Secretary & Prophet

Recall the secretary problem from last lecture.
We have n candidates for a secretary position and we assume there is a universal ordering of them, i.e. if we compare any two candidate i & j we can say which one is better for the job.
The candidates arrive in a u.r. dist. order; each time we interview we must decide to either hire (and terminate the process) or pass and continue.

Goal: Find a strategy to maximize the prob. of hiring the best candidate.

We proposed the following algo:
- pass for the first $\frac{n}{e}$ candidates
- Then hire the candidate that is the best seen so far.

Theorem: This algorithm finds/hires the best candidate with prob. $\frac{1}{e}$.

We showed that if we pass $\underline{k}$ and then hire the best we hire the best with prob. $(1-o(1))\frac{k}{n}\ln\left(\frac{n}{k}\right)$ and with $k \approx \frac{n}{e}$ this is maximized.

We can show that in this problem the best strategy is in fact a wait_and_pick strategy like what we described.

Theorem: The strategy that maximizes picking the best candidate is a wait-and-pick strategy

Proof: Consider some fixed optimal strategy.
let P_i : prob. the strategy picks agent at position i

q_i : prob. pick agent at position i conditioned on it being the best seen so far.

q_i : prob. pick agent at position i conditioned on it being the best seen so far.

So $q_i = \frac{P_i}{1/i} = i \cdot P_i$

$$\Pr[\text{Picking the best}] = \sum_i \Pr[i\text{'th agent is the best \& we pick it}]$$

$$= \sum_i \Pr[i\text{'th agent is the best}] \cdot q_i$$

$$= \sum_i \frac{1}{n} q_i = \sum_i \frac{i}{n} P_i$$

clearly $0 \le P_i \le 1$ and we have:

$$P_i = \Pr[\text{pick agent } i \mid i \text{ is the best so far}] \cdot \Pr[i \text{ is best so far}]$$

$$\le \Pr[\text{did not pick } 1 \dots i-1 \mid i \text{ is best so far}] \cdot \frac{1}{i}$$

But not picking $1 \dots (i-1)$ is independent of "i being the best so far"

So $P_i \le \frac{1}{i}\left(1 - \sum_{j<i} P_j\right)$

So we can upper-bound the success prob. of this (and any) strategy by the following LP:

$$\max \sum_i \frac{i}{n} \cdot P_i$$

$$i \cdot P_i \le 1 - \sum_{j<i} P_j$$

$$0 \le P_i \le 1$$

It can be verified that:

$$\begin{cases} P_i = 0 & \text{for } i < k \\ (k-1)\left(\frac{1}{i-1} - \frac{1}{i}\right) & k \le i \le n \end{cases}$$

is a feasible solution where k = smallest int s.t. $H_{n-1} - H_{k-1} \le 1$

It can be shown that the dual LP has a solution of the same value $\longrightarrow$ must be optimum.

of the same value $\longrightarrow$ must be optimum.

We have a wain-and-pick strategy whose value is the same as this LP solution:

For position i, if we have not picked any & this is the best so far, pick with prob. $\frac{i P_i}{1 - \sum_{j<i} P_j}$

then it can be shown that this is picking the best candidate with prob. $\sum_i \frac{i}{n} \cdot P_i$ which is the same as the LP.

Extension: Multiple agents

- Suppose we want to hire k agents (or pick k items) instead of just 1.

Goal: maximize the expected sum of the values of these

- Say the set of k largest items are $S^* \subseteq [n]$ and their total values is $V^* = \sum_{i \in S^*} a_i$

First approach: Suppose we pass on the first $\frac{n}{2}$ agents. Note that we expect half $(\sim \frac{k}{2})$ of the largest be in the first half and half $(\sim \frac{k}{2})$ be in the second half.

Suppose we use the value $\tilde{a}$ of the $\frac{(1-\varepsilon)k}{2}$ th largest element in the first half as the threshold and pick up to k items from the second half if their values are $\geq \tilde{a}$. We can formalize this.

let $\delta = O(\frac{\log k}{k^{1/3}})$ and $\varepsilon = \delta/2$. Our goal is an algorithm let $\delta = O\left(\frac{\log k}{k^{1/3}}\right)$ and $\varepsilon = \delta/2$. Our goal is an algorithm that gives value $V^*(1-\delta)$ (Note $\delta \to 0$ when $k \to \infty$)

- pass the first δn items
  (we expect $\approx k^{2/3}/\log k$ items from S^* in this set)
- let $\tilde{a}$ be the value of $(1-\varepsilon)\delta k$'th highest among the "passed" items; pick the first k elements in the remaining $(1-\delta)n$ with value $\geq \tilde{a}$

Bad events: E1: if $a' = \min_{i \in S^*} a_i$ is the lowest value of S^* and $\tilde{a} < a'$ (we pick items with lower value)

E2: # of items from S^* in the last $(1-\delta)n$ and greater than $\tilde{a}$ is much smaller than k

Why these bad events happen rarely:

For E1, less than $(1-\varepsilon)\delta k$ items from S^* fall in the first δn locations; this requires the # of them is smaller than $(1-\varepsilon) \times$ expectation.

Using Chernoff-Hoeffding this prob $\leq e^{-\varepsilon^2 \delta k} \approx \frac{1}{\text{poly}(k)}$

For E2 to happen, more than $(1-\varepsilon)\delta k$ of the top $(1-\delta)k$ from S^* must be in the first δn elements.

This means the # of them exceeds $(1+O(\varepsilon))$ times the expectation, again has probability $\leq e^{-\varepsilon^2 \delta k} = \frac{1}{\text{poly}(k)}$

Best possible: Kleinberg'14 showed that we can get

Best possible: Kleinberg '14 showed that we can get $1-O(\sqrt{1/k})$ competitive ratio for the k-Secretary & this is best possible.

The Prophet inequality

- The setting and the goal of this is slightly different
- We have a distribution for the value of each item but don't know the actual value until the candidate is sampled (interviewed).

The arrival of candidates/items is not random but adversarial!

Formally: An ordered set of random variables $X_1 \dots X_n$, with known distribution.

At each step i we sample $x_i \sim X_i$ and get its value: can decide to accept (terminate) or pass (continue).

Goal: maximize the expected value of chosen item compared to $\max_i X_i$

i.e. find an algorithm A s.t. $\mathbb{E}[A] \geq \alpha \cdot \mathbb{E}[\max_i X_i]$ while maximizing α.

First we show we cannot have $\alpha > \frac{1}{2}$

Suppose X_1 is always 1 while $X_2 = 0$ with prob $1-\varepsilon$ and $X_2 = \frac{1}{\varepsilon}$ with prob. ε. Then

$$\mathbb{E}[\max\{X_1, X_2\}] = (1-\varepsilon) \cdot 1 + \varepsilon \cdot \frac{1}{\varepsilon} = 2-\varepsilon$$

Any strategy either picks X_1 (which has always value 1) or picks X_2 and in either case $\mathbb{E}[\text{algorithm}] = 1$

So we cannot have $\alpha > \frac{1}{2}$, but surprisingly we can get $\alpha = \frac{1}{2}$

Theorem (Krengel, Sucheston, Garling) There is a strategy with $\alpha = \frac{1}{2}$

proof: The idea is similar to the secretary problem, however because the arival is not random (could be adversarial) we cannot ignore a fixed collection of values.

Define $X_{max} = \max_i X_i$ and let $\tau = \frac{1}{2}\mathbb{E}[X_{max}]$

The algorithm is: Select the first X_i s.t. $\geq \tau$

i.e. the first time we sample a value $\geq \tau$.

For any random var Y let $Y^+ = \max\{Y, 0\}$

$$\begin{aligned}
\mathbb{E}[X_{max}] &= \mathbb{E}[\tau + X_{max} - \tau] \\
&= \tau + \mathbb{E}[(X_{max} - \tau)^+] \qquad \text{Since } (X_{max} - \tau)^+ \geq X_{max} - \tau \\
&\leq \tau + \mathbb{E}\Big[\sum_i (X_i - \tau)^+\Big] \\
&= \tau + \sum_i \mathbb{E}[(X_i - \tau)^+] \qquad (*)
\end{aligned}$$

let q be the prob. that the alg. returns some value

$$q = \Pr[\exists i : X_i \geq \tau] = 1 - \Pr[\forall i : X_i < \tau]$$

$$\mathbb{E}[\text{Alg}] = \sum_i \underbrace{\Pr\Big[\bigwedge_{j<i} (X_j < \tau) \wedge (X_i \geq \tau)\Big]}_{\text{Alg} = X_i} \cdot \mathbb{E}\Big[X_i \,\Big|\, \bigwedge_{j<i} (X_j < \tau) \wedge (X_i \geq \tau)\Big]$$

since x_i is ind

$$= \sum_i \Pr\Big[\bigwedge_{j<i} (X_i < \tau)\Big] \cdot \underbrace{\Pr[X_i \geq \tau] \cdot \mathbb{E}[\tau + (X_i - \tau)^+ \mid X_i \geq \tau]}_{X_i \text{ ind of } X_j}$$

$$= \sum_i \Pr\Big[\bigwedge_{j<i} (X_i < \tau)\Big] \cdot \underbrace{\Pr[X_i \geq \tau] \cdot \mathbb{E}[\tau + (X_i - \tau) \mid X_i \geq \tau]}_{\tau + \mathbb{E}[(X_i - \tau)^+]}$$

X_i ind of X_j

Note that $q = \sum_i \Pr\Big[\bigwedge_{j<i} (X_j < \tau)\Big] \cdot \Pr[X_i \geq \tau]$

$$\mathbb{E}[\text{Alg}] = q \cdot \tau + \sum_i \Pr\Big[\bigwedge_{j<i} (X_j < \tau)\Big] \cdot \mathbb{E}[(X_i - \tau)^+]$$

$$\geq q \cdot \tau + \sum_i (1-q)\, \mathbb{E}[(X_i - \tau)^+]$$

$$\geq q \cdot \tau + (1-q)\big(\mathbb{E}[X_{max}] - \tau\big) \qquad \text{using } (*)$$

$$= \tau$$

∎

LP-based Generalization

The following LP-based approach works for more general settings of prophet inequalities.

Suppose X_i comes from a known dist. and say it will have value $v_i \geq 0$ with prob. P_i and is zero o.w. (for simplicity we are focusing on disc. version but the same argument works in general; in general we can have: $\Pr[X_i \geq v_i] = P_i$).

Consider the following LP:

$$\max \sum_i x_i \cdot v_i$$

$$\forall i \qquad x_i \leq P_i$$

$$\sum_i x_i \leq 1$$

$$\sum_t x_i \leq 1$$
$$x_i \geq 0$$

let x^* be the opt LP solution

Lemma: $\sum_i x_i^* \cdot v_i \geq \mathbb{E}[X_{max}]$

proof: we give a feasible sol with obj. value $\mathbb{E}[X_{max}]$

let q_i be the prob. that X_i is non-zero and is X_{max}

Note $q_i \leq p_i$ and $\sum_i q_i \leq 1$.

Also $\sum_i q_i \cdot v_i = \mathbb{E}[X_{max}]$. ∎

Using this lemma we can design a $\frac{1}{4}$-competitive alg.

Idea: the LP sol x^* suggests how frequenty we should accept X_i (x_i^*).

if we had no constraints on how many X_i's we can accept we could say: accept X_i when it has value v_i with prob. $\frac{x_i^*}{p_i}$, unless we have accepted something

This could have poor performance:

e.g. $X_1 = 1$ w.p. 1 $\quad X_2 = \begin{cases} \frac{1}{\varepsilon^2} & \text{w.p. } \varepsilon \\ 0 & \text{o.w.} \end{cases}$

$x^* = (1-\varepsilon, \varepsilon)$ but an alg. that always accepts X_1 w.p. $1-\varepsilon$ reaches X_2 without accepting anything at most ε-frac. of time. So

$\Pr[\text{reach } X_2] \cdot \Pr[X_2 = 1/\varepsilon^2] = \varepsilon^2 \longrightarrow$

Expected value is at most $(1-\varepsilon) + \varepsilon^2 \cdot \frac{1}{\varepsilon^2} = 2 - \varepsilon$ but

$\mathbb{E}[X_{max}] = 1/\varepsilon$

What if we ignore each element we reach with prob. $\frac{1}{2}$? i.e. we don't even look at it!

Alg: let x^* be LP solution.

when looking at item i (assuming nothing is selected)

ignore with prob $\frac{1}{2}$ (w.o. even looking),

else accept with prob x_i

Lemma: The alg achieves a value $\frac{1}{4}\mathbb{E}[X_{max}]$

Proof: we say we reach item i if we have not picked any item before. Note that we pick each i with prob at most $\frac{x_i^*}{2}$ ($\frac{1}{2}$ ignore & then pick w.p. x_i^* if we reach i)

$$\mathbb{E}[Alg] \geq \sum_{i=1}^{n} \Pr[\text{reach } i] \cdot \frac{1}{2} \cdot x_i^* \cdot v_i$$

$$\geq \sum_{i=1}^{n} \Pr\left[\bigwedge_{j<i} \text{Not picked } j\right] \cdot \frac{x_i^* \cdot v_i}{2}$$

$$\geq \sum_{i=1}^{n} \left(1 - \sum_{j<i} x_j^*/2\right) \frac{x_i^* \cdot v_i}{2}$$

$$\geq \sum_{i=1}^{n} \frac{1}{2} \cdot \frac{x_i^* v_i}{2} \qquad \text{Since } \sum_j \frac{x_j^*}{2} \leq 1/2$$

$$\geq \frac{1}{4}\mathbb{E}[X_{max}]$$