Lecture 18

Monday, May 22, 2017 2:16 PM

Independent Sets in Bounded-Degree, 3-colorable graphs

Input :-
- $G = (V, E)$
- $\deg(v) \leq d \quad \forall v$
- Promise :- G is 3-colorable (*)

Output :- $I \subseteq V$, I is independent

Obj :- Maximize $|I|$

Randomized Rounding with a fix (even w/o (*))

- Select I_1 where $v \in V$ is sampled w.p. p
- Remove any pair i,j from I_1 if $(i,j) \in E$
- Return remaining set $I = I_1 - I_2$

- $\mathbb{E}[|I_1|] = np$
- $\mathbb{E}[|I_2|] \leq \sum_{(i,j) \in E} \Pr[i \in I_1 \wedge j \in I_1]$

$$\leq \frac{nd}{2} \cdot p^2$$

$$\therefore \mathbb{E}[|I|] \geq np - \frac{nd}{2} \cdot p^2$$

$$\therefore \mathbb{E}[|I|] \geq np - \frac{nd}{2} \cdot p^2$$
$$= n\left(p - \frac{dp^2}{2}\right)$$

$\therefore$ If p was chosen so that

$p = dp^2$ i.e. $p = \frac{1}{d}$; then...

$$\mathbb{E}[|I|] \geq \frac{n}{2d}$$

linear in d.

Today, we see how to get a sublinear dep. on d when (*), ie, G is 3-colorable.
Uses SDPs and a "simple" rounding trick.

<u>Two ideas</u>:

1. Use G is 3-colorable to obtain an embedding of G to S_n, the n-dimn sphere, with edges' epts being "far apart"

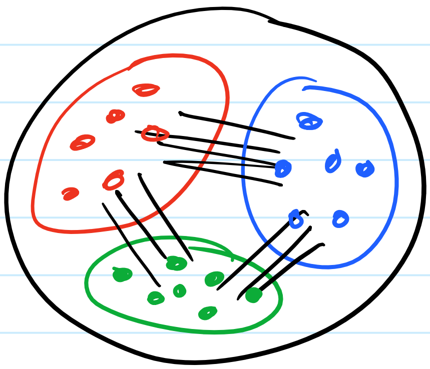
2. Use a similar idea as above (RR with a fix) to get a better IS.

Except of ind sampling, it'll be correlated by the soln.

What can be said about 3-colorable Graphs?

→ Checking if G is 3-COL or not is NP-complete. No exact characterization is known.

→ We use a necessary cond^n.



Note: G is 3-COL

$\Downarrow$

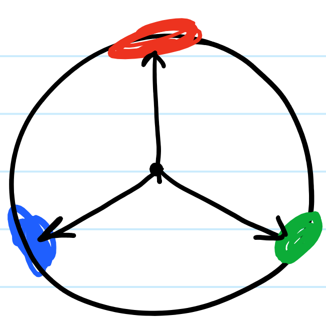
max IS $\geq n/3$

We will be nowhere close to finding this large a IS.

$\therefore$ $\exists$ an embedding $\phi : V \to \mathbb{R}^2$

s.t. $\forall i$, $\|\phi(i)\|_2 = 1$ (unit circle)

if $(i,j) \in E$, $\angle \phi(i), \phi(j) = 120^\circ$



→ If G is 3-COL, the following has a feasible soln:

$$\{ (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n : \quad \forall (i,j) \in E : \langle v_i, v_j \rangle = -1/2, \quad \|v_i\|_2 = 1 \}$$

⇒ <u>SDP-formulation</u>: The following system has a feasible soln

$$\{ X \in \mathbb{R}^{n \times n} : \quad X_{ii} = 1, \ \forall i = 1..n,$$

(SDP-col)

$$X_{ij} = -1/2, \ \forall (i,j) \in E$$

$$X \succcurlyeq 0 \}$$

G is 3-col ⟹ SDP-col is feasible.

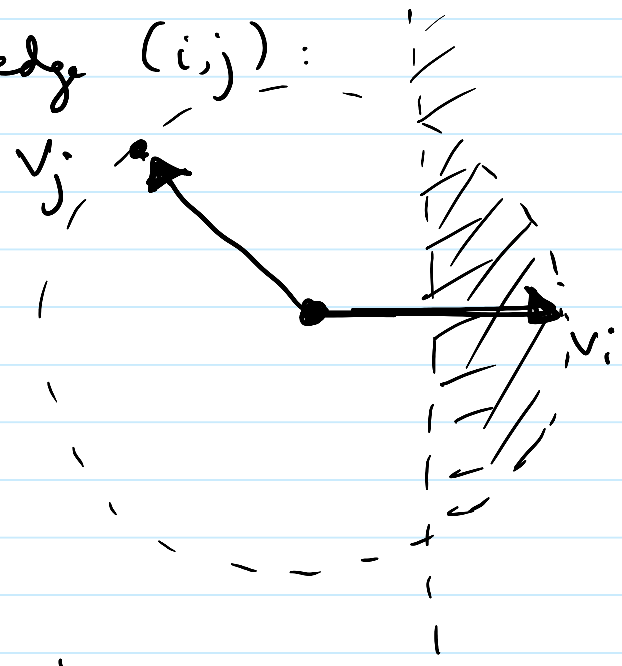
∴ We may assume we have unit vectors $\{v_1, \ldots, v_n\}$ st $\langle v_i, v_j \rangle = -\frac{1}{2}$ for all edges $(i,j) \in E$

— End of Part 1.

Randomized Rounding with a fix... part deux

- Two forces at loggerheads
  - Want to sample so that lots of vertices in I_1.
  - But not so agressively that many edges enter. I_1.

- Pick an edge (i,j):



$\rightarrow \langle v_i, v_j \rangle = -\frac{1}{2}$

$$\Rightarrow \quad \|v_i + v_j\|^2 = \|v_i\|^2 + \|v_j\|^2 + 2\langle v_i, v_j\rangle$$

$$= 2 - 1 = 1$$

$\therefore \|v_i + v_j\| = 1$ .... while for non-edges it could be as large as 2.

→ **ALGO**: (Karger-Motwani-Sudan aka KMS algo)

- Sample a random unit gaussian g in $\mathbb{R}^n$.
- $I_1 := \{ i \mid \langle v_i, g \rangle \geq c \}$ for some c to be chosen later
- $I = I_1 - I_2$, where I_2 are the set of edges in I_1

Facts about Gaussians

⊙ 1-dimn: $X \sim N(0, \sigma)$

$$\Rightarrow \Pr[X = x] = \frac{1}{\sqrt{2\pi\sigma}} e^{-x^2/2\sigma^2}$$

⊙ n-dimn: $g = (g_1, \ldots, g_n)$

each $g_i \sim N(0,1)$ independent

- Unit Gaussian : $\dfrac{g}{\|g\|_2}$

• If g is a unit Gaussian in $\mathbb{R}^n$, and v is any fixed vector in $\mathbb{R}^n$, then

$\langle v, g \rangle$ is a random variable

with $\langle v, g \rangle \sim N(0, \|v\|)$

- Sum of ind gaussian rvs is Gaussian
- Variances add up.

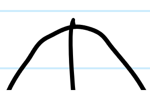
• "Error - function" / "Quantile f^n"

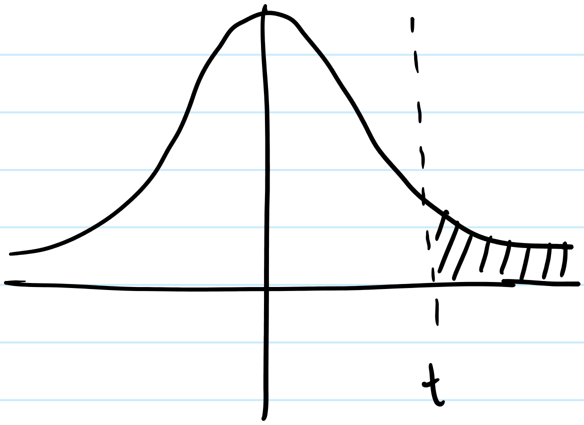
$$\mathrm{erf}(t) := \Pr[\underset{\uparrow \atop \sim N(0,1)}{X} \geq t] = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2}\, dx$$

<u>Bounds</u>:

$\forall t > 0$:

$$\left(\frac{1}{t} - \frac{1}{t^3}\right) e^{-t^2/2} \leq \sqrt{2\pi} \cdot \mathrm{erf}(t) \leq \frac{1}{t} \cdot e^{-t^2/2}$$

 :



Analysis of the KMS algorithm

$$\mathbb{E}[|I_1|] = \sum_{i \in V} \Pr[\langle v_i, g \rangle \geq c]$$

$N(0,1)$ $\because \|v_i\| = 1$

$$= n \cdot \mathrm{erf}(c)$$

$$\mathbb{E}[|I_2|] = \sum_{(i,j) \in E} \Pr[\langle v_i, g \rangle \geq c \ \& \ \langle v_j, g \rangle \geq c]$$

$$\leq \sum_{(i,j) \in E} \mathbb{P}[\langle v_i + v_j, g \rangle \geq 2c]$$

$N(0, \|v_i + v_j\|)$

$$\leq \frac{nd}{2} \cdot \mathrm{erf}\left(\frac{2c}{\|v_i + v_j\|}\right)$$

$$\leq \frac{nd}{2} \cdot \operatorname{erf}(2c)$$

$$\therefore \mathbb{E}[|I|] \geq n \operatorname{erf}(c) - \frac{nd}{2} \operatorname{erf}(2c)$$

$$\approx n \left[\frac{1}{c} e^{-c^2/2} - \frac{d}{2} \cdot \frac{1}{2c} \cdot e^{-2c^2} \right]$$

if $e^{-c^2/2} \approx \frac{d}{2} e^{-2c^2}$

i.e. $e^{\frac{3c^2}{2}} \approx \frac{d}{2}$ i.e. $c = \sqrt{\frac{2}{3} \ln\left(\frac{d}{2}\right)}$

$$\approx \frac{n}{\sqrt{\frac{2}{3}\ln\left(\frac{d}{2}\right)}} \cdot \left[\frac{1}{4} \cdot e^{-\frac{1}{3}\ln(d/2)} \right]$$

$$\geq \frac{n}{A \cdot d^{1/3} \cdot \sqrt{\ln d}}$$ for some constant A

cube-root of d instead of linear!

$\boxed{\text{Thm :- If } G \text{ has max-degree } d \text{ and is 3-colorable, then one can find an independent set of size} \geq \dfrac{n}{\tilde{O}(d^{1/3})}}$

hides log-factors.

Corollary :- Once we find an independent set, we can "pluck" it out giving it one color and repeat. Thus we can color 3-colorable graphs with $\tilde{O}(d^{1/3})$ colors.

Using "another trick", this gives a coloring of 3-col. graphs using $\tilde{O}(n^{0.25})$ colors....

sounds ridiculous?

The best known algorithm till date is $\tilde{O}(n^{0.2038..})$ colors. from 2012!