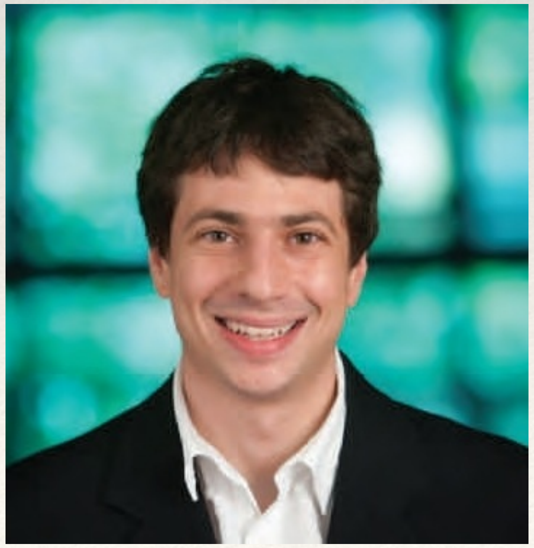
Watch and Learn: Optimizing from Revealed Preferences Feedback

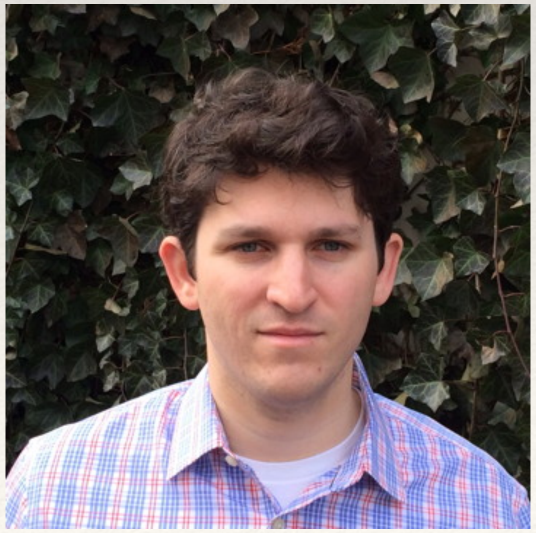
Watch and Learn:

Optimizing from Revealed Preferences Feedback

Steven Wu
University of Pennsylvania

Joint work with Aaron Roth and Jonathan Ullman





What kind of Data?

Prices

Purchase Behavior (Revealed Preferences)

Tolls

Traffic (Equilibrium Flow)

Learning from Revealed Preferences

- d divisible goods
- Producer posts prices:

$$p = (p_1, \ldots, p_d) \in \mathbb{R}^d_+$$

- Buyer purchases utility-maximizing bundle:

$$x^*(p) = \arg \max_{x \in C \subseteq \mathbb{R}^d_+} v(x) - \langle p, x \rangle$$

- v : valuation function ***unknown*** to producer;
- C is the set of feasible bundles
- v : Strongly concave & Lipschitz over C, Non-decreasing

Producer's Goal: Profit Maximization

Learning: The producer can adaptively set prices over rounds, and observe the purchased bundle by the buyer and the profit

Objective: Find the (approximately) optimal price vector under a small number rounds

$$\mathrm{Profit}(p) = \langle p, x^*(p) \rangle - c(x^*(p))$$

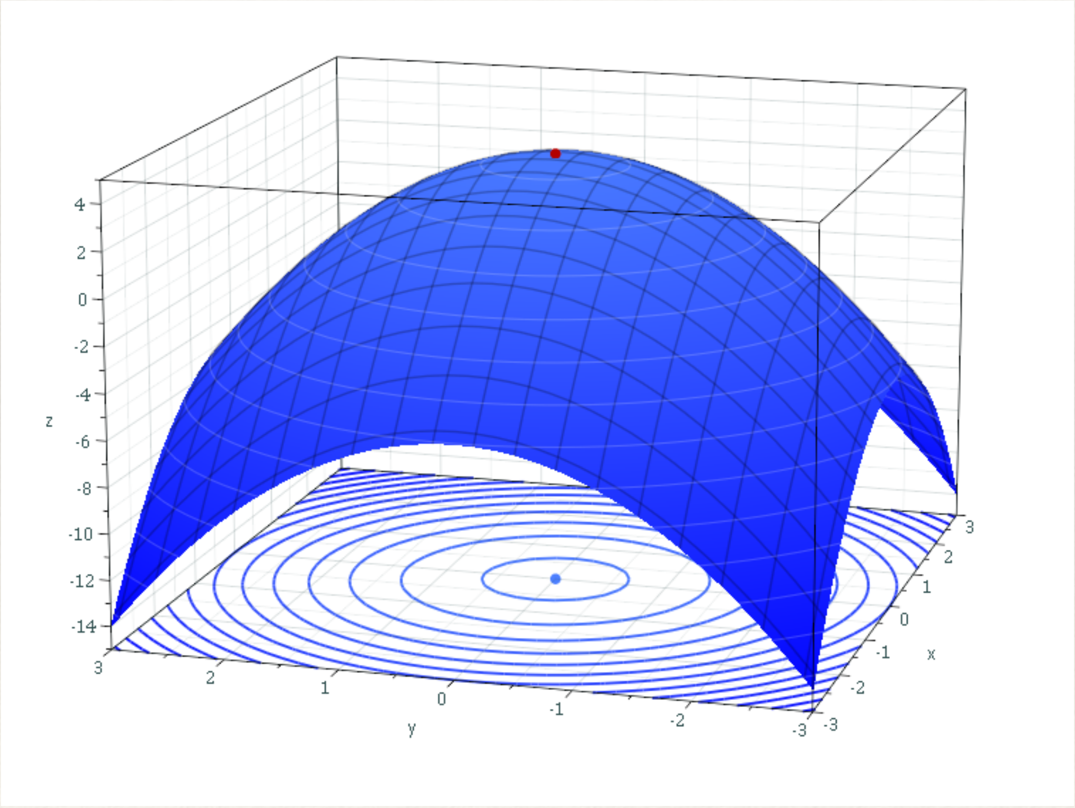
Revenue

Convex Production Cost

Unknown Objective

Zeroth Order Optimization?

Zeroth Order Optimization: given query access to an unknown **concave** function *f*, can find an approximately optimal solution with *poly*(*d*) queries



Unfortunately, the Profit function is not **concave** in the decision variables p

For example, if $v(x) = \sqrt{x}$

then $\text{Profit}(p) = \frac{1}{4p} - \frac{1}{4p^2}$

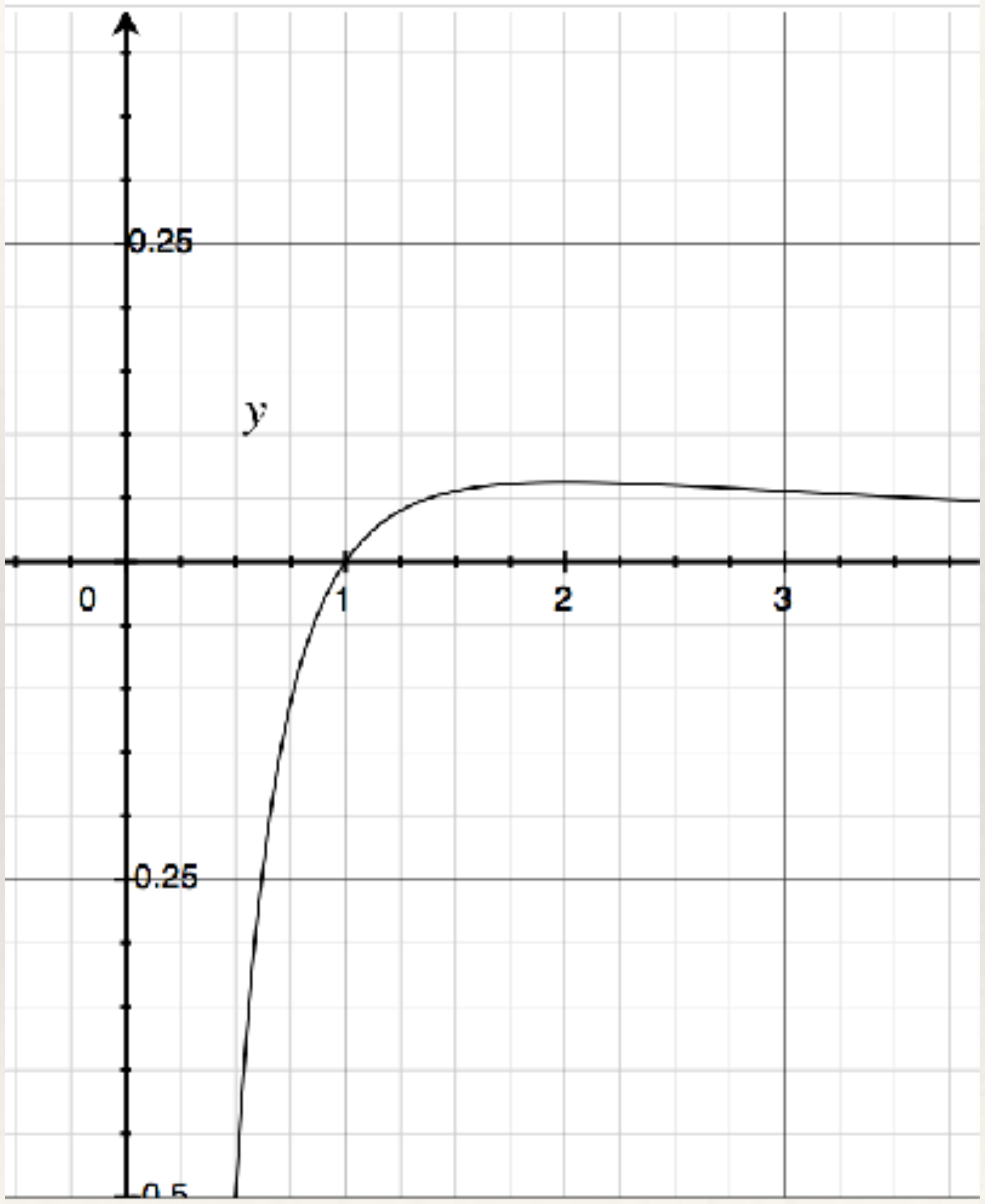
Illustration

$$v(x) = \sqrt{x}$$

$$x^*(p) = \arg\max_{x \in R^+} \sqrt{x} - p \cdot x$$

$$= \frac{1}{4p^2}$$

$$Profit(p) = \frac{1}{4p} - \frac{1}{4p^2}$$



Switching Decision Variables

- What if the producer could magically control what buyer buys (the variable *x*)?

Switching Decision Variables to Bundles

$$\mathrm{Profit}(x) = \max_{p: x^*(p) = x} \langle x, p \rangle - c(x)$$

$p^*(x)$ is the best price vector to induce bundle x

What is the best price vector?

- Lemma: if the buyer is allowed to buy nothing $(\mathbf{0} \in C)$, then

$$p^*(x) = \nabla v(x)$$

Now the profit is simpler!

$$\text{Profit}(x) = \langle x, \nabla v(x) \rangle - c(x)$$

simple is beautiful.

Is Profit(x) concave?

$$\text{Profit}(x) = \langle x, \nabla v(x) \rangle - c(x)$$

- The answer is yes for a large class of economically meaningful valuation functions — *homogeneous functions*

$$\exists k \geq 0, \quad v(a\, x) = a^k\, v(x)$$

- Scale Invariance: $x^*(p)$ is unchanged even switched to different units
- Example: CES & Cobb-Douglas

CES and Cobb-Douglas

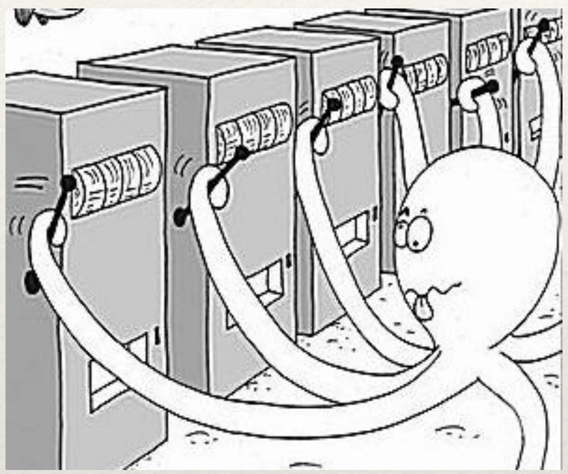
CES:

$$v(x) = \left(\sum_{i=1}^{d} \alpha_i x_i^{\rho} \right)^{\beta}$$

Cobb-Douglas:

$$v(x) = \prod_{i=1}^{d} x_i^{\alpha_i}$$

New Plan



At step t, query $\mathrm{Profit}(x^t)$

Return (approximate) evaluation for

$\mathrm{Profit}(x^t)$

Bandit Algorithm
Optimize over function
Profit(x)

Oracle

Optimal bundle x^*

Price vector to induce bundle x^*

Need to simulate query access to Profit(x)

Technical Problem

- Given a target bundle, find a price vector to (approximately) induce it.

$$\text{for any } \hat{x}, \text{ find } \hat{p} \text{ s.t. } \|\hat{x} - x^*(\hat{p})\| \leq \varepsilon$$

- Due to Lipschitzness,

$$\text{Profit}(\hat{x}) \approx \text{Profit}(x^*(\hat{p}))$$

Tâtonnement

For each good j, $$\hat{p}_j^{t+1} = \left[p_j - \eta(\hat{x}_j - x^*(p^t)_j)\right]$$
$$p^{t+1} = \Pi_P\left[\hat{p}^{t+1}\right]$$

Projected Gradient Descent

- If the buyer buys too much good *j*, raise the price
- If the buyer buys too little good *j*, lower the price

Why does it work?

- Consider the following convex program

$$\max_{x \in C} v(x)$$

$$\text{for each good } j \qquad x_j \leq \hat{x}_j$$

$$\hat{x} \text{ is the optimal solution}$$

- For each good, introduce a price (dual) variable

$$\text{Lagrangian: } \mathcal{L}(x, p) = v(x) - \langle p, x - \hat{x} \rangle$$

Lagrangian Zero-sum Game

- Strong duality

$$\max_{x \in C} \min_{p \in \mathbb{R}^d_+} \mathcal{L}(x,p) = \min_{p \in \mathbb{R}^d_+} \max_{x \in C} \mathcal{L}(x,p) = v(\hat{x})$$

- Minimax theorem continues to hold for

$$P = \{p \in \mathbb{R}^d_+ \mid \|p\| \leq \sqrt{d}\}$$

$$\max_{x \in C} \min_{p \in P} \mathcal{L}(x,p) = \min_{p \in P} \max_{x \in C} \mathcal{L}(x,p) = v(\hat{x})$$

Payoff $\mathcal{L}(x,p)$

Price Player p

Bundle Player x

No Regret vs. Best Response

For $t = 1, \ldots, T$

Price Player plays

online gradient descent

$$p^{t+1} = \Pi_P \left[p^t - \eta(\hat{x} - x^t) \right]$$

Bundle Player plays

best response

$$x^t = \arg\max_{x \in C} \mathcal{L}(x, p^t)$$

Rewrite the Best Response

$$x^t = \arg\max_{x \in C} \left[v(x) - \langle p^t, x - \hat{x} \rangle \right]$$

$$= \arg\max_{x \in C} \left[v(x) - \langle p^t, x \rangle + \langle p^t, \hat{x} \rangle \right]$$

We could always remove the constants in argmax

$$x^t = \arg\max_{x \in C} \left[v(x) - \langle p^t, x \rangle \right] = x^*(p^t)$$

Best response is just
the observed purchased bundle!

No Regret vs. Best Response

Price Player plays
online gradient descent

$$p^{t+1} = \Pi_P \left[p^t - \eta(\hat{x} - x^*(p^t)) \right]$$

Bundle Player plays
best response

$$x^*(p^t)$$

The average plays $p' = 1/T \sum_t p^t$ and $x' = 1/T \sum_t x^t$
forms approximate minimax equilibrium [FreudSchapire'96]

Approximate Equilibrium

Lemma: Let (x', p') be any approximate minimax equilibrium of the Lagrangian zero-sum game, then

$$\|\hat{x} - x^*(p')\| \leq \varepsilon$$

Proof idea: equilibrium condition and strong concavity

$$\mathcal{L}(\hat{x}, p') \approx \mathcal{L}(x^*(p), p')$$

Our Roadmap

Profit(p) non-concave objective

switch decision variables to *bundles*

Profit(x) concave (Homogeneous v)

optimize over x

Bandit Algorithm [BLNR'15]

requires query access Profit(x)

Learning dynamics (OGD by [Zin'03])

Optimizing Traffic Routing from Revealed Behavior

- The same framework can find tolls to induce an approximately *optimal* traffic flow

- In the process, we also solve the problem of finding tolls to induce target flow introduced by [BLSS'14]

Stackelberg Games

- More general settings: a class of ***Stackelberg Games***:
  - Optimize *leader's* utility given observations on the *follower's actions*
  - e.g. Contract design in principal-agent Problems

- Extension to noisy observations

Open Problem

- Stochastic Revealed Preferences
  - producer sets prices p, and a buyer with valuation v drawn from some unknown prior
  - Goal: find the prices that maximize expected profit

Time for Coffee and more Problems!

Watch and Learn:

Optimizing from Revealed Preferences Feedback

Steven Wu
University of Pennsylvania

Joint work with Aaron Roth and Jonathan Ullman