CS 6840 Algorithmic Game Theory November 6, 2024

Lecture 29: No-Envy Learning in Multi-Item Auctions

Instructor: Èva Tardos *Scribe: Ved Sriraman*

Introduction and Motivation

In the previous lecture, we discussed the complexity of achieving no-regret learning in multi-item auctions where each bidder wants a single item. In such settings, ensuring no-regret is NP-Hard when the number of items is large. In fact, even running a single step of Follow The Perturbed Leader (FTPL) becomes computationally infeasible due to the exponential size of the action space.

In today's lecture, we introduce a more tractable alternative called *no-envy learning*. This framework relaxes the stringent requirements of no-regret learning by allowing bidders to compare their cumulative utility against historical averages, rather than aiming for the best possible utility in hindsight. We will see that despite this relaxation, no-envy learning can still guarantee approximate social welfare optimality and can be efficiently implemented in polynomial time.

Auction Model and Definitions

Definition 1 (Multi-Item Auction)**.** Recall the setup in a multi-item auction with multiple bidders. Over a set of items A, each bidder i has a value $v_i(A) = \max_{j \in A} v_{ij}$. This is saying that when multiple items are won, the bidder keeps the highest-valued item and discards the rest (free disposal). Each bidder's objective is to maximize utility, defined as their value for the item they retain minus the price paid.

In no-regret learning, fluctuating prices made it difficult for our learner to compete, as a median bid could lead to winning some items and not others. Instead, we will consider a more manageable benchmark in that the bidders will compare their utility against historical prices. We formalize the no-envy condition as follows:

Definition 2 (No-Envy Benchmark)**.** A no-envy benchmark allows a bidder to compare their utility to what they would have achieved with the best possible single-item choice under historical averages. Formally, the no-envy condition holds if:

$$\sum_{t=1}^{T} u_i^t(s^t) \geq \max_{j \in A} \left(\lambda \cdot \sum_{t=1}^{T} v_{ij}^t - \sum_{t=1}^{T} p_j^t \right) - \mathsf{Regret} \tag{1}$$

where $u_i^t(s^t)$ is the utility of bidder i at time t with strategy s^t, $v_{ij}^t = v_{ij}$ (like last time) is the value of item j at time t, p_j^t is the price of item j (as it was in the auction) at time t, and $\lambda \in (0, 1]$ is a smoothness factor.

The no-envy benchmark is more permissive than the no-regret benchmark because it only requires the bidder to compete against the best fixed bundle (single item) at the average price, rather than all possible strategies. More formally, we require that the average utility of the bidder across T rounds is within $o(1)$ of what he would have achieved by purchasing the optimal bundle at its average price in hindsight.

Price of Anarchy Bound for No-Envy Learning

Let the social welfare (SW) be the sum of the utilities of all the bidders and the revenue of the auction. Then, the social welfare achieved by the bidders is given by:

$$SW = \sum_{t=1}^{T} \sum_{i=1}^{n} u_i^t(s^t) + \mathsf{Rev}.$$

Theorem 1 (PoA Bound under No-Envy)**.** *If each player achieves no-envy, then the total social welfare over T auction rounds satisfies:*

$$SW \geq \lambda T \cdot \mathsf{OPT} - n \cdot \mathsf{Regret},$$

where OPT *is the welfare under an optimal matching, and each of the n players suffers a little bit of regret.*

Proof. We assume that each bidder's valuations are fixed over time, i.e., $v_{ij}^t = v_{ij}$ for all t. This allows each bidder to compare their utility against the best single-item choice over time. Each bidder knows their own valuations for the items.

Let M be the optimal matching that maximizes the total welfare, i.e., $\mathsf{OPT} = \sum_{(i,j) \in M} v_{ij}$. While we use M in our analysis, bidders do not need to know this matching. By the no-envy condition (1) for each bidder i, we have:

$$\sum_{t=1}^{T} u_i^t(s^t) \geq \sum_{t=1}^{T} \left(\lambda \cdot \sum_{t=1}^{T} v_{ij_i} - \sum_{t=1}^{T} p_{j_i}^t - \mathsf{Regret} \right)$$

where j_i is the item assigned to bidder i in the optimal matching M. For bidders not assigned an item in M, we have $v_{ij_i} = 0$ and their utility is non-negative: $u_i^t(s^t) \geq 0$.

Summing over all bidders i, we get:

$$\sum_{i=1}^{n} \sum_{t=1}^{T} u_i^t(s^t) \geq \lambda \cdot T \cdot \sum_{i=1}^{n} v_{ij_i} - \sum_{i=1}^{n} \sum_{t=1}^{T} p_{j_i}^t - n \cdot \mathsf{Regret} = \lambda \cdot T \cdot \mathsf{OPT} - \sum_{t=1}^{T} \sum_{i=1}^{n} p_{j_i}^t - n \cdot \mathsf{Regret}.$$

Note that the sum $\sum_{i=1}^{n} p_{j_i}^t$ is the total revenue collected from the items in the optimal matching at time t. The total revenue collected over all items and all rounds is:

$$\mathsf{Rev} = \sum_{t=1}^{T} \sum_{j=1}^{m} p_j^t \geq \sum_{t=1}^{T} \sum_{j=1}^{n} p_{j_i}^t$$

Therefore,

$$\begin{aligned} \mathrm{SW} &= \sum_{i=1}^{n} \sum_{t=1}^{T} u_i^t(s^t) + \mathsf{Rev} \\ &\geq \left[\lambda \cdot T \cdot \mathsf{OPT} - \sum_{t=1}^{T} \sum_{i=1}^{n} p_{j_i}^t - n \cdot \mathsf{Regret} \right] + \mathsf{Rev} \\ &= \lambda \cdot T \cdot \mathsf{OPT} + \underbrace{\left(\mathsf{Rev} - \sum_{t=1}^{T} \sum_{i=1}^{n} p_{j_i}^t \right)}_{\geq 0} - n \cdot \mathsf{Regret} \end{aligned}$$

$$\geq \lambda \cdot T \cdot \mathsf{OPT} - n \cdot \mathsf{Regret},$$

since $\mathsf{Rev} \geq \sum_{t=1}^{T} \sum_{i=1}^{n} p_{j_i}^t$. ■

Algorithms for No-Envy

Having established that no-envy dynamics can lead to approximately optimal social welfare, we now focus on how bidders can achieve no-envy learning in practice. In this first section, our objective is to apply a modified version of *Follow the Leader* that enables bidders to select single items maximizing utility against historical average prices.

Be The Leader

Similar to previous lectures, we will first consider the "slight cheat" version of the algorithm, where we know what other bidders bid at the current time step.

Let us set $\lambda = \frac{1}{2}$. In this setting, we will try to find an item set S at time t that maximizes the following value:

$$\max_S \frac{1}{2} \cdot v_i(S) - \sum_{j \in S} \bar{p}_j^t$$

where $\bar{p}_j^t$ is the average price of item j up to time t, $\sum_{\tau=1}^{t} p_j^\tau$. Note that the bidder will suggest the set of items with smallest average price relative to $\frac{1}{2}$ his value for the bundle. Therefore, the bidder selects the item j^t that maximizes $\frac{1}{2} \cdot v_{ij} - \bar{p}_j^t$, and bids $\frac{1}{2} v_{ij}$ on item j^t, and zero on all other items.

Proof Sketch for Follow the Perturbed Leader

We will provide a proof sketch, which borrows elements from the proof of Follow the Perturbed Leader (FTPL) in Lecture 21. However, we omit some of the details that were not covered in class.

Step 1. Show that Be The Leader is a no-envy algorithm. That is,

$$\sum_{\tau=1}^{T} u_i^\tau(s^\tau) \geq \max_j \frac{1}{2} T \cdot v_{ij} - \sum_{\tau=1}^{T} p_j^\tau$$

Proof. This mirrors the style of the proof of Claim 1 in Lecture 21. The proof is by induction.

<u>Base case.</u> At $t = 1$, the bidder selects the item j^1 that maximizes $\frac{1}{2} v_{ij} - p_j^1$, i.e.,

$$j^1 = \operatorname*{argmax}_j \frac{1}{2} v_{ij} - p_j^1$$

The bidder bids $\frac{1}{2} v_{ij^1}$ on item j^1 and 0 on all other items.

For any item j, if the bidder bids $\frac{1}{2} v_{ij}$ on item j and 0 on all others, the utility satisfies:

$$u_i^1(s_i^1) \geq \frac{1}{2}v_{ij} - p_j^1$$

This holds because:

- If the bidder wins item j, the utility is $v_{ij} - p_j^1 \geq \frac{1}{2}v_{ij} - p_j^1$.
- If the bidder does not win item j, the utility is $0 \geq \frac{1}{2}v_{ij} - p_j^1$ (since $p_j^1 \geq \frac{1}{2}v_{ij}$ for not winning).

Since this inequality holds for all j, it holds in particular for our chosen item j^1:

$$u_i^1(s_i^1) \geq \frac{1}{2}v_{ij^1} - p_{j^1}^1$$

Inductive step: Assume that the no-envy condition holds up to time $t-1$. At time t, the bidder selects the item j^t that maximizes $\frac{1}{2}v_{ij} - \bar{p}_j^t$, where $\bar{p}_j^t = \frac{1}{t}\sum_{\tau=1}^t p_j^\tau$.

As in the base case, for any item j, bidding $\frac{1}{2}v_{ij}$ on item j yields utility:

$$u_i^t(s_i^t) \geq \frac{1}{2}v_{ij} - p_j^t$$

Therefore, for our chosen item j^t, we have:

$$u_i^t(s_i^t) \geq \frac{1}{2}v_{ij^t} - p_{j^t}^t$$

Now, the cumulative utility up to time t satisfies:

$$\begin{aligned}\sum_{\tau=1}^t u_i^\tau(s_i^\tau) &= \left(\sum_{\tau=1}^{t-1} u_i^\tau(s_i^\tau)\right) + u_i^t(s_i^t) \\ &\geq \left(\frac{1}{2}(t-1)v_{ij^t} - \sum_{\tau=1}^{t-1} p_{j^t}^\tau\right) + \left(\frac{1}{2}v_{ij^t} - p_{j^t}^t\right),\end{aligned}$$

where we used the inductive hypothesis for $t-1$ and $j = j^t$ the fact that

$$u_i^t(s_i^t) \geq \frac{1}{2}v_{ij^t} - p_{j^t}^t$$

(as shown above). Summing, we have:

$$\sum_{\tau=1}^t u_i^\tau(s_i^\tau) \geq \frac{1}{2}t \cdot v_{ij^t} - \sum_{\tau=1}^t p_{j^t}^\tau.$$

Finally, since j^t maximizes $\frac{1}{2}v_{ij} - \bar{p}_j^t$, it follows that:

$$\sum_{\tau=1}^t u_i^\tau(s_i^\tau) \geq \max_j \left(\frac{1}{2}t \cdot v_{ij} - \sum_{\tau=1}^t p_j^\tau\right).$$

This completes the induction.

Step 2. Next, we mirror Claim 2 from Lecture 21. For each item, we will let the learner generate a random noise term ξ_j at the first timestep, and then choose the item such that $j^t = \text{argmax}_j \sum_{\tau=1}^{t} u^\tau(s^\tau) + \xi_j$. Essentially, we keep the "cheating" and add in some noise. We can then show the following:

$$\sum_{\tau=1}^{T} u_i^\tau(s^\tau) + \max_{j \in S} \xi_j = \max_S \left(\lambda t \cdot v_i(S) + \sum_{j \in S} \xi_j - \sum_{\tau=1}^{T} p_j^\tau \right) \geq \max_j \left(\lambda T \cdot v_{ij} + \xi_j - \sum_{\tau=1}^{T} p_j^\tau \right)$$

where the last inequality follows from the proof above. The proof of this follows directly from the proof of Claim 2 in Lecture 21.

Step 3. The real algorithm needs to select an item at time t without knowing the prices at time t. We can reconcile this "cheating" by using noise to "cover up" the fact that we are missing timestep t in comparison to FTL. We can show that if the noise is chosen after all the other bidders behave (they are independent), then the probability that we choose the same item at time t and time $t-1$ is large enough. That is, if $\xi_s \sim Geo(\epsilon)$ is chosen randomly and independent of s_t, then $\Pr(j^t = j^{t-1}) \geq 1 - \epsilon$.

Step 4. After this, we must consider a better version of the cheat algorithm that generates noise at each timestep (generating ξ_s^t's). To prove the final result, we must make sure the noise remains independent of the previous choices so that we can choose the randomness every step, which relies on linearity of expectation. See Theorem 7 and Lemma 8 from Lecture 21 for more details. ■

Concluding Remarks

- **Relaxing the Benchmark**: No-envy is more computationally feasible than no-regret and still approximates socially optimal behavior.
- **Item-Level Noise**: We did not add as much noise as before, adding noise on the items as opposed to the subsets. Adding noise at the item level suffices to maintain no-envy in most scenarios, especially when strategy spaces expand exponentially.
- **Extensions**: Though we focused on first-price auctions, the no-envy benchmark and item-level noise can extend to other auction formats, like second-price auctions. Refer to the original paper for details.