

Learning Theory and Algorithms for Second-Price Auctions with Reserve

Joint work with **Andrés Muñoz Medina**

MEHRYAR MOHRI

MOHRI@

COURANT INSTITUTE & GOOGLE RESEARCH

Auctions

- Standard method for buying or selling goods:
 - U.S. government: Treasury bills.
 - Christie's or Sotheby's: art.
 - eBay: everything, e.g., 'honeymoon wife replacement'.
 - search engine companies: advertising rights.



Auctions

■ Interaction between buyers and sellers:

→ game-theoretical analysis.

- mechanism design.
- study of properties.

■ This talk:

→ learning theory analysis.

- repeated auctions.
- leveraging data.

Some Auction Types

- English auctions: interactive format; seller gradually increases the price until a single bidder is left.
- Dutch auctions (flowers in the Netherlands): interactive format; seller gradually decreases the price until some bidder accepts to pay.
- First-price sealed-bid auctions (e.g. NYC apartments): non-interactive; simultaneous bids, highest bidder wins and pays the value of his bid.

Second-Price Auctions

(William Vickrey, 1961)

- aka Vickrey auctions: e.g., eBay.
 - bidders submit bids simultaneously.
 - highest bidder wins and pays the value of the second-highest bid.
 - truthful bidding is a dominated strategy.

Truthfulness

- Bidder i with value v_i , other bids fixed.
 - if $b_i > v_i$: change only if bidder wins and wasn't before and second-highest bid is $b_j \in [v_i, b_i]$; payoff is $v_i - b_j \leq 0$.
 - if $b_i < v_i$: change only if bidder loses and used to win. and second-highest bid $b_j \in [b_i, v_i]$; payoff was $v_i - b_j \geq 0$.

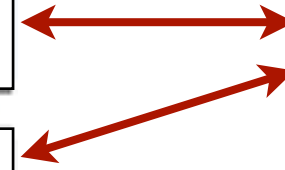
SPA with Reserve

- Second-price auctions with reserve: e.g., Ad Exchanges.
 - seller announces a **reserve price** r and,
 - bidders submit bids simultaneously.
 - winning bidder (if any) wins and pays the maximum of the value of the second-highest bid and r .
 - truthful bidding is a dominated strategy.

Example

- Suppose the seller's value is 0 and there is a single bidder whose value is uniformly distributed over $[0, 1]$.
 - no reserve price: item sold at value 0.
 - reserve price: how should it be chosen?
 - probability $(1 - r)$ for bid being above r .
 - expected revenue $r(1 - r)$, thus $r = \frac{1}{2}$ is optimal.

Ad Exchanges



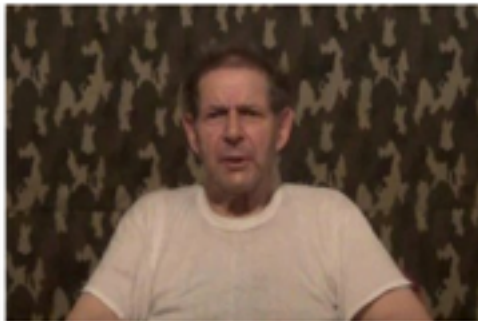
Drone Killed Hostages From U.S. and Italy

Obama Apologizes for Deaths in a U.S. Strike in Pakistan

By PETER BAKER, JULIE HIRSCHFELD DAVIS and ERIC SCHMITT 3:13 PM ET
Intelligence agencies said that Warren Weinstein, held by Al Qaeda since 2011, and Giovanni Lo Porto, held since 2012, died in Pakistan in January. Two American Qaeda members were also killed in attacks in the region.

429 Comments

RELATED COVERAGE
- Video: Obama Statement
- Killing Deepens Debate on Use of Drone Strikes 2:09 PM ET
- 'Kill List' Tests Obama (May 2012)
- A Gracious Host, Immersed in Pakistani Life 4:16 PM



Warren Weinstein was one of the two hostages killed in a counterterrorism operation. NYT Intelligence Group, via Reuters

Experienced Aid Worker, Devoted to Pakistan

By MICHAEL D. SHEAR 3:50 PM ET

Mr. Weinstein, 73, was employed by a contractor for a U.S. aid agency and worked tirelessly to help the Pakistani people.

- A Gracious Host, Immersed in Pakistani Life 4:16 PM

Italian Known as Aid Worker Drawn to Needy

By ELISABETTA POVOLEDO 3:56 PM ET

Giovanni Lo Porto worked for a German aid group that hired him to manage a project in Pakistan.

PRIVATE LIVES

How a Gambling Addict Became a Bookie

By BERT DRAGON

Here was my father, giver of values, such as they were.



- Editorial: Smokeless Tobacco
- Druckerman: Eat Up. You'll Be Happier.
- The Stone: What's Wrong With Inequality?
- Room for Debate: A Less Punitive Child Support System

The Opinion Pages

TAKING NOTE | RINA NORTH

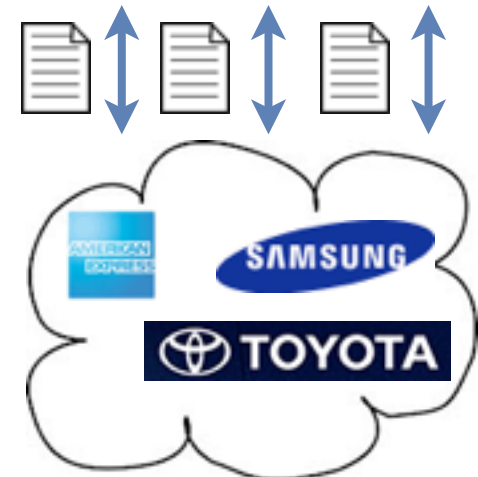
Black Women Want Top Jobs (but They Aren't Getting Them)

Programs that focus on convincing women to shoot for powerful jobs may not always be useful for black women.

- Cohen: Long Live Europe
- Kristof: Beyond Education Wars
- Bobby Jindal: I'm Holding Firm Against Gay Marriage
- Opinion on Facebook »

Watching

12m Eight women completed the initial, grueling four-day Ranger Assessment Phase, Army officials said. Ranger School opened to women for the first time ever this week. The Washington Post »



Ad Exchanges

- Significant fraction of the revenue of search engine and popular online sites:
 - Microsoft, Yahoo!, Google, OpenX, AppNexus.
 - Multi-billion dollar industry.
 - Choice of reserve price:
 - main mechanism through which the auction revenue can be influenced.
 - if set too low, winner may end up paying too little; if set too high, the ad slot could be lost.
- how can we select the reserve price to optimize revenue?

This Talk

- Learning formulation.
- Theoretical guarantees.
- Algorithms.
- Experimental results.

Previous ML Work

- Incentive compatible auctions (Balcan et al., 2008; Blum et al., 2004).
- Predicting bid landscapes (Cui et al., 2011).
- Revenue optimization for sponsored ads (Zhue et al., 2009; He et al., 2013; Devanur & Kakade, 2009).
- Bandit setting with no feature (Cesa-Bianchi et al., 2013; see also Ostrovsky & Schwarz, 2011).
- Strategic regret minimization (Amin et al., 2013; Munoz & MM, 2014).

Loss Function

- Auction revenue can be defined in terms of the pair of highest bids $\mathbf{b} = (b^{(1)}, b^{(2)})$:

$$\text{Rev}(r, \mathbf{b}) = b^{(2)} 1_{r < b^{(2)}} + r 1_{b^{(2)} \leq r \leq b^{(1)}}.$$

- Equivalently, loss define by

$$L(r, \mathbf{b}) = -\text{Rev}(r, \mathbf{b}).$$

Learning Formulation

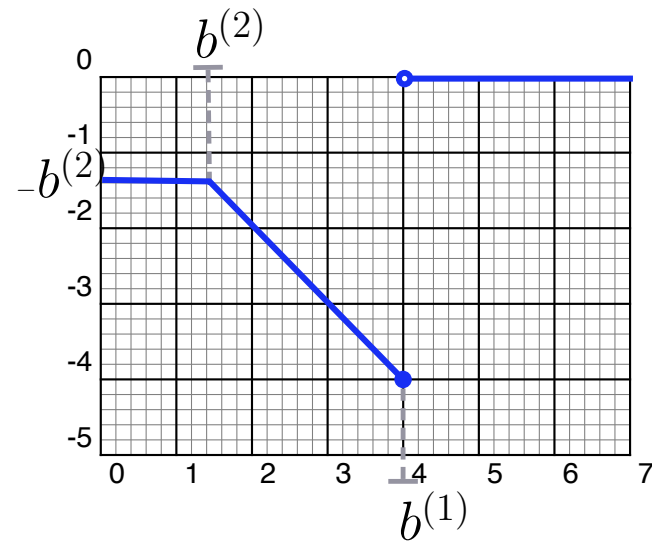
- $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^N$: public information about auction (features).
- $\mathcal{B} \subseteq \mathbb{R}_+^2$: bid space.
- $H \subseteq \mathbb{R}^{\mathcal{X}}$: hypothesis set.
- \mathcal{D} distribution over $\mathcal{X} \times \mathcal{B}$.
- **Problem**: find $h \in H$ with small generalization error,

$$\mathbb{E}_{(\mathbf{x}, \mathbf{b}) \sim \mathcal{D}} [L(h(\mathbf{x}), \mathbf{b})].$$

Loss Function

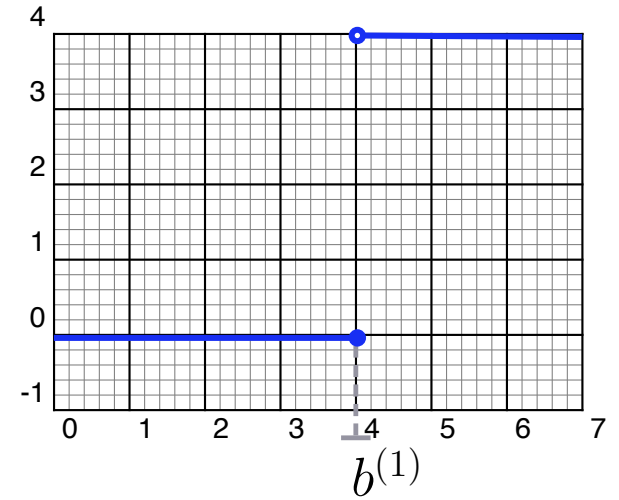
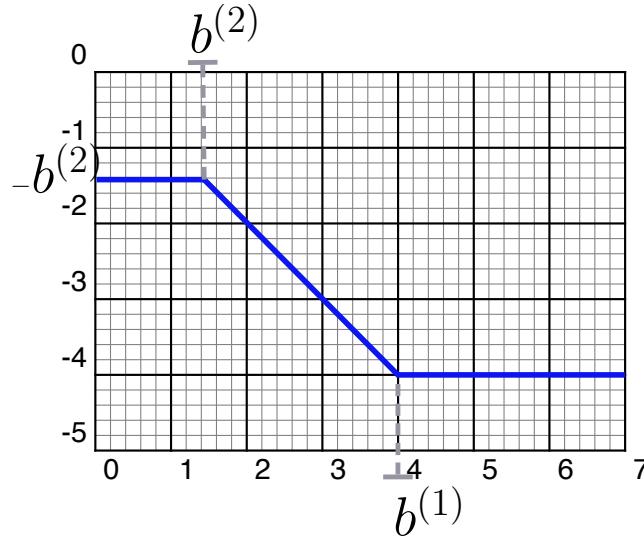
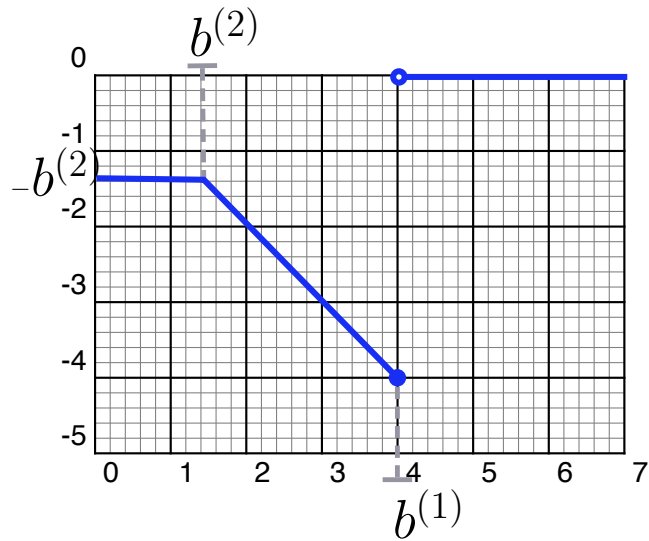
■ Properties:

- discontinuous.
- non-differentiable.
- non-convex.



Can we derive guarantees for learning with this loss function?

Loss Decomposition



Generalization Bound

- **Theorem:** let $M = \sup_{\mathbf{b} \in \mathcal{B}} b^{(1)}$ and let H be a hypothesis set with pseudo-dimension $d = \text{Pdim}(H)$. Then, for any $\delta > 0$, with probability $1 - \delta$ over the choice of a sample S of size m ,

$$\mathcal{L}(h) \leq \hat{\mathcal{L}}_S(h) + 2\mathfrak{R}_m(H) + 2M\sqrt{\frac{2d \log \frac{em}{d}}{m}} + M\sqrt{\frac{\log \frac{1}{\delta}}{2m}}.$$

Can we design algorithms minimizing the right-hand side?

This Talk

- Learning formulation.
- Theoretical guarantees.
- Algorithms.
- Experimental results.

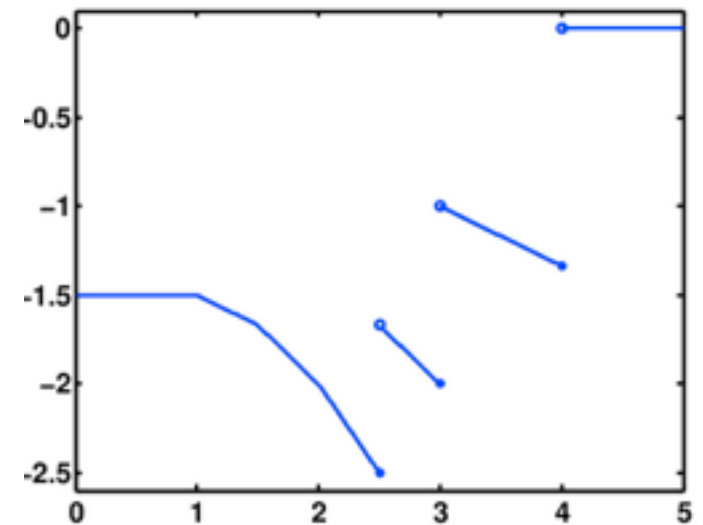
No Feature Case

■ **Problem:** find optimal reserve price,

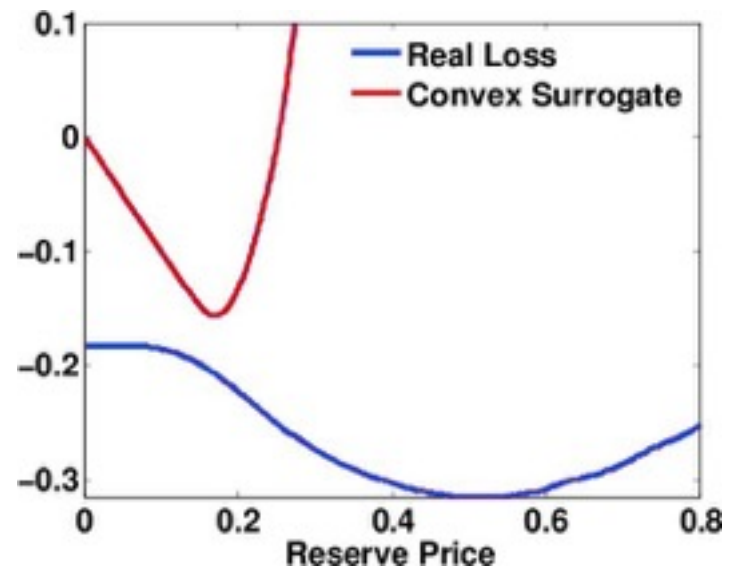
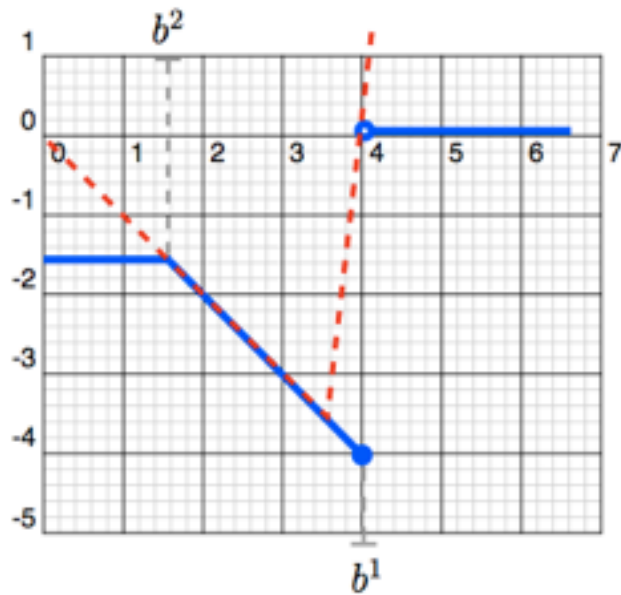
$$\min_{r \in \mathbb{R}} \sum_{i=1}^n L(r, \mathbf{b}_i).$$

■ **Algorithm:**

- optimum one of highest bids.
- naive in $O(m^2)$.
- sorting solution in $O(m \log m)$.



Convex Surrogate Loss

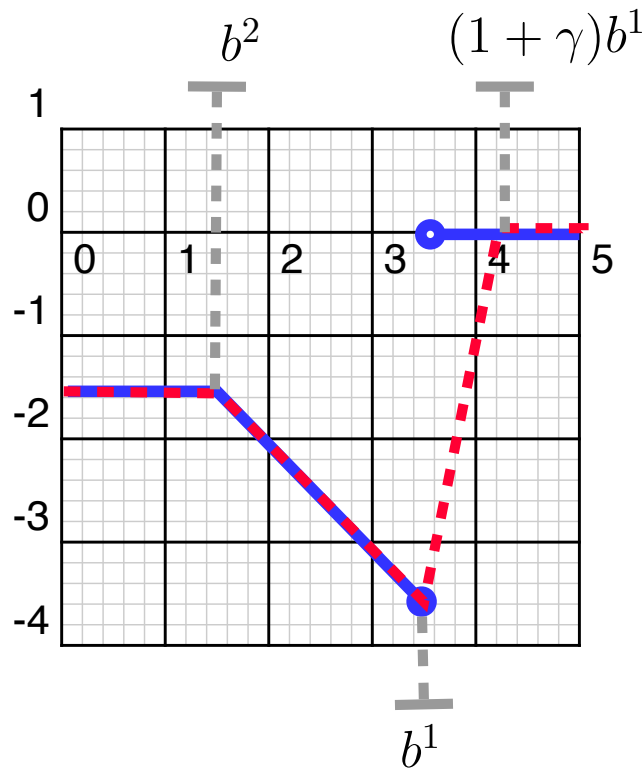


Convex Surrogate Loss

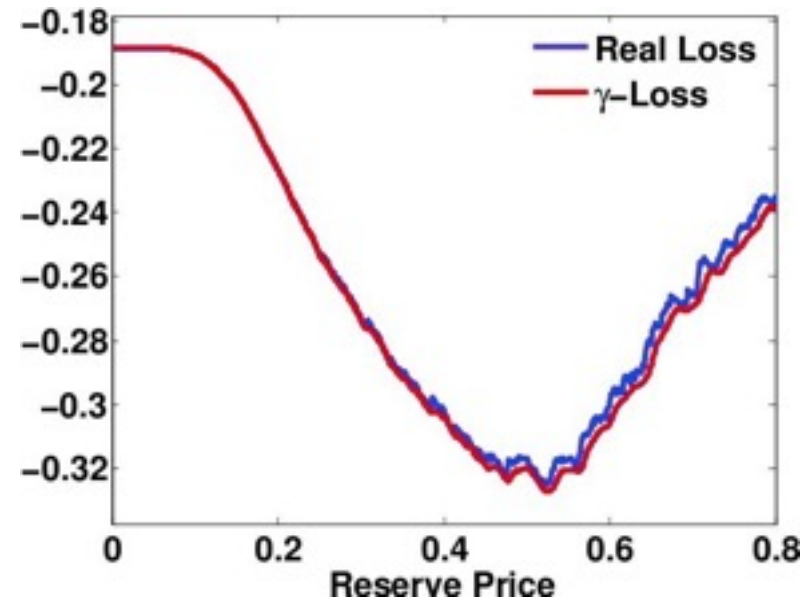
- No useful convex surrogate loss.
- **Theorem:** Let $L_c: [0, M] \times [0, M] \rightarrow \mathbb{R}$ be a bounded function, convex with respect to its first argument. If L_c is consistent with $(r, b) \mapsto -r1_{r \leq b}$, then $L_c(\cdot, b)$ is constant for every $b \in [0, M]$.

Which loss function should we use?

Continuous Surrogate Loss



loss function L_γ



Consistency Results

- **Theorem:** let $M = \sup_{\mathbf{b} \in \mathcal{B}} b^{(1)}$ and let H be a closed convex subset of a linear space of functions containing 0. Then,

$$\mathcal{L}(h^*) \leq \mathcal{L}(h_\gamma^*) \leq \mathcal{L}_\gamma(h_\gamma^*) + \gamma M.$$

Learning Guarantees

- **Theorem:** fix $\gamma \in (0, 1]$. Then, for any $\delta > 0$, with probability at least $1 - \delta$ over the choice of a sample S of size m ,

$$\mathcal{L}_\gamma(h) \leq \hat{\mathcal{L}}_{\gamma,S}(h) + \frac{2}{\gamma} \mathfrak{R}_m(H) + M \sqrt{\frac{\log \frac{1}{\delta}}{2m}}.$$

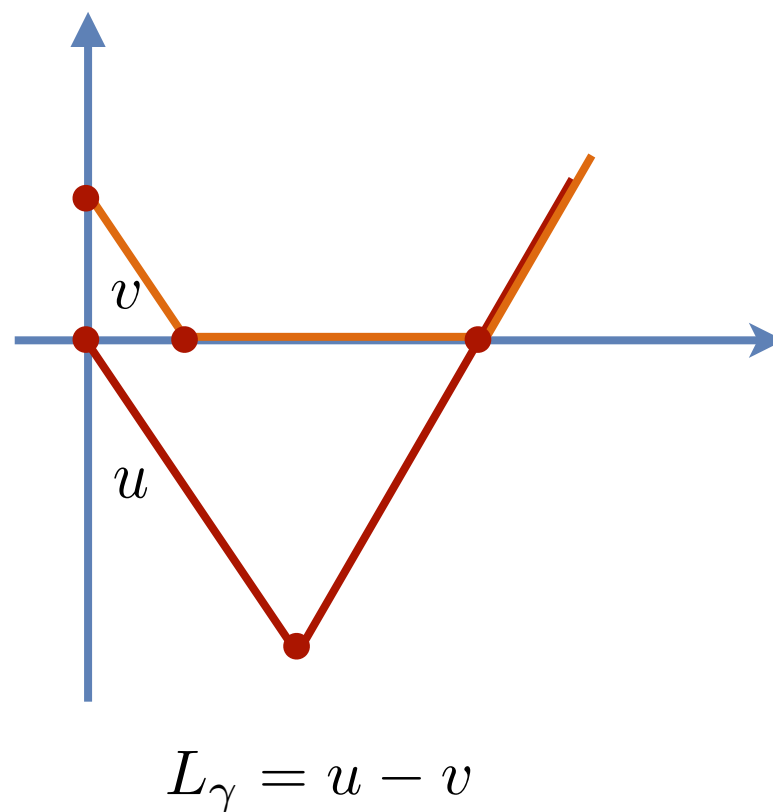
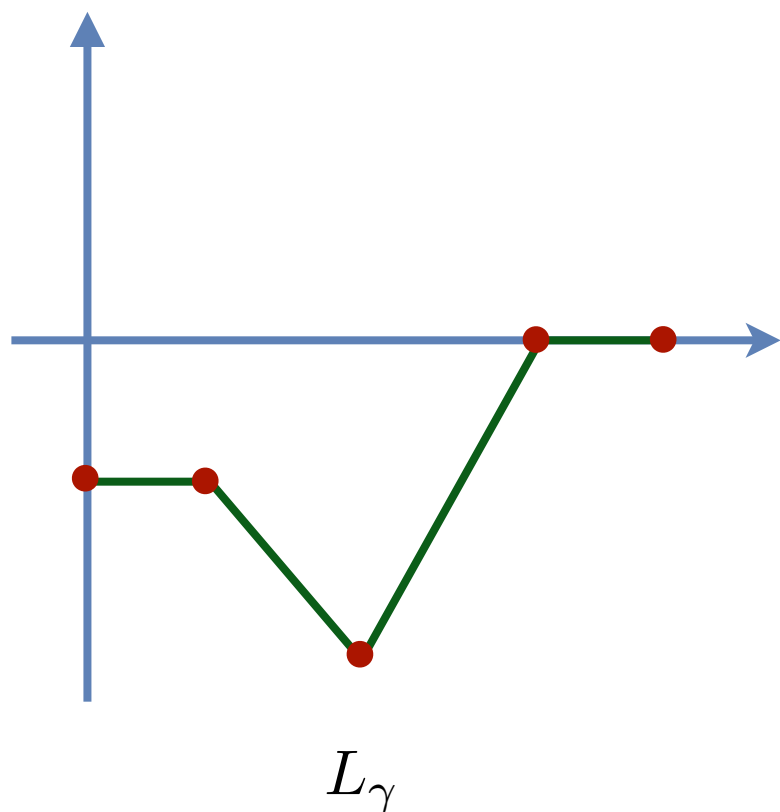
Algorithm

- Optimization problem: for fixed $\gamma \in (0, 1]$.

$$\min_{\|\mathbf{w}\| \leq \Lambda} \sum_{i=1}^m L_{\gamma}(\mathbf{w} \cdot \mathbf{x}_i, \mathbf{b}_i).$$

- difficulty: optimizing sum of non-convex functions.
- solution: DC-programming (Difference of Convex).

Difference of Convex Functions



DC-Programming

(Tao and Hoai, 1997; Yuille and Rangarajan, 2002)

- Convex-concave procedure: replace $F(\mathbf{w}) = f(\mathbf{w}) - g(\mathbf{w})$ at iteration $(t + 1)$ with upper bound

$$\hat{F}(\mathbf{w}) = f(\mathbf{w}) - g(\mathbf{w}_t) - \delta g(\mathbf{w}_t) \cdot (\mathbf{w} - \mathbf{w}_t),$$

with $\delta g(\mathbf{w}_t) \in \partial g(\mathbf{w}_t)$.

Algorithm

SECONDPRICERESERVE()

```
1   $\mathbf{w} \leftarrow \mathbf{w}_0$ 
2  for  $t \leftarrow 1$  to  $T$  do
3       $\mathbf{v} \leftarrow \text{DCA}(\mathbf{w}_{t-1})$ 
4       $\mathbf{u} \leftarrow \frac{\mathbf{v}}{\|\mathbf{v}\|}$ 
5       $\eta^* \leftarrow \min_{0 \leq \eta \leq \Lambda} \sum_{\mathbf{u} \cdot \mathbf{x}_i > 0} L_\gamma(\eta \mathbf{u} \cdot \mathbf{x}_i, \mathbf{b}_i)$ 
6       $\mathbf{w}_t \leftarrow \eta^* \mathbf{v}$ 
7  return  $\mathbf{w}$ 
```

Line Search

- Observation: L_γ is positive homogenous, for all $\eta > 0$,

$$L_\gamma(\eta r, \eta \mathbf{b}) = \eta L_\gamma(r, \mathbf{b}).$$

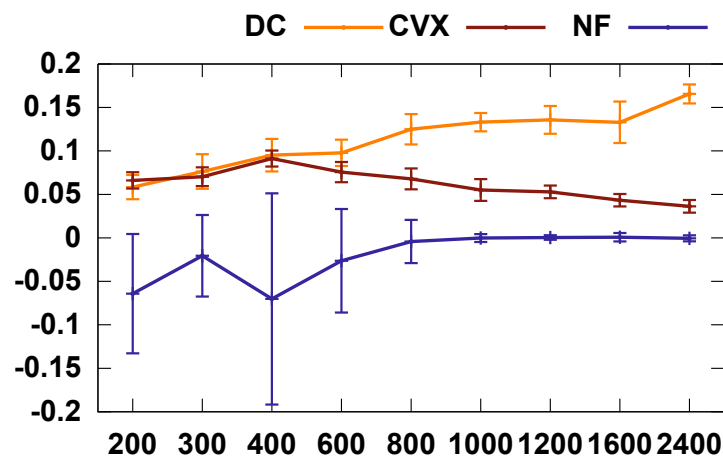
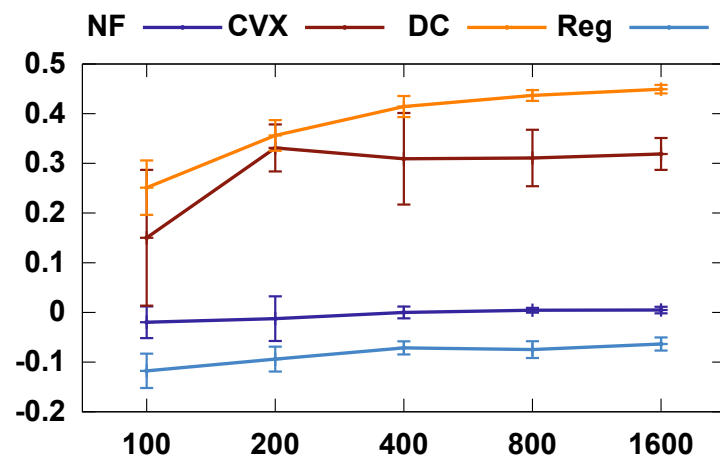
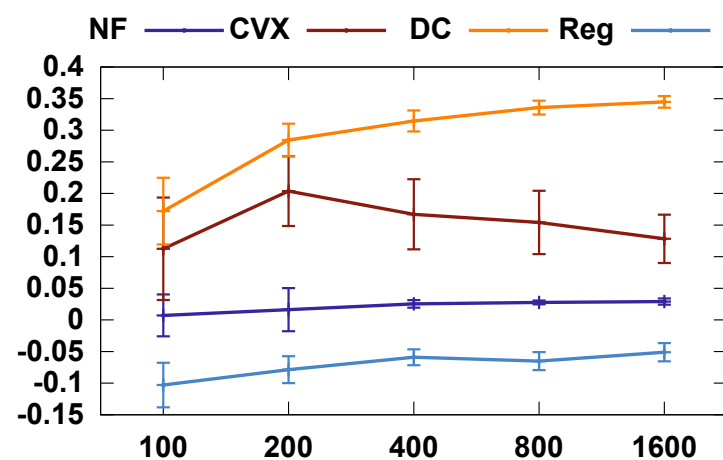
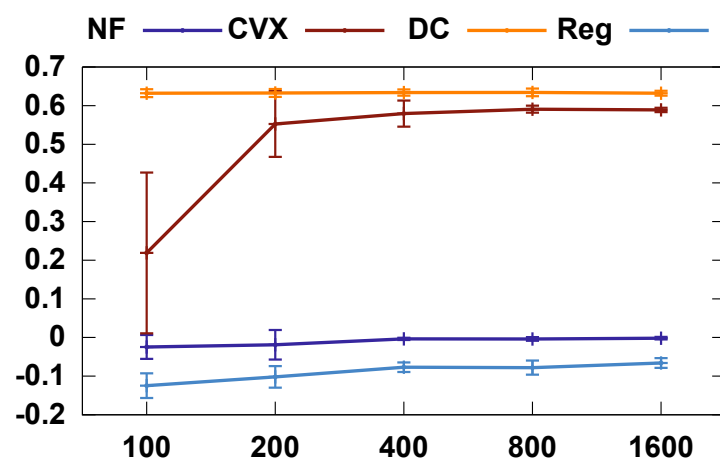
- Consequence: line search equivalent to no-feature minimization algorithm; for $\mathbf{w}_0^\top \mathbf{x}_i > 0$

$$\sum_{i=1}^m L_\gamma(\eta \mathbf{w}_0^\top \mathbf{x}_i, \mathbf{b}_i) = \sum_{i=1}^m (\mathbf{w}_0^\top \mathbf{x}_i) L_\gamma \left(\eta, \frac{\mathbf{b}_i}{\mathbf{w}_0^\top \mathbf{x}_i} \right).$$

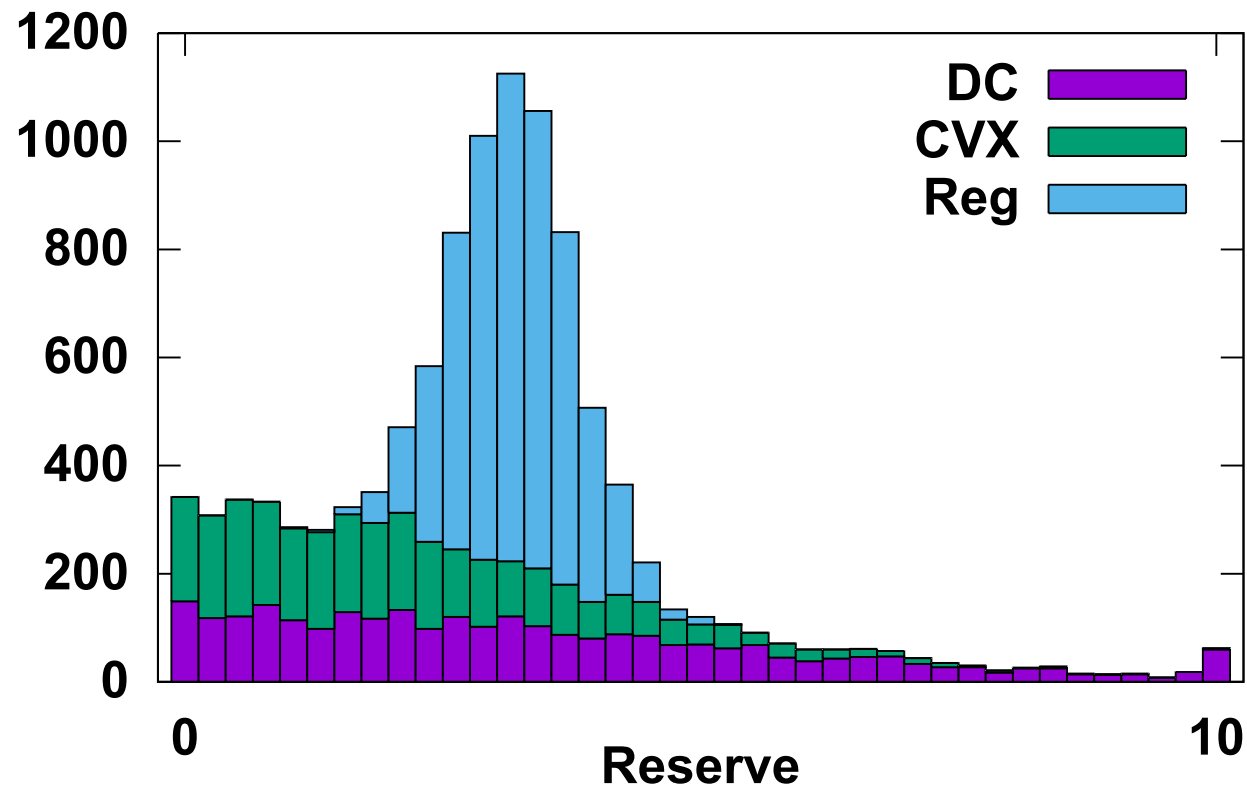
This Talk

- Learning formulation.
- Theoretical guarantees.
- Algorithms.
- Experimental results.

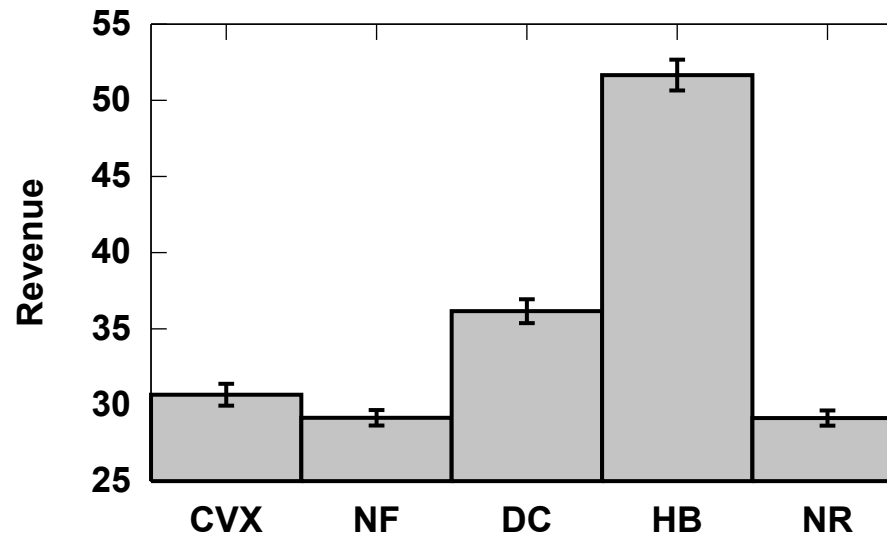
Experimental Results



Distribution of Reserve Prices



eBay Sport-Card Data Set



Data: <http://cims.nyu.edu/~munoz/data>.

Random 2000 training pts - 2000 test pts.

Conclusion

- Theory, algorithms, and experiments for second-price auctions with reserve.
 - scaling up DC algorithm.
 - study of dependencies.
 - effect of using revenue optimization algorithm.
 - better initialization.
- Learning and auctions:
 - many other scenarios and types of auctions.
 - Example: Generalized Second-Price auctions (GSPs).