Online Matching with unknown iid arrivals

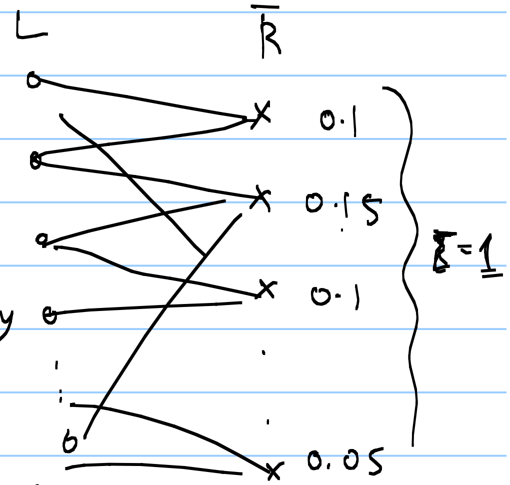
Note Title 1/29/2013

- Vertices in R are sampled (independently & identically) from a distribution.

- Given the offline vertex set L.

- A probability distribution from which vertices in R are sampled from. Say the support of this distribution is $\bar{R}$. Each element $\bar{j} \in \bar{R}$ is identified by its neighbors in L. The probability of $\bar{j}$ is $p(\bar{j})$

Distribution graph.



- Each vertex $j \in R$ is sampled independently from $p()$.

- Both ALG & OPT are r.v.s.
  $\therefore$ We'd like $E[ALG] \geq \gamma \cdot E[OPT]$.

Defn. $\overline{OPT}$ = optimum value of the "expected instance" deterministic value, and $\overline{OPT} \geq E[OPT]$.
$\therefore$ Sufficient to prove $E[ALG] \geq \gamma \cdot \overline{OPT}$.

m = # vertices in R.

$\overline{OPT} = \max \sum_{i,j} x_{ij}$ st. — expected LP

$\forall\, i \in L \quad \sum_{j \in \bar{R}} x_{ij} \le 1.$

$\forall\, j \in \bar{R} \quad \sum_{i \in L} x_{ij} \le p(j) \cdot m$ = expected # of times j appears in R

$x_{ij} \geq 0.$

Lemma:- $\overline{OPT} \geq E[OPT]$.

Proof:- Let $X_{ij}^* =$ # of times j is matched to i in OPT, a r.v.

& $x_{ij}^* = E[X_{ij}^*]$

$\sum_i X_{ij}^* \le$ # times j appears in R

$\therefore E[\sum_i X_{ij}^*] \le E[\text{———}''\text{———}] = p(j) \cdot m$

$\therefore \sum_i x_{ij}^* \le p(j) \cdot m$

$\text{III}^{ly} \quad \sum_j x_{ij}^* \le 1 \quad \forall i$

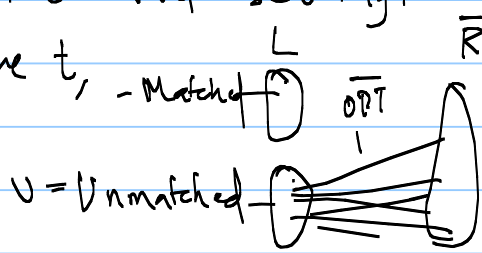
& $E[OPT] = \sum_i x_{ij}^*$

$\left.\begin{matrix} \\ \\ \\ \end{matrix}\right\}$ x_{ij}^* is a feasible soln. to the expected LP.

$\therefore E[OPT] \le \overline{OPT}$ $\square$.

Theorem:- The greedy algorithm is $1 - 1/e$ competitive in the iid setting.

Proof:- At time t,



$\Pr[j$ has a neighbor in $U]$

Use only $\overline{OPT}$

$$\overline{OPT} = \sum_{i,j} x_{ij} = \sum_{i\in U, j} x_{ij} + \underbrace{\sum_{i \in L\setminus U, j} x_{ij}}_{= ALG(t)}$$

$$\therefore \overline{OPT} - ALG(t) = \sum_{i\in U, j} x_{ij}$$

j is chosen with probability $p(j)$.

Suppose we Match j to i with prob. $\frac{x_{ij}}{m \cdot p(j)}$

Then $Pr[i \text{ gets matched}] = \sum_j \cancel{p(j)} \cdot \frac{x_{ij}}{m\,\cancel{p(j)}}$

$$= \sum_j x_{ij}/m.$$

$$\therefore Pr[\text{Some } i \in U \text{ gets matched}] = \frac{\sum_{j, i\in U} x_{ij}}{m}$$

$$\therefore Pr \underset{\text{in step } t}{ALG \text{ finds a match}} \geq \frac{\overline{OPT} - ALG(t)}{m}$$

$$\therefore E[ALG(t+1) \mid ALG(t)] \geq ALG(t) + \frac{\overline{OPT} - ALG(t)}{m}$$

$$\therefore E[\overline{OPT} - ALG(t+1)] \leq (\overline{OPT} - ALG(t))\left(1 - \frac{1}{m}\right)$$

$$\therefore E[\overline{OPT} - ALG(m)] \leq \overline{OPT}\left(1 - \frac{1}{m}\right)^m \leq \frac{\overline{OPT}}{e}$$

$$\therefore E[ALG] \geq \overline{OPT}\left(1 - \frac{1}{e}\right)$$

Suggested Exercise:

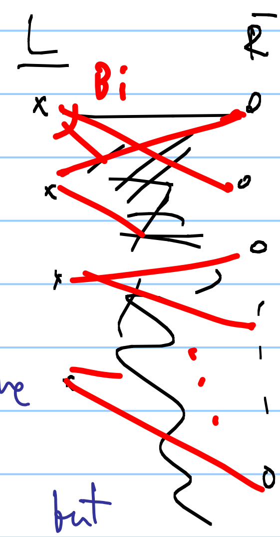
- Generalize to (Integral) Budgeted Allocation without the assumption $b_{ij} \ll B_i$.

B-matching: each $i \in L$ can be matched B_i times. All $B_i \geq k$.

Assume:- $|\bar{R}| = m$, & $p(j) = \frac{1}{m}$ $\forall$ $j \in \bar{R}$.

$\therefore$ $p(j) \cdot m = 1$. The expected instance is simply an integral matching problem.
Further, suppose $\exists$ a perfect matching in the expected instance. i.e each j is matched to $M(j)$ & each $i \in L$ is matched B_i times.

$\therefore$ $OPT = m = \sum_i B_i$.



Pure-Random Algorithm:

- Knows expected instance, G, & the matching.
- Is non-adaptive, makes all the decisions ahead of time.
- Always matches j to $M(j)$, but gets credit only if $\leq B_i$ matches

$\forall$ steps, $\Pr[i \text{ is matched in } 1 \text{ step}] = \frac{B_i}{m} = \frac{\# j\text{'s matched to } i}{\text{total } \# \text{ of } j\text{'s}}$

$\because$ uniform distribution

$\equiv$ Balls and bins procedure. **Independent of the Graph!**

- Each $i \equiv$ bin with capacity B_i. $\sum_i B_i = m$.
- In each round, throw a ball in bin i with prob. B_i/m.
- Repeat m times.

Q:- How many balls are in bin i? (X_i)

A:- There are ℓ balls with prob.

$$\binom{m}{\ell}\left(\frac{B_i}{m}\right)^{\ell}\left(1-\frac{B_i}{m}\right)^{m-\ell}$$

Want:- $E[\min\{X_i, B_i\}] = \sum_{\ell=1}^{B_i} \binom{m}{\ell}\left(\frac{B_i}{m}\right)^{\ell}\left(1-\frac{B_i}{m}\right)^{m-\ell} \cdot \ell$

$$+ B_i \sum_{\ell=B_i+1}^{m} \binom{m}{\ell}\left(\frac{B_i}{m}\right)^{\ell}\left(1-\frac{B_i}{m}\right)^{m-\ell}$$

- monotonically decreasing in m
- as $m \to \infty$, $E[\min\{X_i, B_i\}] \to B_i - \sqrt{\frac{B_i}{2\pi}}$.

$\therefore E[PR] = \sum_i B_i \left(1 - \sqrt{\frac{1}{2\pi B_i}}\right)$

$$\geq \sum_i B_i \left(1 - \frac{1}{\sqrt{2\pi k}}\right) \quad \text{Since } B_i \geq k \ \forall i$$

$$= OPT\left(1 - \frac{1}{\sqrt{2\pi k}}\right) \geq OPT(1-\varepsilon)$$

if $k \geq \frac{1}{2\pi\varepsilon^2} \iff \varepsilon \geq \frac{1}{\sqrt{2\pi k}}$

(Compare to $B_i \geq c\,\frac{n\log(mn)}{\varepsilon^2}$
 i .max

But we started out with unknown distribution!

Algorithm for unknown distribution: Define inductively

Say we've already matched $t-1$ vertices, denoted by

$$\mathcal{H}^t := A_1\, A_2, \dots\dots A_{t-1} \;?\; P_{t+1}\, P_{t+2} \dots\dots P_m.$$

Suppose after this step, we could magically run the pure-Random algorithm. Let $\mathcal{H}^t$ be the "Hybrid Algorithm".

Given j in the t^{th} step, $\forall$ choices of $i \sim j$, i unmatched, evaluate the expected # of matches in the remaining time for $\mathcal{H}^t$.

Match j to i that maximizes this. i.e

Match j to $\displaystyle \operatorname*{arg\,max}_{i:\, i\sim j} \left\{ E\left[\mathcal{H}^t \mid A_t = i\right] \right\}$

This defines A_t, & hence the algorithm.

Compare:

$$\begin{array}{ll} \mathcal{H}^{t-1} & A_1\, A_2, \dots A_{t-1} \boxed{P_t} P_{t+1}, \dots, P_m \\ \mathcal{H}^t & A_1\, A_2, \dots A_{t-1} \boxed{A_t} P_{t+1}, \dots, P_m \end{array}$$

$\to$ only difference.

Claim:- $E\left[\mathcal{H}^t\right] \geq E\left[\mathcal{H}^{t-1}\right]$, almost by definition.

$\therefore \quad E[\mathcal{H}^m] \geq E[\mathcal{H}^{m-1}] \geq \cdots \cdots \geq E[\mathcal{H}^0]$

$\quad\quad \| \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \|$

$E[ALG] \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad E[PR]$

$\therefore$ ALG is atleast as good as PR ! D

How can we perform the "magical step"?

We can estimate the expected # of matches $\because$ it is $\equiv$ some balls & bins procedure.

Given remaining capacity, prob. of match in 1 step & # of steps, can calculate exp. # of matches.

X

Suggested Exercise:

- Generalize to Budgeted Allocation. (Integral).

Sketch: - X_i = sum of b_{ij}'s. $E[X_i] = B_i$. $E[X_i^t] = \frac{B_i}{m}$

$E[\min\{X_i, B_i\}]$ is smallest when $b_{ij} \in \{0, b_i^{max}\}$

Algo:- evaluate profit in this step + expected remaining profit.

estimate assuming

$ALG \geq OPT\left(1 - \frac{1}{\sqrt{2\pi k}}\right)$ where $\frac{B_i}{b_i^{max}} \geq k$.

- Observe everything works for non-uniform distb'n.