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

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

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

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

Today

- finish up online matching analysis
- second moment method.
  - random graphs
  - percolation on trees

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

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

$x_{ij} \uparrow$ by Δ

$\uparrow \alpha_i$ by $g(y_i)dx$
$\uparrow \beta_j$ by $(1-g(y_i))dx$

y_i WL online:

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

Showed that $\forall (i,j)\in E$
$\forall \; 0\le y\le 1$

$$\alpha_i + \beta_j \ge \underbrace{\int_0^y g(z)dz + 1 - g(y)}_{= 1-\frac{1}{e} \approx 0.63}$$

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

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

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

Integral Online Matching

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.

Reinterpretation

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

when j arrives, match it to $\arg\min \left\{ Y_i \;\middle|\; \begin{array}{l} i \in N(j) \\ i \text{ unmatched} \end{array} \right\}$

Proof that this alg has c.r. $1 - \frac{1}{e}$

- Build feasible primal soln $\{x_{ij} \mid (i,j) \in E,\ x_{ij} \in \{0,1\}\}$

  and (not necessarily feasible) randomized dual soln $\{\alpha_i, \beta_j \mid i \in L, j \in R\}$

s.t.
(1) $P = \sum_{(i,j) \in E} x_{ij} \geq D = \sum_{i \in L} \alpha_i + \sum_{j \in R} \beta_j$

(2) $E\left(\frac{D}{c}\right)$ feasible dual for $c = 1 - \frac{1}{e}$

This suffices since

$$P \geq D \Rightarrow \frac{E(P)}{c} \geq \frac{E(D)}{c} \underset{\text{weak duality}}{\geq} OPT$$

$$\Rightarrow E(P) \geq c\,OPT = \left(1 - \tfrac{1}{e}\right) OPT$$

Initialize α_i's to be 0, β_j' to be 0.

To construct dual use $g : [0,1] \longrightarrow [0,1]$ $g \uparrow$

$g(1) = 1$

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

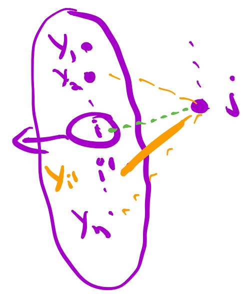
$\alpha_i = g(Y_i)$ i — j

$\alpha_{i'} = 0$ i' $\beta_j = 1 - g(Y_i)$

$\alpha_{i'} + \beta_j < 1$

Need to show $\dfrac{E(D)}{c}$ is feasible. $c = 1 - \frac{1}{e}$

Fix edge (i,j). Fix $Y_{-i} = (\underbrace{Y_1 \cdots Y_{i-1}, Y_{i+1} \cdots Y_n})$



Consider execution of alg on G_{-i}

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

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

$Y_i > Y_{i'}$

$Y_i < Y_{i'}$

Lemma

① In actual execution $\beta_j \geq \beta_j^{-i}$

② In actual execution i certainly matched if $Y_i < y^{-i}$

Lemma $\Rightarrow$ $\forall Y_{-i}$ $\forall (i,j) \in E$ $\boxed{\dfrac{E_{Y_i}[\alpha_i + \beta_j]}{1 - \frac{1}{e}} \geq 1}$

Pf

if $\underline{Y_i < y^{-i}}$ $\Rightarrow$ i will be matched

$$E_{Y_i}(\alpha_i) \geq \int_0^{y^{-i}} g(y)\, dy$$

Y_i gets matched for sure if $Y_i < y^{-i}$

for fixed Y_{-i}, $E_{Y_i}(\alpha_i) = \int_0^1 \underbrace{E(\alpha_i \mid Y_i = y)}_{g(y)\mathbb{1}_{i \text{ is matched}}} \underbrace{f_{Y_i}(y)}_{1} dy \geq \int_0^{y^{-i}} g(y)\, dy$

$$\boxed{\beta_j \geq \underline{\beta}_j^{-i} = 1 - g(y^{-i})}$$

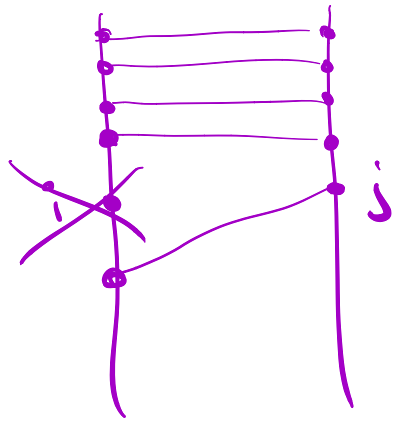
set $\alpha_i = g(Y_i)$ $\beta_j = 1 - g(Y_i)$

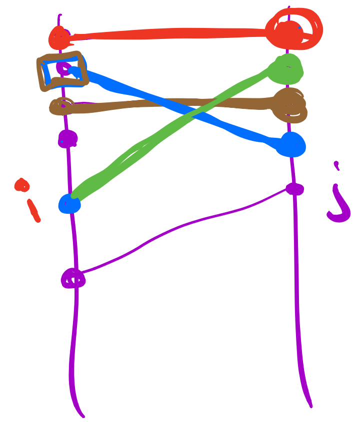
$$E_{Y_i}(\alpha_i + \beta_j) \geq \int_0^{y^{-i}} g(y)\, dy + 1 - g(y^{-i})$$

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

$$\mathop{E}_{Y_1, Y_2, \dots, Y_n} (\alpha_i + \beta_j) \geq 1 - \frac{1}{e}$$

To show $\beta_j \geq \beta_j^{-i}$





every vertex up to j has either same options or 1 additional option

$\Longrightarrow$

$$\beta_j = \underbrace{1 - g(Y_{\text{matched } j})} \geq \beta_j^{-i}$$

fractional version $\xRightarrow[\text{total mystery.}]{}$ randomized alg

2nd Moment Method.

First moment method.

- r.v. takes some value $\geq E(X)$
  $\leq E(X)$

- $X \geq 0$ int valued $\Rightarrow$ $Pr(X \geq 1) \leq E(X) = \sum_{k=1}^{\infty} k \, Pr(X=k)$
  $E(X) \ll 1 \Rightarrow X = 0$ w.h.p.

Proving $Pr(X \geq 1)$ large?
$E(X)$ large insufficient - could be HUGE with negligible prob

2nd Moment Method

- Chebychev's Inequality
  $\forall \lambda > 0 \quad Pr(|X-\mu| \geq \lambda\sigma) \leq \frac{1}{\lambda^2}$

- Another version $\quad Pr(X=0) \leq \frac{Var(X)}{E(X)^2}$

  Pf $\quad Pr(X=0) \leq Pr(|X-\mu| \geq \mu) \leq \frac{\sigma^2}{\mu^2} = \frac{Var(X)}{E(X)^2}$

Corollary: If $Var(X) = o(E(X))^2$ then $Pr(X>0) = 1 - o(1)$

Common case

$X = X_1 + X_2 + \dots + X_n$

Use $\quad Var(X) = \sum_{i=1}^{n} Var(X_i) + \sum_{i \neq j} Cov(X_i, X_j)$

$Cov(X,Y) = E[(X-E(X))(Y-E(Y))]$
$= E(XY) - E(X)E(Y)$
$Cov(X,Y) = 0$ if X & Y indep

Pf
$Var(X+Y) = E[(X+Y - E(X+Y))^2]$
$= E[(X-E(X))^2 + (Y-E(Y))^2 + 2(X-E(X))(Y-E(Y))]$
$= Var(X) + Var(Y) + 2\,Cov(X,Y)$

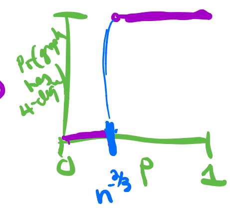
Often X_i's are indicator r.v.'s $\quad X_i \sim Ber(p_i)$
$Var(X_i) = p_i(1-p_i) \leq p_i = E(X_i)$

$\Rightarrow Var(X) \leq E(X) + \sum_{i \neq j} Cov(X_i, X_j)$

Example 1: Consider $G_{n,p}$, graph on n vertices where each edge present with prob p (independently)

For what p is $G_{n,p}$ likely to have clique with 4 or more vertices

Claim: ① almost surely doesn't for $p = o(n^{-2/3})$
② almost surely does for $p = \omega(n^{-2/3})$

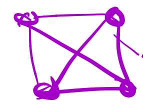


Proof of ①

$\binom{n}{4}$ possible 4-cliques $\qquad C_1, \dots, C_{\binom{n}{4}}$ enumeration

$X_i = \begin{cases} 1 & C_i \text{ clique} \\ 0 & \text{o.w.} \end{cases} \qquad X = \sum_{i=1}^{\binom{n}{4}} X_i \quad$ # 4-cliques

$$E(X) = \binom{n}{4} E(X_i) = \binom{n}{4} p^6 = c\, n^4 p^6$$



if $p = o(n^{-2/3})$ $\qquad n^{-2/3}$ $\quad p^6 = n^{-4}$

$E(X) = o(1) \implies \Pr(X \geq 1) \leq E(X) = o(1)$

Proof of ②

For $p = \omega(n^{-2/3})$ $\qquad E(X) \underset{\text{as } n \to \infty}{\longrightarrow} \infty$ $\qquad$ insufficient by itself

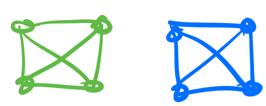
So we'll use 2nd moment method $\qquad \Pr(X=0) \leq \frac{\text{Var}(X)}{E(X)^2}$

we'll show $\quad \text{Var}(X) = o(E(X)^2)$

need to compute $\text{Var}(X)$

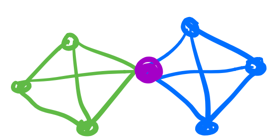
$\text{Cov}(X_i, X_j)$

$\boxed{\text{type 1}}$ $\quad C_i \cap C_j = \emptyset$



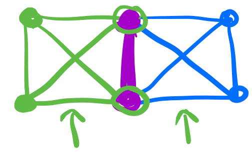
$\text{Cov}(X_i, X_j) = 0$

$\boxed{\text{type 2}}$ $\quad |C_i \cap C_j| = 1 \qquad \text{Cov}(X_i, X_j) = 0$



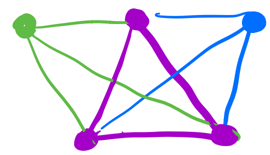
$\boxed{\text{type 3}}$ $|c_i \cap c_j| = 2$

$$Cov(X_i, X_j) \le E(\underline{X_i X_j}) = Pr(c_i \cap c_j \text{ in graph}) = p^{11}$$



$\boxed{\text{type 4}}$ $|c_i \cap c_j| = 3$

$$Cov(X_i X_j) \le p^9$$



$$Var(X) \le E(X) + \sum_{\substack{1 \le i,j \le m \\ i \ne j}} Cov(X_i X_j)$$

$$\le \binom{n}{4} p^6 + \underbrace{\binom{n}{6}\binom{6}{2,2,2} p^{11}}_{\text{type 3}} + \underbrace{\binom{n}{5}\binom{5}{3,1,1} p^9}_{\text{type 4}}$$

$$\le c\left(n^4 p^6 + n^6 p^{11} + n^5 p^9\right) = o\Big(\underbrace{E(X)^2}_{\left(\binom{n}{4} p^6\right)^2 = \Theta(n^8 p^{12})}\Big)$$

$$Pr(X=0) \le \frac{Var(X)}{[E(X)^2]} = o(1)$$

for $p = \omega(n^{-2/3})$

Evolution of Random Graphs

The following statements hold almost surely $\equiv$ with prob $\to 1$ as $n \to \infty$

$p = o\left(\frac{1}{n}\right)$
forest
no cycles

$\longrightarrow$

$c<1$
$p \le \frac{c}{n}$
largest connected component has size $\Theta(\log n)$ (a)

$\longrightarrow$

$p = \frac{1}{n}$
largest CC has size $\Theta(n^{2/3})$

$\longrightarrow$

$c>1$
$p \ge \frac{c}{n}$
largest CC has size $\Theta(n)$
"giant component" (c)

$\longrightarrow$

$p = \frac{\ln n}{n}$
graph is connected
has Hamiltonian cycle
contains perfect matching
....

(a) $p = \frac{0.99999}{n}$

(c) $\frac{1.00001}{n}$

Graph property (closed under isomorphism)
(if property holds & rename the nodes, still holds)

$Pr(G(n,p) \models A) \equiv Pr(\text{random graph from } G(n,p) \text{ has property } A)$

Monotone graph property: $\uparrow$ or $\downarrow$ as add edges
e.g.
- being connected
- having min degree 10
- containing a perfect matching
- containing Hamiltonian cycle

Cool: Every monotone graph property has a sharp threshold

A fn $p^* = p^*(n)$ is a threshold for a monotone $\uparrow$ property A if

$$\lim_{n\to\infty} Pr(G(n,p) \models A) = \begin{cases} 0 & \text{if } \frac{p}{p^*} \to 0 \\ 1 & \text{if } \frac{p}{p^*} \to \infty \end{cases}$$

Ex $p = n^{-2/3}$ threshold fn for K_4

$p \ll n^{-2/3}$
no K_4

$p \gg n^{-2/3}$
K_4

Graph properties expressible in "First order theory" of graphs

language vars $(x, y, z, \dots)$ represent vertices

equality $x = y$

adjacency $x \sim y$

$\wedge, \vee, \neg$ $\forall, \exists$

Examples:

G contains triangle $\exists x \exists y \exists z \; x \sim y \wedge y \sim z \wedge x \sim z$

no isolated pt

diameter 2 $\forall x \forall y \exists z \; (x = y) \vee (x \sim y) \vee (x \sim z \wedge z \sim y)$

G Hamiltonian, G connected not FO properties

$p = \frac{1}{2}$

$G(n, \frac{1}{2})$ $\frac{1}{2^{\binom{n}{2}}}$

uniform distn overall graphs

$\boxed{\text{Thm [Fagin] [Glebskii et al]}}$

$\forall$ fixed p, $0 < p < 1$, and any FO graph property A

$$\lim_{n \to \infty} \Pr(G(n,p) \models A) = 0 \text{ or } 1$$

$\boxed{\text{Thm [Shelah, Spencer]}}$

$\forall$ irrational α, $0 < \alpha < 1$, setting $p = p(n) = n^{-\alpha}$

and for any FO property A

$$\lim_{n \to \infty} \Pr\left(G(n,p) \models A\right) = 0 \text{ or } 1$$

$n^{-\pi/7}$

$\alpha = \frac{2}{3}$ rational.

this is about threshold fns

Ex $p = n^{-2/3}$ threshold fn for K_4

$p \ll n^{-2/3}$ no K_4

$p = n^{-2/3}$ $\Pr(K_4) = 1 - e^{-1/24}$

$p \gg n^{-2/3}$ K_4

Another 2nd moment inequality

$$\boxed{\Pr(X>0) \geq \frac{(E(X))^2}{E(X^2)}}$$

Follows from Cauchy-Schwartz

$$\boxed{[E(XY)]^2 \leq E(X^2)\,E(Y^2)}$$

Proof:

wlog $E(X^2)>0$, $E(Y^2)>0$

Let $U = \frac{X}{\sqrt{E(X^2)}}$ $V = \frac{Y}{\sqrt{E(Y^2)}}$

$$2|UV| \leq U^2 + V^2$$

$$\Rightarrow 2|E(UV)| \leq 2E(|UV|) \leq E(U^2) + E(V^2) = 2$$

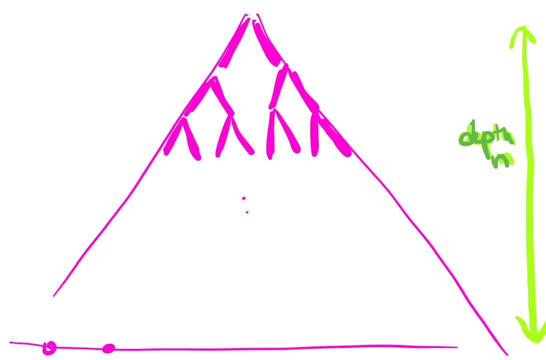
$$\Rightarrow [E(UV)]^2 \leq 1$$

$$\equiv [E(XY)]^2 \leq E(X^2)E(Y^2)$$

Percolation on Tree

complete binary tree of depth n
each edge present with prob p

Is there a surviving path from root to a leaf?



X #leaves reachable from root

$$X = X_1 + X_2 + \dots + X_n$$

X_i indicates if i^{th} leaf is reachable

$$E(X) =$$

$\boxed{p > \frac{1}{2}}$

$$\Pr(X>0) \geq \frac{[E(X)]^2}{E(X^2)}$$

$$X^2 = \left(\sum_i X_i\right)^2 = \sum_i X_i^2 + \sum_i \sum_{j\neq i} X_i X_j$$

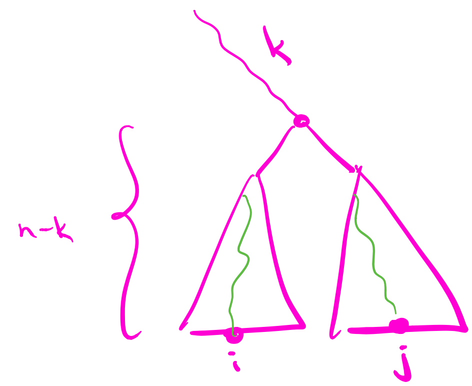
$$E(X_i X_j) = p^{2n-k}$$

$$\sum_i \sum_{j \neq i} E(X_i X_j) = \sum_{k=0}^{n-1} 2^k 2^{n-k} 2^{n-k-1} p^{2n-k}$$

$$= \frac{1}{2} \sum_{k=0}^{n} 2^{2n} \frac{1}{2^k} p^{2n} \frac{1}{p^k}$$

$$= \frac{1}{2} (2p)^{2n} \underbrace{\sum_{k=0}^{n-1} \frac{1}{(2p)^k}}_{\frac{1-\frac{1}{(2p)^n}}{1-\frac{1}{2p}}}$$

$$\leq \frac{1}{2}(2p)^{2n} \frac{2p}{2p-1} = \frac{p}{2p-1}(2p)^{2n}$$



$\boxed{p = \frac{1}{2}}$

$$\sum_i \sum_{j \neq i} E(X_i X_j) = \sum_{k=0}^{n} 2^k 2^{n-k} 2^{n-k-1} p^{2n-k}$$

$$= \frac{n}{2}$$

$E(X^2) = 1 + \frac{n}{2}$ $\qquad$ $[E(X)]^2 = 1$

Let's prove that this is the right order of magnitude