$ sales
- The total regret per round is $O(1)$, since there are $O(\log \log T)$ rounds before $\Delta_k < 1 / T$ the total regret is $O(\log \log T)$. 
Kleinberg and Leighton search

- Regret $\mathcal{R} \in O(\log \log T)$
- Lower bound $\Omega(\log \log T)$
Multiple valuations
Bandits

Expected revenue curve $\mathcal{R}(p) = \mathbb{E}_v[Rev(p, v)]$

Discretize

Apply Bandits
Random valuation

- Valuation $v_t \sim \mathcal{D}$

- Regret $\mathcal{R} = T \max_p \mathbb{E}_p[Rev(p, v_t)] - \mathbb{E} \left[ \sum_{t=1}^{T} Rev(p_t, v_t) \right]$

- General strategy: discretize prices and treat each prices as a bandit
  - without any assumptions $\tilde{O}(T^{2/3})$: balance the discretization error and error in UCB
  - can be improved for special families of distributions
Random valuation

- Expected revenue function $\mathbb{E}_{v \sim D}[\text{Rev}(p, v)]$ is unimodal
  - Unimodal Lipschitz bandits [Combes, Proutiere 2014] $\tilde{O}(\sqrt{T})$

- If the revenue curve is quadratic around the maximum, then Kleinberg and Leighton also give a $\tilde{O}(\sqrt{T})$ regret algorithm which is tight in this class.
Adversarial Valuations

- Compete against the best fixed price policy

\[ \mathcal{R} = \mathbb{E} \left[ \max_{p^*} \sum_{t=1}^{T} Rev(p^*, v_t) - \sum_{t=1}^{T} Rev(p_t, v_t) \right] \]

- General approach: discretize prices in K intervals and treat each as an arm. Use EXP3: [Kleinberg and Leighton 07]

\[ \mathcal{R} = \tilde{O}(\sqrt{KT}) + O(T/K) = \tilde{O}(T^{2/3}) \]

EXP3 discretion regret regret
Contextual Pricing

- Each product represented by a context $x_t \in \mathbb{R}^d; \|x_t\|_2 \leq 1$

- Buyer valuation is a dot-product: $v_t = \langle \theta, x_t \rangle$

- The weight vector $\theta$ is fixed but unknown, $\|\theta\|_2 \leq 1$

- Regret is: $R = \sum_{t=1}^{T} v_t - Rev(p_t, v_t)$

- Can we draw a connection with online learning?
Contextual Pricing

- Stochastic gradient give regret $\tilde{O}(\sqrt{T})$ [Amin et al. 2014]
- Cohen, Lobel, Paes Leme, Vladu, Schneider: $\mathcal{R} = O(d \log T)$
- Algorithm based on the ellipsoid method

Keep knowledge sets:

$$S_0 = \{ \theta \in \mathbb{R}^d; \|\theta\|_2 \leq 1 \}$$

For each $x_t$ we know: $v_t \in [a_t, b_t]$

$$a_t = \min_{\theta \in S_t} \langle \theta, x_t \rangle$$
$$b_t = \max_{\theta \in S_t} \langle \theta, x_t \rangle$$
Contextual Pricing

- Stochastic gradient give regret $\tilde{O}(\sqrt{T})$ [Amin et al. 2014]
- Cohen, Lobel, Paes Leme, Vladu, Schneider: $R = O(d \log T)$
- Algorithm based on the ellipsoid method

If $|a_t - b_t| \leq 1/T$ then we are done.
If not, guess $p_t \in [a_t, b_t]$

Update the knowledge set to either:

$$S_{t+1} = \{ \theta \in S_t; \langle \theta, x_t \rangle \leq p_t \}$$

$$S_{t+1} = \{ \theta \in S_t; \langle \theta, x_t \rangle \geq p_t \}$$
Contextual Pricing

- Stochastic gradient give regret $\tilde{O}(\sqrt{T})$ [Amin et al. 2014]
- Cohen, Lobel, Paes Leme, Vladu, Schneider: $\mathcal{R} = O(d \log T)$
- Algorithm based on the ellipsoid method
  
  **Theorem**: Setting $p_t = \frac{1}{2}(a_t + b_t)$ has $\Theta(2^d \log T)$ regret.
  
  **Theorem**: Ellipsoid regularization has $O(d^2 \log T)$ regret.
  
  **Theorem**: Cylindrification regularizer has $O(d \log T)$ regret.
  
  **Theorem**: Squaring trick has regret $O(d^4 \log \log T)$.
Strategic Buyers
Strategic buyers

- What happens if buyers know the seller will adapt prices?
Setup

- Buyer’s valuation $v_t$

- Seller offers price $p_t$

- Buyer accepts $a_t = 1$ or rejects $a_t = 0$

- Discount factor $\gamma$

- Buyer optimizes $\mathbb{E} \left[ \sum_{t=1}^{T} \gamma^t a_t (v_t - p_t) \right]$

- Seller maximizes revenue $\mathbb{E} \left[ \sum_{t=1}^{T} a_t p_t \right]$
Three scenarios

- Fixed value $v_t = v$ [Amin et al. 2013, Mohri and Muñoz 2014, Drutsa 2017]
- Random valuation $v_t \sim D$ [Amin et al. 2013, Mohri and Muñoz 2015]
- Contextual valuation $v_t = \langle \theta, x_t \rangle$ with $x_t \sim D$ [Amin et al. 2014]
Game setup

- Seller selects pricing algorithm
- Announces algorithm to buyer
- Buyer can play strategically
Measuring regret

- Best fixed price in hindsight?

- $p_t = 4, 2, 1, 1, 1, 1, 1, \ldots$
- $a_t = 0, 0, 1, 1, 1, 1, 1, \ldots$

real value = 8
fake value = 1

No
No
YES!
Strategic Regret

- Compare against best possible outcome

- Fixed valuation \( \mathcal{R} = T v - \sum_{t=1}^{T} a_t p_t \)

- Random valuation \( \mathcal{R} = T \max_p \mathbb{E}_p[Rev(p, v_t)] - \mathbb{E}[a_t p_t] \)

- Contextual valuation \( \mathcal{R} = \mathbb{E}\left[ \sum_{t=1}^{T} v_t - a_t p_t \right] \)
The Buyer

- Knowledge of future incentivizes buyer to lie

- **Lie:** Buyer rejects even if his value is greater than reserve price
How can we reduce the number of lies?
Warm up

- Monotone algorithms [Amin et al. 2013]
- Choose $\beta < 1$
- Offer prices $p_t = \beta^t$
- If accepted offer price for the remaining rounds
Warm up

- Decrease slowly to make lies costly
- Not too slow or accumulate regret
- Regret in $O\left(\frac{\sqrt{T}}{1-\gamma}\right)$
- Lower bound $\Omega\left(\log \log T + \frac{1}{1-\gamma}\right)$
Better guarantees

- Fast search with penalized rejections [Mohri and Muñoz 2014]
  - Every time a price is rejected offer again for several rounds
  - Regret in $O\left(\frac{\log T}{1 - \gamma}\right)$

- Horizon independent guarantees [Drutsa 2017]
  - Regret in $O\left(\frac{\log \log T}{1 - \gamma}\right)$
Random valuations

- Valuation $v_t \sim D$

- Regret $\mathcal{R} = T \max_p \mathbb{E}_p[Rev(p, v_t)] - \mathbb{E}[a_t p_t]$

- UCB type algorithm with slow decreasing confidence bounds [Mohri and Muñoz 2015]

- Regret in $O\left(\sqrt{T} + \frac{1}{\log 1/\gamma} T^{1/4}\right)$
Contextual Valuation

- Explore exploit algorithm with longer explore time
- Amin et al. 2014
- Regret in $O\left(\frac{T^{2/3}}{\sqrt{\log(1/\gamma)}}\right)$
Related Work

- Revenue optimization in second price auctions [Cesa-Bianchi et al. 2013]
- Modeling buyers as regret minimizers [Nekipelov et al. 2015]
- Selling to no regret buyers [Heidari et al. 2017, Braverman et al. 2017]
- Selling to patient buyers [Feldman et al. 2016]
Open problems

- Contextual valuations without realizability assumptions
- Strategic buyers with adversarial valuations
- Online learning algorithms in general auctions [Roughgarden 2016]
- Multiple strategic buyers
Revenue from Multiple Buyers
(Pricing -> Auctions)
Multiple buyers

$100

$1000

$50
Multi-buyer Setup

- N buyers with valuations $v_i \in [0, 1]$ from distribution $D_i$

- Auction $A$ is an allocation $x_i : [0, 1]^N \rightarrow \{0, 1\}$ and payment $p_i : [0, 1]^N \rightarrow \mathbb{R}$

- Revenue: $Rev(A) = \sum_{i=1}^{N} p_i$

- Goal: Maximize $\mathbb{E}_{v_1, \ldots, v_N}[Rev(A)]$

- Notation: Given valuation vector $(v_1, \ldots, v_N)$
  $$(v, v_{-i}) = (v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_N)$$
Conditions on auction

- Object can only be allocated once $\sum_{i=1}^{N} x_i \leq 1$
- Individual rationality (IR): $u_i = v_i x_i - p_i \geq 0$
- Incentive compatibility (IC):
  $$v_i x_i(v_i, v_{-i}) - p_i(v_i, v_{-i}) \geq v_i x_i(v, v_{-i}) - p_i(v, v_{-i})$$
Why IC?

- Buyers truly reveal how much they are willing to pay.
- Makes auction stable
- Allows learning
Some IC auctions

- Second price auction: allocate to the buyer with highest $v_i$ and charge second highest value

- $x_i = 1 \leftrightarrow v_i = \max_j v_j$

- $p_i = \max_{j \neq i} v_j$ if $x_i = 1$; 0 otherwise
Second price auction

$100

$1000

$50
IC auctions

- Second price with reserve price $r$: allocate to the highest bidder if $v_i \geq r$. Charge $p_i = \max(r, \max_{j \neq i} v_j)$

  - $x_i = 1$ if $v_i \geq \max_j \max v_j, r$

  - $p_i = \max_{j \neq i} \max v_j, r$ if $x_i = 1$
Second Price Auction With Reserve

$r = \$2000$

$r = \$900$
Myerson Auction

\[ \phi_1 \]
\[ \phi_2 \]
\[ \phi_3 \]

$100 \quad $1000

$600 \quad $500 \quad $300

$90$
Some IC auctions

- Myerson's auction: pick a monotone bid deformation $\phi_i(\cdot)$

- $x_i = 1 \iff \phi_i(v_i) = \max_j \phi_j(v_j) \text{ and } \phi_i(v_i) > 0$

- $p_i = \phi_i^{-1}(\max(\max_{j \neq i} \phi_j(v_j), 0))$ if $x_i = 1$, 0 otherwise

- If $\phi_i = \phi \forall i$

- $x_i = 1 \iff v_i = \max_j v_j$

- $p_i = \phi^{-1} \max(\max_{j \neq i} \phi(v_j), 0) = \max(\max_{j \neq i} v_j, \phi^{-1}(0))$
Myerson Auction

- Optimal auction if \( v_i \sim D_i \) independently

- If \( D_i \) is known, functions \( \phi_i \) can be calculated exactly

- What about unknown distributions?

- Can we learn the optimal monotone functions?

- What is the sample complexity?
Sample Complexity of Auctions

- $N$ bidders
- Valuations $v_i \sim D_i$ independent
- Observe $Nm$ samples $v_{i,1} \ldots v_{i,m} \sim D_i$, $i \in \{1, \ldots, N\}$
- Find auction $A$ such that
  \[
  \mathbb{E}[Rev(A)] \geq (1 - \epsilon) \max_A \mathbb{E}[Rev(A)]
  \]
- Can we use empirical revenue optimization?
  \[
  \max_A \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{N} p_i(v_{1j}, \ldots, v_{Nj})
  \]
Lower bounds on sample complexity

- Proof for a single buyer [Huang et al. 2015]
- Problem reduces to finding the optimal price for a distribution
- Need at least $\Omega\left(\frac{1}{\epsilon^2}\right)$ samples to get a $1 - \epsilon$ approximation
Idea of the proof

- Two similar distributions
- $KL(D_1 \| D_2) = \epsilon$
- Need $\frac{1}{\epsilon^2}$ samples to distinguish them w.h.p
Revenue curves

- Approximately optimal revenue sets disjoint
- If algorithm optimizes revenue for both distributions. It must be able to distinguish them

\[ E_{v \sim D_2}[\text{Rev}(r, v)] \]

\[ E_{v \sim D_1}[\text{Rev}(r, v)] \]
Upper bounds on sample complexity

- Auctions are parametrized by increasing functions $\phi_i$
- Pseudo-dimension of increasing functions is infinite!
- Restrict the class and measure approximation error
t-level auctions

$100

$1000

$50
t-level auctions

- Morgenstern and Roughgarden 2016
- Rank candidates using t-step functions
- Pseudo dimension bounded $O(Nt \log Nt)$
- Best t-level auction is a $\frac{1}{t}$ approximation
t-level auctions

**Theorem:** Let \( t = \Omega\left(\frac{1}{\epsilon}\right) \), using a sample of size \( m = \Omega\left(\frac{N}{\epsilon^3}\right) \) the t-level auction \( \hat{A} \) maximizing empirical revenue is a \( 1 - \epsilon \) approximation to the optimal auction
Algorithm

- Cole and Roughgarden 2015, Huang et al. 2017
- In summary, optimize auctions over all increasing functions
- Proof for finite support
- Extension by discretization
- \( O\left(\frac{1}{\epsilon^3}\right) \) samples
Is this enough?
Features in auctions

- In practice valuations are not i.i.d.
- They depend on features (context)
- Dependency is not realizable in general
- Algorithm of Huang et al. can be generalized to 1 feature
Display ads

- Millions of auctions
- Parametrized by publisher information, time of day, ...
- Dependency of valuations on features is not clear
Setup

- Single buyer auction, find optimal reserve price

- Observe sample \((x_1, v_1), \ldots, (x_m, v_m)\) from distribution \(D\) over \(\mathcal{X} \times [0, 1]\)

- Hypotheses \(h: \mathcal{X} \rightarrow \mathbb{R}\)

- Goal: Find \(\max_{h \in H} \mathbb{E}_{(x,v) \sim D} [\text{Rev}(h(x), v)]\)
Revenue function

- Non-concave
- Non-differentiable
- Discontinuous
- Is it possible to learn?
Learning Theory

- Theorem [Mohri and Muñoz 2013] given a sample of size $m$, with high probability the following bound holds uniformly for all $h \in H$:

$$\left| \mathbb{E}[\text{Rev}(h(x), v)] - \frac{1}{m} \sum_{i=1}^{m} \text{Rev}(h(x_i), v_i) \right| \leq O\left(\sqrt{\frac{\text{PDim}(H)}{m}}\right)$$

Space of linear functions?
Can we do empirical maximization?
The revenue function
Revenue function

- Non-concave
- Non-differentiable
- Discontinuous
- Is it possible to optimize?
Surrogates

- Loss similar to 0-1 loss
- Can we optimize a concave surrogate reward?
Calibration

We say a function $R: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is calibrated with respect to $\text{Rev}$ if for any distribution $D$ we have

$$\arg\max_r \mathbb{E}_v[R(r, v)] \subset \arg\max_r \mathbb{E}_v[\text{Rev}(r, v)]$$
Surrogates

- Theorem [Mohri and Muñoz 2013]: Any concave function that is calibrated is constant.
Continuous Surrogates

- Remove discontinuity
- Difference of concave functions
- DC algorithm for linear hypothesis class [Mohri and Muñoz 2013]
Optimization Issues

- Sequential algorithm
- Not scalable
Other class of functions?
Clustering

- Muñoz and Vassilvitskii 2017
- Show attainable revenue is related to variance of the distribution
- Cluster features to have low variance of valuations
- Revenue related to quality of cluster
Related problems

- Dynamic reserves for repeated auctions [Kanoria and Nazerzadeh 2017]
- New complexity measures [Syrgkanis 2017]
- Combinatorial auction sample complexity [Morgenstern and Roughgarden 2016, Balcan et al. 2016]
- Optimal auction design with neural networks [Dütting et al. 2017]
Conclusion

- Revenue optimization is a crucial practical problem
- Machine learning techniques have yielded new theory and algorithms on this field
- We need to better understand the relationship of buyers and sellers
- There are several open problems still out there
Thank you!