Algorithms, Games, and Networks March 7, 2013

Lecture 15

Lecturer: Avrim Blum *Scribe: Christian Kroer*

1 Unlimited supply revenue maximization (digital good pricing)

In this lecture we consider the case where we have an unlimited supply of one or more types of items that we want to sell, and the goal is to set the price of the items such that we maximize revenue. It is immediately obvious that if we want to maximize social welfare we can give everyone their desired items at a price of 0.

1.1 Single item

We first consider the case where we're selling a single item with unlimited supply. Let's assume that valuations are between 1 and H. The question we will examine is what fraction of the social welfare can we hope to extract as revenue?

First, one useful thing to notice is that if buyers were arriving from a fixed probability distribution, then we would optimize revenue by charging:

$$\arg\max_p [p \cdot \Pr(v_i \geq p)]. \tag{1}$$

What can we say about how this revenue compares to social welfare? First, we have the following lower bound.

Claim 1 *There exist distributions over* $[1, H]$ *such that the expected revenue for any price is at most a* $\frac{1}{\log H}$ *fraction of the expected social welfare*

Proof: Consider buyers drawn from the following distribution. For each buyer, her valuation is drawn uniformly from $H, H/2, H/3, ..., H/H$.

If we set the price to $p = \frac{H}{i}$, then the probability that a buyer will have valuation $v \geq p$ is $\frac{i}{H}$, hence we get expected revenue of 1. We can do no better than this, as choosing a price $\frac{H}{i} > p > \frac{H}{i+1}$ has the same probability for $v \geq p$ as $\frac{H}{i}$, and hence can only yield lower revenue.

But the expected social welfare is $1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{H} \approx \log H$, and hence we get the result. ■

We now give an upper bound: in fact, a randomized algorithm such that no matter what the buyer's valuation—so long as it is in the range $[1, H]$—the expected revenue is within $\Omega(1/\log H)$ of the social welfare (i.e., of the buyer's valuation).

Claim 2 *There exists a randomized algorithm that for any buyer (or any sequence of buyers) gets revenue of* $\Omega(\frac{1}{\log H})$ *of the optimal social welfare.*

Proof: In order to do this we need a randomized strategy, since for any fixed price, we can have an adversary that gives us exactly the guy who has valuation just too low to buy.

We start out by noting that we would like to be able to give a price between $\frac{v}{2}$ and v. In order to do this, we pick uniformly from powers of 2 in $[1 : \frac{H}{2}]$: $1, 2, 4, 8, 16, ..., \frac{H}{2}$. This gives us a $\frac{1}{log_2 H}$ chance of picking a number in $[\frac{v}{2}, v]$, and hence we get expected revenue of $\frac{1}{log_2 H} \cdot \frac{v}{2}$. ■

1.2 Multiple items

Now, we consider the case where we have n items. Buyers have an arbitrary valuation function over subsets $v_i : S \subseteq \{1, 2, ..., n\} \to \mathbb{R}^+$. Where the maximum valuation for any subset is $1 \leq \max_S v_i(S) \leq H$.

First, suppose we were allowed to assign bundle prices. Then, this problem can be converted to the original setting, by bundling everything together and selling it as a single item.

However, suppose we can only assign prices to items, not bundles. It turns out that we can do almost as well. We can get $\Omega(\frac{1}{\log(nH)})$ of the optimal social welfare.

Before describing the algorithm, we first present a lower bound showing why the $1/\log n$ term is necessary: in particular, this is needed even if you know the buyers' valuations completely. Specifically, suppose the following. Any one item is worth 1 to the buyers. Any two items is worth $1 + \frac{1}{2}$ to the buyers. Any three items is worth $1 + \frac{1}{2} + \frac{1}{3}$ to the buyers, and so on.

If we price everything at \$1 dollar, we make \$1 per buyer, as everybody only wants to buy one. If we price everything at \$0.5, then we make \$1 per buyer, since everybody wants two items.

In fact, we can never make more than \$1. For the buyer, she always want to sort items by price, and buy cheapest first. She only buys the i'th item if its price is less than or equal to $\frac{1}{i}$, and this means that any previously bought items are sold for at most this much. On the other hand, the maximum social welfare for each buyer is $1 + 1/2 + 1/3 + \ldots + 1/n \approx \log(n)$.

We now present the upper bound.

Consider one buyer and the function $f(p) : \mathbb{R} \to \mathbb{Z}^+$, giving the number of items purchased for any given price p. This function is a step function, and an example is shown in Figure 1,

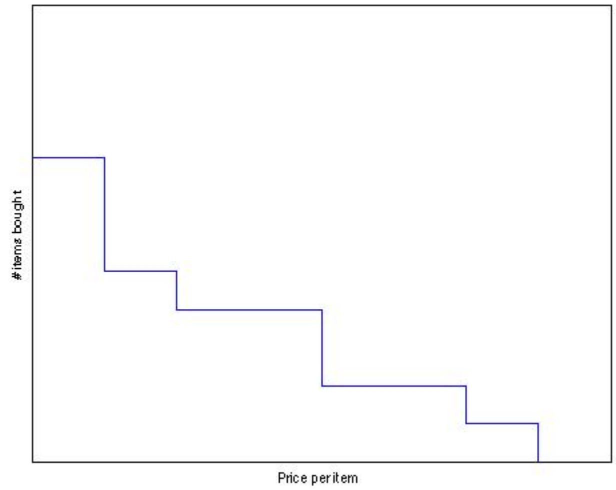

Figure 1: A function showing the number of items a buyer would purchase as a function of the price.

where the x axis is the price of the item and the y axis is the number of items purchased.

Claim 3 *$f(p)$ is non-increasing.*

Proof: For any given price p, let S be the bundle purchased at the price. For any bundle S' such that $|S'| > |S|$, the bundle only becomes less attractive compared to S as prices are increased, since an increase in price Δp increases the price of S' by $|S'| \cdot \Delta p$, whereas the price of S only increases by $|S| \cdot \Delta p$. Hence, if the buyer was not buying the set of bigger size before, she will not want to switch to that set once the price increases, as his relative utility for the previously preferred smaller set has increased. ■

Claim 4 *$\max_S v(S)$ is equal to the area under the curve of $f(p)$.*

Proof: The buyer's utility at price p is $\max_S[v(S) - p|S|]$. So, consider price $p = 0$ where the buyer's utility is $\max_S v(S)$. If we increase the price from zero to p_1, where we don't change the bought set, then the buyer's happiness decreases by $p_1 \cdot |S_1|$, where S_1 is $\arg\max_S[v(S) - p|S|]$. This is the the area under the curve for the interval $[0; p_1]$. Now suppose that we increase the price until we hit the first step. Then the buyer is indifferent between S_1 and S_2, where S_2 is the preferred set at this price. Say that the buyer switches to buying S_2. We can then continue increasing prices, and again for each increase in price Δp, the decrease in utility of the buyer is exactly the area swept out, $|S_2|\Delta p$, until again the buyer switches to some new set S_3. More generally, the decrease in utility to the buyer caused by increasing prices by some Δp for a given set of items purchased is exactly the area swept out, and when the buyer changes sets at some price where she is indifferent there is no change in utility. So, continuing this process until utility hits zero, we have covered the entire area under the curve. Therefore the area under the curve equals the initial utility, $\max_S v(S)$. ■

Now, for every price p, we earn $p \cdot \#items\ bought$. We choose a price uniformly from $H, H/2, H/4, ..., \frac{1}{4n}$.

The sum of revenues for each price is then $\geq \frac{1}{2}$(area under curve)$-\frac{1}{4}$, since for each maximal price p where a given set S is bought, we have some price H/i such that it gets at least half the price, and at least that big of a set is bought. The term "$-\frac{1}{4}$" is due to the fact that we stop at price $\frac{1}{4n}$, so we potentially lose the area of a rectangle of that width and height at most n. It follows that picking uniformly between these prices gives us average revenue $\frac{\frac{1}{2}(social\ welfare)-\frac{1}{4}}{\log(4nH)}$.

Comparison to revenue from best set of item prices

We are going to look at the following problem:

Each buyer is single-minded over sets of size at most 2, where the valuations are known. We can represent this as a graph, where items are vertices, and buyers are edges with their valuation being the edge weight. Now we want to set prices on vertices such that we maximize revenue. We are going to compare our performance with the optimal revenue achievable by any item pricing.

It is NP-hard to find the best set of prices yielding optimal revenue OPT. However, there is a fast algorithm for getting $\geq \frac{1}{4}OPT$. First, the vertices are randomly split into two groups. We then have 2 types of bidders. Those who want an item from each of the two groups, and those who are only interested in items in their own group. For each edge, there's a 50% chance that the edge will be between the two groups. Thus, in expectation half the revenue is made from people who want an item from each group. Let OPT' denote this revenue, we have $E[OPT'] = OPT/2$. The algorithm will optimize revenue for the edges going across. This immediately costs us half the revenue, as the other edges are ignored.

The revenue-maximizing pricing scheme over just the edges between the groups will make some amount of revenue from the first group, call this OPT_1, and some amount of revenue from the second group, call this OPT_2. The max of these will be at least $OPT'/2$. Notice that if we take these optimal prices and zero out the prices in the first group, then the revenue made will be at least OPT_2 since all previous buyers still buy and perhaps new buyers as well. This is good because we can efficiently determine the best prices for the second group (given that the first group is priced at zero and we are only looking at buyers who want one item in each group) since the problem splits into a collection of single-item problems and so we can use equation (1). So, we can efficiently get at least OPT_2. Similarly, zeroing out prices in group 2 and setting prices to items in group 1 we can efficiently get at least OPT_1. Overall, we immediately get expected revenue of at least $OPT/4$ by taking the max of the two.

Note that we did not consider incentive compatibility here. One way to convert this into an incentive compatible setting is to split the buyers into two buyer-groups. Set the prices for buyer-group 1 based on valuations of buyer-group 2, and vice versa. Then act on their behalf, buying if the price is less, otherwise not. This is clearly incentive compatible, as they buy if they get positive utility, otherwise not, and they can't affect their own price. Given a large enough number of buyers, you can make a probabilistic argument saying that with high probability it will yield close to optimal revenue.