CSC2421: Online and other myopic algorithms Spring 2021

Allan Borodin

March 25, 2021

Week 10

Agenda for today and following two weeks

- For the last meeting, we had overview presentations of the reading projects by Kosha (min cost) and Xiaoxu (caching). As I said two weeks ago, . I am basing the grade on the quality of these presentations. The criteria are : Was the presentation clear and how informative was the presentation.
- Today Jinman will give her presentation on coloring. I will follow her presentation by continuing the discussion on prophet inequalities and prophet secretaries. Then time permitting. I will discuss ML advice.
- Next week (April 1)
  Alex (temporary jobs) and Koko (streaming) will give their presentations (streaming).

Recap of March 11 discussion

- We discussed three related problems: the secretary problem, prophet inequalities, and the prophet secretary problem.
- These all can be considered within the framework of online bipartite matching.
- The prophet secretary results are relatively new and based mainly on papers by Esfandiari et al [SICOMP 2017], Ehsani et al [SODA 2018] and Correa et al [Math Programming 2020].
- The secretary and prophet inequalities results are more "classic". The first published algorithm and proof for the secretary problem is attribued to Lindley (1961). The prophet inequalities bounds is due to Krengel and Sucheston (1977) whereas the threshold algorithm presented presented on March is due to Samuel-Cahn (1984) using the proof by Kleinberg and Weinberg (2012) as presented in Lucier (2017).

Summary of ressults for single item secretary, prophet inequalities, and prophet secretary

All of the results that follow are asymptotic ratios.

- For the secretary problem, the optimal competitive ratio is $\frac{1}{e} \approx .3679$
- For the prophet inequalities problem, the optimal competitive ratio is $\frac{1}{2}$ and this is achieved by a simple single threshold algorithm. This improvement over the seretary bound is made possible since now the input values are being drawn from known i.d. distributions even though the items are arriving in adversarial order vs random order in the secretary problem.
- Recall, it is not possible to get a constant competitive ratio for adversarial values and adversarial order even using randomized algorithms.

Summary of single item results continued

- For the prophet secretary problem, there is a single threshold algorithm (breaking ties using randomization) that achieves competitive ratio $1 - \frac{1}{e} \approx .6321$. Here we have the benefit of values from i.d. distributions and random order. This bound is optimal for non-adaptive thresholds.
  NOTE: I was following a conference version of the Ehsani et al paper. A later arXiv version presents this by introducing a time variable $t \in [0, 1]$ and then inducing a random order on the arriving buyers by choosing a random time T_i for each buyer i. The proof of the lemma and the theorem is in terms of $q(t)$ which is defined as the probability that the item has not been sold before time t. The corresponding useful lemma is that $q(t) \leq q_{-i}(t)$ defined as the probability that the item has not been sold before time t conditioned on $T_i = t$ and then showing $q(t) \geq e^{-t}$.

Summary continued

- Using adaptive thresholds, the $1 - \frac{1}{e} \approx .6321$ bound can be improved to .669 and this can be compared with the best known $\sqrt{3} - 1 \approx .732$ inapproximation. The current method (based on "blind strategies") has a .675 limitation.
- These results can all be improved for i.i.d. distributions in which case there is no difference between adversarial order and random order. The optimal bound is .745 for i.i.d. vs the best known .732 inapproximation for i.d.

Multi item extensions of the secretary problem, prophet inequalities and the prophet secretary problems

So far we have considered the secretary problem, prophet inequalities and prophet secretary problems in their basic form where only one item is being selected.

All of these problems have been extended to the settings where a set of items are being selected subject to various downward closed constraints (i.e., if S is a feasible set of items, then S' is feasible for any $S' \subseteq S$).

In this regard, matroid constraints, knapsack constraints and (bipartite) matching constraints are the most common constraints. We will focus on matching constraints.

The multi item secretary problem subject to a matching constraint

We first note that the Secretary problem is a special case of edge weighted online bipartite matching when there is only one offline vertex. For our first extension, it is natural to consider the edge weighted bipartite matching problem in the ROM model.

That is, upon arrival of an online vertex, we know the edge weights of all its adjacencies. We cannot do better than the $\frac{1}{e}$ competitive ratio, so the question is whether or not we can achieve this ratio.
The following theorem and algorithm due to Kesselheim et al [ESA 2013] achieves the optimal ROM bound.

Theorem

The online edge weighted bipartite Algorithm has expected approximation ratio $\frac{1}{e}$ in the ROM model. More spefically

$$\mathbb{E}[w(M)] \geq (\frac{1}{e} - \frac{1}{n}) \cdot OPT$$

The edge weighted bipartite matching algorithm

Algorithm 49 The EDGE-WEIGHTED BIPARTITE MATCHING ALGORITHM

```
procedure WEIGHTED MATCHING
    ▷ V is the set of offline vertices
    ▷ Online vertices u_1, ..., u_n arrive according to the ROM model
    U' ← {u_1, ..., u_{⌊n/e⌋}}                ▷ U' is the current set of online vertices
    M ← ∅                                      ▷ M will be the constructed matching
    ℓ ← ⌈n/e⌉
    while ℓ ≤ n do
        U' ← U' ∪ {u_ℓ}
        M^(ℓ) ← optimal matching on edge weighted graph with online vertices U' and offline vertices V
        if (ℓ, r) ∈ M^(ℓ) and r not yet matched then
            M ← M ∪ {(ℓ, r)}
        ℓ ← ℓ + 1
```

Note: The matching $M^{(\ell)}$ is in terms of the induced graph defined by the vertices in U' ignoring whatever matches have already been made.

Proof for the edge weighted bipartite matching algorithm

The proof of the Theorem establishing the $\frac{1}{e}$ competitive ratio relies on the following lemma for estimating the expected contribution of each online node u_ℓ for $\ell \geq \lceil n/e \rceil$. We'll prove this lemma after seeing how it yields the theorem.

Lemma

Let A_ℓ denote the contribution (i.e. the weight added to the solution) of online vertex $u_\ell \in U$ for $\lceil n/e \rceil \leq \ell \leq n$. Then $\mathbb{E}[A_\ell] \geq \frac{\lfloor n/e \rfloor - 1}{\ell - 1} \cdot \frac{OPT}{n}$.

Proof for the edge weighted bipartite matching algorithm continued

Using the Lemma, the proof of the Theorem follows easily.

Proof.

Theorem 1 is obtained by summing up the individual contributions $\mathbb{E}[A_\ell]$. That is,

$$\mathbb{E}[w(M)] = \mathbb{E}[\sum_{\ell=1}^{n} A(\ell) \geq \sum_{\ell=\lceil n/e \rceil}^{n} \frac{\lfloor n/e \rfloor - 1}{\ell - 1} \cdot \frac{OPT}{n}$$
$$= \frac{\lfloor n/e \rfloor - 1}{n} \sum_{\ell=\lfloor n/e \rfloor}^{n-1} \frac{OPT}{\ell}$$

This can be simplified to yield the desired bound on $\mathbb{E}[w(M)]$ by observing that

$\frac{\lfloor n/e \rfloor - 1}{n} \geq (\frac{1}{e} - \frac{1}{n})$ and $\sum_{\ell=\lfloor n/e \rfloor}^{n-1} \frac{1}{\ell} \geq \ln(\frac{n-1}{\lfloor n/e \rfloor} \geq 1$

□

Proof of the lemma needed for the edge weighted bipartite competitive ratio

It is helpful to view the random order of online vertices so that u_ℓ is chosen uniformly at random from U'. Then, conditioned on r being unmatched thus far,
(1) the expected weight $\mathbb{E}[(u_\ell, r)]$ of the edge (u_ℓ, r) is $\frac{w(M^{(\ell)})}{\ell}$ where $w(M^\ell)$ is the weight of the optimal matching on the current set of vertices.
Furthermore, U' is a uniformly at random set of size ℓ chosen from U so that
(2) $\mathbb{E}[w(M^{(\ell)}] \geq \frac{\ell}{n} OPT$. It follows that

$$\mathbb{E}[w(u_\ell, r)] \geq \frac{OPT}{n}$$

The expectation of the above inequality is in terms of the random choice of U' and the choice of ℓ as the last arrival in U'.

Proof of the lemma continued

We now need to consider the randomness in the preceding $\ell - 1$ arrivals to determine the probability that the intended match r for u_ℓ was not already matched.

Using the same view that the last element in any initial input sequence is being chosen randomly from the initial set of inputs and independent of the order of the previous elements, the probability that r is not chosen in the k^{th} iteration (for $k = \ell - 1, \ell - 2, \ldots 1$) is $\frac{k-1}{k}$ and hence Pr(r is unmatched) when u_ℓ arrives is equal to $\prod_{k=\lceil n/e \rceil}^{\ell-1} \frac{k-1}{k} = \frac{\lceil n/e \rceil - 1}{\ell - 1}$

Summarizing, the expected contribution of the ℓ^{th} online vertex is $\mathbb{E}[A_\ell] = \mathbb{E}[(u_\ell, r) | r \text{ is not yet matched}] \cdot \Pr(r \text{ is not yet matched})$. Namely $\mathbb{E}[A_\ell] \geq \frac{\lfloor n/e \rfloor - 1}{\ell - 1} \cdot \frac{OPT}{n}$ as claimed.

The bipartite matching secretary problem in terms of matroids

The weighted edge bipartite theorem is clearly selecting a set of winning candidates subject to a constraint, namely the constraint is that the set of selected online vertices can be matched (1-1) to the offline vertices. This simpler constraint constitutes the definition of independence in a *transversal matroid*.

Definition

Matroids Let U be a set of elements and $\mathcal{I}$ be a collection of subsets of U. $(U, \mathcal{I})$ is a matroid if the following hold:

- (Hereditary property) If $I \in \mathcal{I}$ and $I' \subset I$, then $I' \in \mathcal{I}$.
- (Exchange property) If $I', I \in \mathcal{I}$ and $|I'| < |I|$, then $\exists u \in I \setminus I'$ such that $I' \cup \{u\} \in \mathcal{I}$.

An *hereditary set system* $(U, \mathcal{I})$ is any set system satisfying the hereditary property so that a matroid is an hereditary set system that also satisfies the exchange property.

The sets $I \in \mathcal{I}$ are referred to as the *independent sets*.

Matroid secretary problem contiuned

We note that there are alternative equivalent definitions. In particular, an alternative to the exchange property is that every maximal independent set has the same size, and this maximum size is call the *rank* of the matroid.

Note: The bipartite matching constraint in terms of edges being a matching is the intersection of two matroids and is *not* a matroid constraint.

A special case of edge weighted matching is when the online vertices are weighted which is equivalent to saying that so all edges adjacent to an online node u have the same weight.

Let $(U, \mathcal{I})$ be a matroid whose elements u_i are weighted by $w : U \to \mathbb{R}^{\geq 0}$. The *matroid secretary problem* is to choose an independent set $I \subseteq U$ in the matroid so as to maximize $\sum_{u \in I} w_u$.

Matroid secretary problem continued

As an immediate consequence of the Kesselheim et al secretary matching result, we obtain a constant competitive ratio for three secretary problems , namely choosing a set of candidates so as to maximize the sum of the element weights subject to the following matroid independence constraints.

- A uniform matroid; that is, where the independent sets I have cardinality at most k for some fixed k. Here it is immediate to see that the rank of such a matroid is the cardinality constraint k.
- A partition matroid; that is, there is a partion $(U_1, \ldots, U_m)$ (for some m) and for each U_j there is a capacity k_j; the independent sets I are those satisfying $|u \in U_j \cap I| \leq k_j$. Clearly, every uniform matroid is a partition matroid.
- A transversal matroid. Every partition matroid is a transversal matroid.

The uniform matroid

Of course, since these particular matroid constraint problems are special cases of the bipartite matching problem it may be possible to obtain better constant approximation ratios. That is, multi item secretary problem might have a better ratio for some constraints.

Indeed this is the case for the uniform matroid with cardinality constraint k for which there is a $(1 - \frac{5}{\sqrt{k}})$ approximation. Hence for uniform matroids the approximation ratio limits to 1 as k increases.

It is an open problem whether or not there is a constant approximation for all matroid constraints. Currently, the best known approximation guarantee for an arbitrary matroid constraint is $\frac{1}{\Omega(\log \log k)}$ for matroids of rank k.

Knapsack secretary problem

Another noteworthy extension is the secretary problem subject to a knapsack constraint. This is the same thing as the knapsack problem in the ROM input model.) In this problem, candidates u_j have both a value v_j and a size s_j (sometimes called a size or a budget), and a size bound W.

The goal is to select online (in the ROM model) a set S of candidates satisfying the constraint $\sum_{j \in S} s_j \leq W$ so as to maximize $\sum_{j \in S} v_j$. Note that this constraint subsumes the uniform matroid constraint and is incomparable with an arbitrary matroid constraint and incomparable with the bipartite matching constraint. The best known competitive ratio for the knapsack secretary problem is $\frac{1}{6.65}$ achieved by a randomized algorithm.

The multi item prophet inequalities and prophet secretary problem subject to a matching constraint

As I probably mentioned, the matching constraint for both prophet inequalities and the prophet sectretary problem is of particular, interest due to the interpretation for posted price mechanisms for auctions. The matching constraint is a special case of a combinatorial auction, namely the valuations are *unit demand*.

More generally, in a posted price combinatorial auction (CA) buyers arrive online and take the best bundle of items (in terms of the agents valuation) that is still available. A unit demand CA is one where each agent only has value for single items. This is then clearly a matching problem. In the posted price unit demand CA, each buyer is offered a single price for each item. I self-interest buyer wll take the best item whose value (for the buyer) exceeds the price.

End of meeting on March 26

Next week we will have presentations by Alex and Koko. Our last meeting will be on April 8 and I will continue with the proof of the secretary matching problem and then discuss online algorithms with ML predictions.