Bipartite Matching

Given bipartite graph (L, R, E)
a matching $M \subseteq E$ is a set of edges that share no common endpts

Goal: find matching of maximum size

Announcements
Project preproposal due Feb 15
Pset 3 partially available, due Feb 22

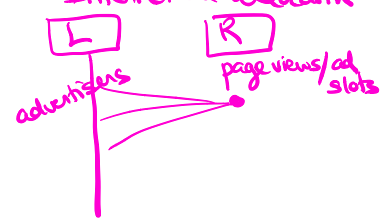
Today
online bipartite matching

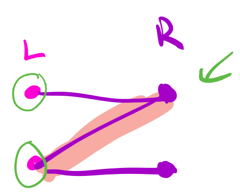
Online bipartite matching

online c.r. of alg A: $\min_{\text{instances } I} \dfrac{\sum_{(i,j) \in E} \hat{x}_{ij}}{OPT(I)}$

L is known ahead of time
vertices in R arrive one at a time
when $j \in R$ arrives
learn which vertices in L are neighbors of j
make an irrevocable decision as to which neighbor of any to match i to

Killer app:
Internet ad allocation





Are you familiar with LP duality
a) yes
b) somewhat
c) no

Greedy Alg

match arriving node to any available neighbor

Claim: Greedy always obtains a matching of size $\geq \frac{1}{2}$ OPT

Analysis: every time Greedy adds an edge to matching
think of it as earning \$1
Place 50¢ on each endpt

Now consider any edge (i,j) matched by OPT (when j arrives)
If Greedy can't match j, then i is already matched.
"charge" OPT's edge to 50¢ on i

$\Rightarrow$ $\forall$ \$1 OPT earns, Greedy earns $\geq$ 50¢

the competitive ratio of Greedy $\left[:= \max_{I} \dfrac{\text{size of Greedy matching on } I}{\text{size of OPT matching on } I} \right]$
$= \frac{1}{2}$

Instance
Graph
and ordering of arrival for $j \in R$

And no deterministic alg can do better.
So we turn to randomization

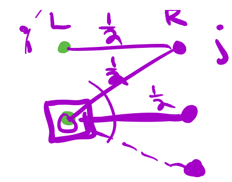
Fractional matching problem

When a node $j \in R$ arrives, can allocate it fractionally

x_{ij} fraction of j matched to i

constraints $\sum_{j \in R} x_{ij} \le 1$ total amt of i matched ≤ 1

$\sum_{i \in L} x_{ij} \le 1$ total amt of j matched ≤ 1



$\frac{3}{4}$ of OPT

Claim: Let A be a randomized alg for integral matching

$\Rightarrow$ $\exists$ deterministic fractional alg D s.t.

$\forall$ instance I

$$\sum_{i,j} x_{ij}^{D}(I) = E\Big(\sum_{i,j} \underbrace{X_{ij}^{A}(I)}\Big)$$

indicator of event (i,j) matched by alg A on instance I

Proof: D "simulates A"

when j arrives, set $x_{ij}^{D} = \Pr(X_{ij}^{A} = 1)$

$\forall i \quad \sum_{j \in N(i)} X_{ij}^{A} \le 1 \quad \Rightarrow E\Big(\sum_{j \in N(i)} X_{ij}^{A}\Big) \le 1$

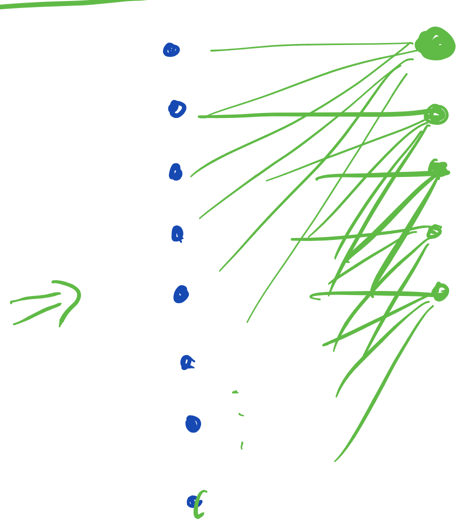
$\Rightarrow \sum_{j \in N(i)} x_{ij}^{D} \le 1$

$$\Rightarrow \sum_{(i,j) \in E} x_{ij}^{D} = E\Big(\sum_{(i,j) \in E} X_{ij}^{A}\Big) = E(\text{performance of } A)$$

$\Rightarrow$ upper bound on c.r. of deterministic fractional alg

$\Rightarrow$ upper bound on c.r. of randomized alg

det fractional



OPT max matching n.

i^{th} vertex on left.

$$\frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \dots \frac{1}{i}$$

$$= H_n - H_i = \ln\left(\frac{n}{i}\right)$$

$$i = \frac{n}{e} \qquad \ln\left(\frac{n}{i}\right) = 1$$

$n - i$

size of online mtchg $n(1 - \frac{1}{e})$.

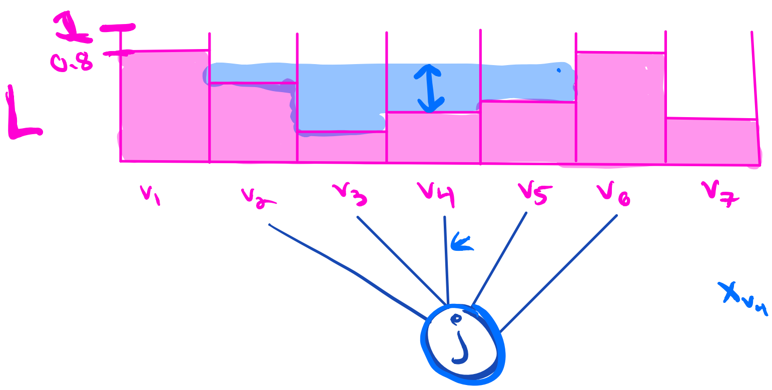
$\simeq 0.63\,n$

Water-level alg

Imagine vertices on LHS are water containers, with capacity 1

RHS - each vertex is source of 1 unit of water

When $j \in R$ arrives "fill up" neighbors in obvious way



pour 1 unit of water in

$x_{v_4,j}$

Primal dual analysis

primal LP

$$\max \sum_{(i,j)\in E} x_{ij}$$

$$\left(\sum_{j\in N(i)} x_{ij} \le 1\right)\cdot \alpha_i \quad i\in L$$

$$\left(\sum_{i\in N(j)} x_{ij} \le 1\right)\cdot \beta_j \quad j\in R$$

$$x_{ij} \ge 0$$

dual LP

$$\min \sum_i \alpha_i + \sum_j \beta_j$$

$$\boxed{\alpha_i + \beta_j \ge 1 \quad \forall (i,j)\in E}$$

$$\alpha_i \ge 0,\ \beta_j \ge 0$$

α_i $i\in L$, β_j $j\in R$

$$\sum_{i\in L}\alpha_i + \sum_{j\in R}\beta_j \geq \sum_i \alpha_i \sum_{j\in N(i)} x_{ij} + \sum_j \beta_j \sum_{i\in N(j)} x_{ij}$$

$$= \sum_{(i,j)\in E} (\alpha_i+\beta_j)\, x_{ij} \geq \sum_{(i,j)\in E} x_{ij}$$

$\forall$ feasible α_i's β_j's for "dual LP" $\quad \sum \alpha_i + \sum \beta_j \geq \sum_{(i,j)\in E} x_{ij} \quad \forall$ feasible x_{ij}'s

$$\underbrace{\sum \alpha_i + \sum \beta_j}_{\text{feasible}} \geq OPT$$

weak duality.

$$\underline{P \text{ feasible obj}} \le \underline{D \text{ feasible obj}}$$

$OPT(P) = OPT(D)$ aside.

Primal dual analysis of online matching problem

as vertices in R arrive, we will maintain $\underbrace{\text{primal "soln"}\ x_{ij}\text{'s}}_{\text{always be feasible}}$, dual soln" α_i, β_j's

update both solns as each $j\in R$ arrives

when j arrives $\forall i$ s.t. $\underline{x_{ij}} \uparrow$ by Δ

increase $\alpha_i + \beta_j$ by Δ also

$\Rightarrow$ $P_{\text{value}} = \underbrace{D_{\text{value}}}_{\text{if this was}}$ feasible at end

we would have proved optimality of primal.

If D is not feasible

show $\exists$ const $c > 0$

s.t. $\boxed{\alpha_i + \beta_j \geq c}$ $\forall (i,j) \in E$

$\alpha_i' = \frac{\alpha_i}{c}$ $\beta_j' = \frac{\beta_j}{c}$ $\forall i, j$

$\alpha_i' + \beta_j' \geq 1$ $\forall (i,j) \in E$

$\frac{\sum_{(i,j)} x_{ij}}{c} = \frac{D^{final}}{c}$ $\quad \boxed{\frac{P^{final}}{c}} = \boxed{\frac{D^{final}}{c}} \geq \underline{OPT} \Rightarrow P^{final} \geq c\, OPT$

feasible. c.r. $\geq c$

---

perform updates as $j \in R$ arrives so that. $\alpha_i + \beta_j \geq \underline{c}$

for c as large as possible

define fn $g: [0,1] \to [0,1]$ increasing fn.

to be determined

when $\boxed{x_{ij} \uparrow \text{ by } dx}$ $\boxed{\begin{array}{l} \uparrow \alpha_i \text{ by } g(y_i)dx \\ \uparrow \beta_j \text{ by } (1-g(y_i))dx \end{array}}$

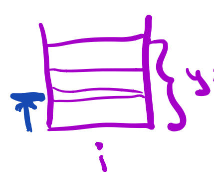
How to choose $g(\ )$?

$y_i := \sum_{j \in N(i)} x_{ij}$

Case 1: $y_i^f = 1$ (i is fully matched $\sum_{j \in N(i)} x_{ij} = 1$)

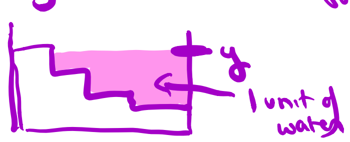
$\alpha_i = \int_0^1 g(y)dy = G(1) - G(0) \geq c$

$G'(y) = g(y)$



Case 2 $y_i^f < 1$ & suppose j is a neighbor of i

$\Rightarrow j$ gets fully matched $\sum_{i \in N(j)} x_{ij} = 1$



neighbors of j
(i is one of them)

$y_i \geq y$

$\alpha_i \geq \int_0^y g(z)dz = G(y) - G(0)$

$\beta_j = \int_{\tilde{y} \text{ in pink region}} (1 - g(\tilde{y}))\, d\tilde{y} \geq \underline{1 - g(y)}$

$1 - g(\tilde{y}) \geq 1 - g(y)$

— $y_{final, i}$

$$\beta_j = \sum_{i \in N(j)} \underbrace{\int_{y_{st,i}} (1-g(z))\,dz}_{\geq (y_{end,i} - y_{st,i})(1-g(y))}$$

$y_i^f < 1$

$\alpha_i + \beta_j$

$\geq G(y) - G(0) + 1 - g(y) \geq c \quad \forall y$

$\underline{y_i^f = 1}$ $\quad \alpha_i \geq G(1) - G(0) \geq c$

choose g w/ c as large as possible.

$$G(y) - G(0) + 1 - g(y) = c \quad \forall y$$

$$g(y) - g'(y) = 0$$

$$g(y) = g'(y)$$

$$\boxed{g(y) = \boxed{ke^y}}$$

$$\underbrace{\int_0^y ke^y\,dy}_{ke^y} - k + 1 - ke^y = \underbrace{k(e-1)}_{G(1)-G(0)}$$

$$1 - k = k(e-1)$$

$$\Rightarrow k = \boxed{\frac{1}{e}}$$

$1 - \frac{1}{e}$

$g(y) = e^{y-1}$

$$\boxed{c = 1 - \frac{1}{e}}$$

$$\alpha_i + \beta_j \geq 1 - \frac{1}{e} \quad \forall (i,j) \in E$$

$$\Rightarrow \quad \boxed{\frac{\alpha_i}{c}} + \boxed{\frac{\beta_j}{c}} \geq 1$$

$$\frac{P}{c} = \boxed{\frac{D}{c}} \geq OPT \Rightarrow P \geq c\,OPT$$

weak duality

$C.R \geq 1 - \frac{1}{e}$

Summary of what we did: "online" primal-dual analysis

Primal

$$\max \sum_{(i,j)\in E} x_{ij}$$

$$\sum_{j\in N(i)} x_{ij} \le 1 \quad \forall i$$

$$\sum_{i\in N(j)} x_{ij} \le 1 \quad \forall j$$

$$x_{ij} \ge 0 \quad \forall (i,j)\in E$$

Dual

$$\min \sum_i \alpha_i + \sum_j \beta_j$$

$$\alpha_i + \beta_j \ge 1 \quad \forall (i,j)\in E$$

$$\alpha_i \ge 0 \quad \forall i$$

$$\beta_j \ge 0 \quad \forall j$$

Weak duality $\Rightarrow$ feasible primal $\le$ $\boxed{\text{OPT}}$ $\le$ feasible dual

$$\sum_{(i,j)\in E} x_{ij} \le \boxed{\sum_{(i,j)\in E} x^*_{ij}} \le \sum_i \alpha_i + \sum_j \beta_j$$

As vertices arrived, we constructed near-feasible dual by ensuring that $\forall (i,j)\in E \quad \alpha_i + \beta_j \ge 1-\frac{1}{e}$

$$\Rightarrow \quad \alpha_i' = \frac{\alpha_i}{1-\frac{1}{e}} \ \forall i \qquad \beta_j' = \frac{\beta_j}{1-\frac{1}{e}} \ \forall j$$

is feasible dual

$$\Rightarrow \text{OPT} \le \sum_i \hat{\alpha}_i + \sum_j \hat{\beta}_j = \frac{1}{\left(1-\frac{1}{e}\right)}\left[\sum_i \alpha_i + \sum_j \beta_j\right]$$

$$= \frac{1}{\left(1-\frac{1}{e}\right)} \sum_{(i,j)\in E} x_{ij}$$

$$\Rightarrow \quad \sum_{(i,j)\in E} x_{ij} \ge \left(1-\frac{1}{e}\right)\text{OPT}$$

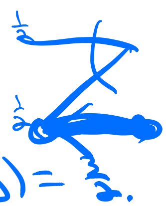
how to go from <u>fractional online algo</u> $\rightarrow$ <u>randomized</u> online algo.

Integral Online Matching

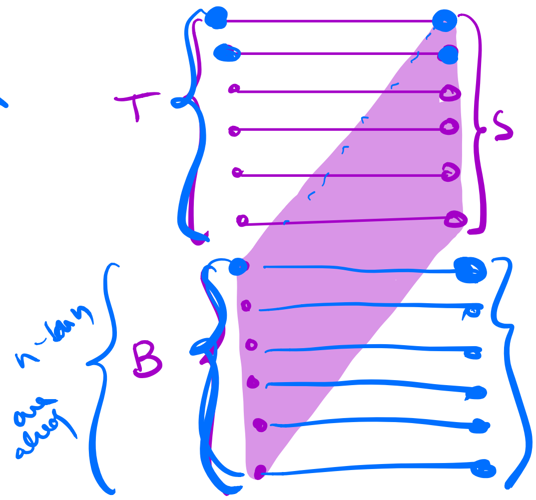
$|T| = |B| = n$.

when $j \in R$ arrives

match to a random (unif) free neighbor.



$E(\text{matching}) = \frac{3}{2}$.



arrival order

$\frac{1}{n+1} + \frac{1}{n} + \frac{1}{n-1} \ldots$

$\ln n$

$\ln n$

$\frac{n + \ln n}{2n}$

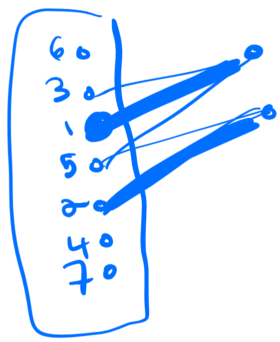
Ranking Algorithm

- Select random total order π of elts of L
- when new vertex in R arrives, match it to lowest ranked neighbor according to π

[Karp, Vazirani, Vazirani] 1990

series of papers simplifying

[Devanur, Jain, Kleinberg] 2013

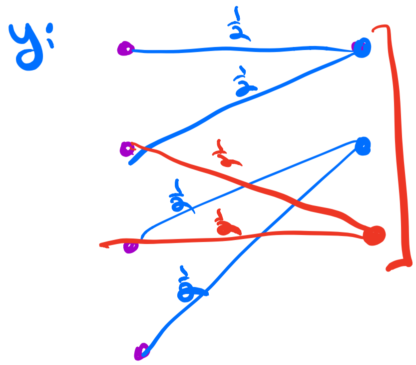


First idea: online fractional $1 - \frac{1}{e}$ competitive.

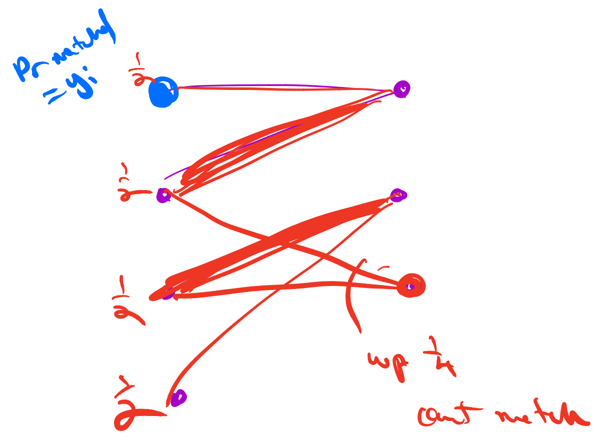
every fractional matching is convex comb of integer soln s.

Can we maintain integral matching whose expected size = value of fractional matching produced by Water Level?

Fractional



Randomized



Reintrep alg.

$\forall i \in L$, pick $Y_i \sim U[0,1]$ independently

when j arrives, match it to $\arg\min \{ Y_i \mid i \in N(j),\ i \text{ unmatched} \}$

feasible primal

and as we go, build (not necessary feasible) dual

$D = \{\alpha_i, \beta_j \ldots\}$ random depend on Y_i's.

s.t. (1) $P \geq D$

(2) $E\left(\frac{D}{c}\right)$ feasible. $\quad E_{\vec{Y}}\left[\frac{\sum_i \alpha_i(\vec{Y}) + \sum_j \beta_j(\vec{Y})}{c}\right]$

$P \geq D \Rightarrow \frac{E(P)}{c} \geq \frac{E(D)}{c} \geq OPT$

$c = 1 - \frac{1}{e}$.

$\frac{D}{c}$

$g : [0,1] \to [0,1]$

when $x_{ij} = 1$ set $\alpha_i = g(Y_i)$ $\beta_j = 1 - g(Y_i)$

if j unmatched $\beta_j = 0$

i unmatched $\alpha_i = 0$

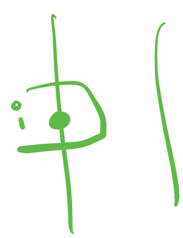
$\Delta P = \Delta D$

need to show $\boxed{\frac{E(D)}{c}}$ feasible.

$\alpha_i' = \frac{E(\alpha_i)}{c}$ $\beta_j' = \frac{E(\beta_j)}{c}$

Fix edge (i, j) Fix Y_{-i} $(Y_1, Y_2, .. Y_{i-1}, Y_{i+1} .. Y_n)$

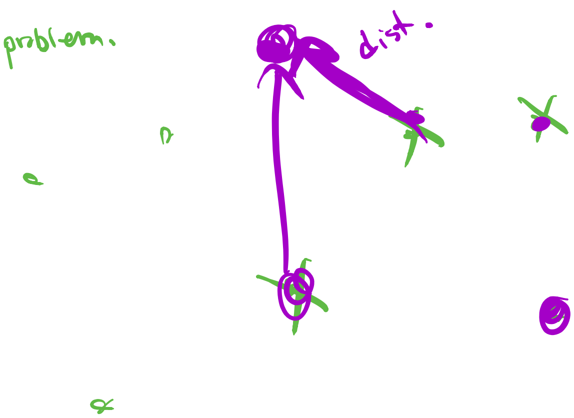
consider execution of alg on $\boxed{G_{-i}}$



$$y^{-i} = \begin{cases} 1 & j \text{ not matched} \\ Y_{i'} & j \text{ is matched to } i' \end{cases}$$

$$\beta_j^{-i} = 1 - g(y^{-i})$$

k-server problem.



min cost flow.

k servers c.r. k

randomized c.r. $\boxed{\ln k}$

$(\ln k)^{c} \ln n$.

Linear programming duality

FYI

(P)

$$\min\ c_1x_1 + c_2x_2 + \dots + c_nx_n$$

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \geq b_1$$
$$a_{21}x_1 + a_{22}x_2 + \dots a_{2n}x_n \geq b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \geq b_m$$
$$x_1, x_2, \dots, x_n \geq 0$$

$$\boxed{\begin{array}{l} \min\ c\cdot x \\ Ax \geq b \\ x \geq 0 \end{array}}$$

(D)

$$\max\ b_1y_1 + b_2y_2 + \dots + b_my_m$$

$$a_{11}y_1 + a_{21}y_2 + \dots + a_{m1}y_m \leq c_1$$
$$a_{12}y_2 + a_{22}y_2 + \dots + a_{m2}y_m \leq c_2$$
$$\vdots$$
$$a_{1n}y_1 + a_{2n}y_2 + \dots + a_{mn}y_m \leq c_n$$
$$y_1, y_2, \dots, y_m \geq 0$$

$$\boxed{\begin{array}{l} \max\ b\cdot y \\ y^TA \leq c^T \\ y \geq 0 \end{array}}$$

$\vec{x}, \vec{c} \in \mathbb{R}^n$
$\vec{y}, \vec{b} \in \mathbb{R}^m$
$A \in \mathbb{R}^{m\times n}$

By construction OPT (D) $\leq$ OPT (P) "weak duality"

x feasible for (P), y feasible for (D)

$\Rightarrow b\cdot y \leq c\cdot x$ $\quad (y^Tb \leq y^TAx \leq c^Tx)$

$\Rightarrow$ if $by = cx$ both are optimal

Duality Thm If (P) & (D) are primal-dual pair of LPs then one of following holds

1. both infeasible
2. (P) unbounded, (D) infeasible
3. (D) unbounded, (P) infeasible
4. both feasible & $\exists$ opt solns x^*, y^* s.t. $cx^* = by^*$

Complementary Slackness $\equiv$

$$\forall i \quad x_i^*\left(c_i - \sum_j y_j^* a_{ji}\right) = 0$$
$$\forall j \quad y_j^*\left(\sum_i a_{ji} x_i^* - b_j\right) = 0$$