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

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

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

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

is feasible dual

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

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

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

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

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

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

Probabilistic Method.

First moment method.

- r.v. takes some value $\geq E(X)$
  $\leq E(X)$
- $X \geq 0$ int valued $\Rightarrow$ $Pr(X \geq 1) \leq E(X)$
  $E(X) \ll 1 \Rightarrow X = 0$ 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
$Var(X+Y) = E[(X+Y - E(X+Y))^2]$
$= E[(X-E(X))^2 + (Y-E(Y))^2 + 2(X-E(X))(Y-E(Y))]$
$= Var(X) + Var(Y) + 2Cov(X,Y)$

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 $\quad C_1, \dots, C_{\binom{n}{4}}$ 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

Proof of (2)

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

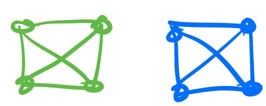
So we'll use 2nd moment method

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

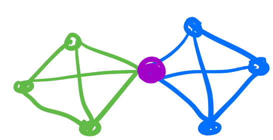
need to compute $Var(X)$

$Cov(X_i, X_j)$

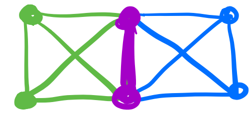
$\boxed{\text{type 1}}$ $\quad C_i \cap C_j = \emptyset$



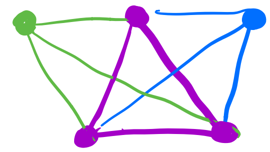
$\boxed{\text{type 2}}$ $\quad |C_i \cap C_j| = 1$



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



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



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

Evolution of Random Graphs

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

$p = o\left(\frac{1}{n}\right)$ forest, no cycles → $p \le \frac{c}{n}$, $c<1$: largest connected component has size $\Theta(\log n)$ → $p = \frac{1}{n}$: largest CC has size $\Theta(n^{2/3})$ → $p \ge \frac{c}{n}$, $c>1$: largest CC has size $\Theta(n)$ "giant component" → $p = \frac{\ln n}{n}$: graph is connected, has Hamiltonian cycle, contains perfect matching, ...

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: ↑ or ↓ 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 ↑ 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 $(x, y, z, \ldots)$ represent vertices

equality $x = y$

adjacency $x \sim y$

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

Examples:

G contains triangle

no isolated pt

diameter 2

G Hamiltonian, G connected not FO properties

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

$\forall$ 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

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

this is about threshold fns

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

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

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

$\Rightarrow 2|E(UV)| \leq 2E(|UV|)$

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

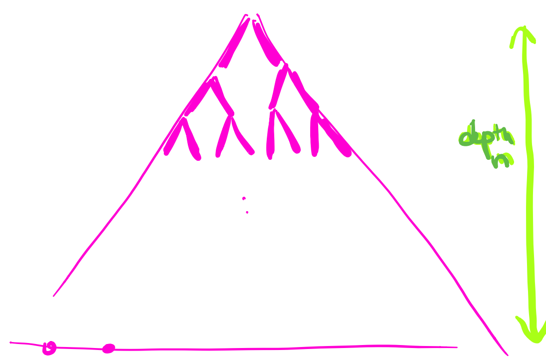
$\Rightarrow [E(UV)]^2 \leq 1$

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

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

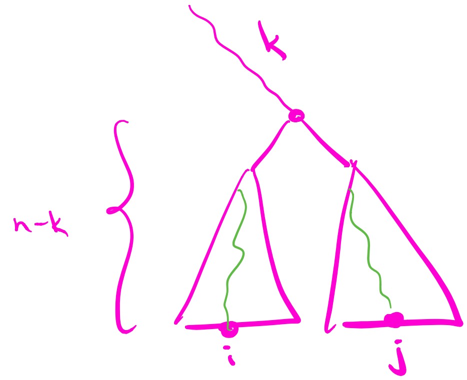
$X = X_1 + X_2 + \dots + X_n$

X_i indicates if i^{th} leaf is reachable

$E(X) =$

$\boxed{p > \frac{1}{2}}$ $\quad Pr(X>0) \geq \frac{[E(X)]^2}{E(X^2)}$

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



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

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

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

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

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

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

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

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

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

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