Wednesday 8/20/14

Probability theory basics:

Example: n coin flips

Sample space: $\Omega = \{H, T\}^n$

Event: $\mathcal{E} \subset \Omega$

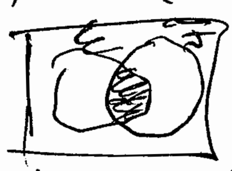
example: 1^{st} flip is H, $\geq \frac{n}{4}$ flips are H

Outcome: $x \in \Omega$, e.g. $x = HHTH\ldots$

$$\Pr(\mathcal{E}) = \frac{|\mathcal{E}|}{|\Omega|} = \sum_{x \in \mathcal{E}} \Pr(x)$$

for events $\mathcal{E}, \mathcal{F}$,

$$\Pr(\mathcal{E} \cup \mathcal{F}) = \Pr(\mathcal{E}) + \Pr(\mathcal{F}) - \Pr(\mathcal{E} \cap \mathcal{F})$$



Hence, $\Pr(\mathcal{E} \cup \mathcal{F}) \leq \Pr(\mathcal{E}) + \Pr(\mathcal{F})$

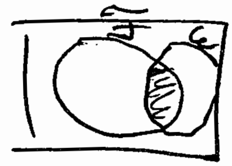
More generally,

for events $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_\ell$

Union bound: $\Pr(\mathcal{E}_1 \cup \mathcal{E}_2 \cup \cdots \cup \mathcal{E}_\ell) \leq \Pr(\mathcal{E}_1) + \cdots + \Pr(\mathcal{E}$

Bayes Theorem: $\Pr(\mathcal{E} \mid \mathcal{F}) = \dfrac{\Pr(\mathcal{E} \cap \mathcal{F})}{\Pr(\mathcal{F})}$

Prob. of $\mathcal{E}$ occurring given that $\mathcal{F}$ has occurred



Example: roll 2 dice

$\mathcal{E}$ = 1st roll is 4

$\mathcal{F}$ = sum is ≥ 7

$$\Pr(\mathcal{E} \mid \mathcal{F}) = \frac{\Pr(1^{st} \text{ is } 4, \text{sum is} \geq 7)}{\Pr(1^{st} \text{ is } 4)} = \frac{\left(\frac{1}{6}\right)\left(\frac{4}{6}\right)}{\left(\frac{1}{6}\right)} = \frac{2}{3}$$

Rearranging, $\Pr(\mathcal{E} \cap \mathcal{F}) = \Pr(\mathcal{E} \mid \mathcal{F}) \Pr(\mathcal{F})$

$$= \Pr(\mathcal{F}) \Pr(\mathcal{E} \mid \mathcal{F}) = \Pr(\mathcal{E}) \Pr(\mathcal{F} \mid \mathcal{E})$$

More generally, for events $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_\ell$

$$\Pr(\mathcal{E}_1 \cap \mathcal{E}_2 \cap \cdots \cap \mathcal{E}_\ell)$$
$$= \Pr(\mathcal{E}_1) \Pr(\mathcal{E}_2 \cap \mathcal{E}_3 \cap \cdots \cap \mathcal{E}_\ell \mid \mathcal{E}_1)$$
$$= \Pr(\mathcal{E}_1) \Pr(\mathcal{E}_2 \mid \mathcal{E}_1) \Pr(\mathcal{E}_3 \cap \cdots \cap \mathcal{E}_\ell \mid \mathcal{E}_1 \cap \mathcal{E}_2)$$

Chain rule: $= \Pr(\mathcal{E}_1) \Pr(\mathcal{E}_2 \mid \mathcal{E}_1) \Pr(\mathcal{E}_3 \mid \mathcal{E}_1 \cap \mathcal{E}_2) \times \cdots \times \Pr(\mathcal{E}_\ell \mid \mathcal{E}_1 \cap \cdots \cap \mathcal{E}_{\ell-1})$

Analysis of Karger's algorithm.

Fix a min cut $S^* \leftrightarrow \overline{S^*}$

<u>Lemma</u>: $\Pr\left(\begin{array}{c}\text{Karger's alg.}\\ \text{finds } S^* \leftrightarrow \overline{S^*}\end{array}\right) \geq \frac{1}{\binom{n}{2}}$

<u>Proof</u>:

Let $e_1, e_2, \ldots, e_{n-2}$ denote the contracted edges

Let $G_0, G_1, \ldots, G_{n-2}$ denote the sequence of graphs.

G_0 = input graph G, & for $i > 0$, $G_i = G_{i-1} / e_i$.

We need that we never contract an edge in $\delta_G(S^*)$.

Let $\mathcal{E}_i$ = event that $e_i \notin \delta_G(S^*)$

Let $k = |\delta_G(S^*)|$ = size of min cut $S^* \leftrightarrow \overline{S^*}$.

We need that:

$$e_1, e_2, \ldots, e_{n-2} \notin \delta_G(S^*) = \mathcal{E}_1 \cap \mathcal{E}_2 \cap \cdots \cap \mathcal{E}_{n-2}$$

$$\Pr\left(\begin{array}{c}\text{algorithm outputs}\\ \text{cut } S^* \leftrightarrow \overline{S^*}\end{array}\right) = \Pr(\mathcal{E}_1 \cap \mathcal{E}_2 \cap \cdots \cap \mathcal{E}_{n-2})$$

$$= \Pr(\mathcal{E}_1) \Pr(\mathcal{E}_2 | \mathcal{E}_1) \Pr(\mathcal{E}_3 | \mathcal{E}_1 \cap \mathcal{E}_2) \times \cdots \times \Pr(\mathcal{E}_{n-2} | \mathcal{E}_1 \cap \cdots \cap \mathcal{E}_{n-3})$$

What's $\Pr(\mathcal{E}_1)$?

k edges in the min cut

$m = |E|$ edges.

$$\Pr(\mathcal{E}_1) = 1 - \frac{k}{m}$$

Want a bound on m in terms of n & k:

Since min cut is size $= k$,

for a vertex v its degree is $\geq k$

(otherwise the cut $v \leftrightarrow V - v$ is size $< k$)

Thus min degree is $\geq k$ and hence

$$m = |E| = \frac{1}{2} \sum_{v \in V} \deg(v) \geq \frac{nk}{2}$$

$$m \geq \frac{nk}{2}$$

Therefore,

$$\Pr(\mathcal{E}_1) \geq 1 - \frac{k}{nk/2} = 1 - \frac{2}{n}.$$

Suppose $\mathcal{E}_1$ occurred, so $(S^*, \overline{S^*})$ is still preserved.

What's the min cut size in G_1?

$(S^*, \overline{S^*})$ is still there so there is a cut of size $= k$.

Is there a new smaller cut?

No! Every cut in G_1 corresponds to a cut in $G_0 = G$

(the reverse is not necessarily true)

So if there's a cut in G_1 of size $< k$ then there is also a cut in $G_0 = G$ of size $< k$ which contradicts the assumption that the min cut size in G is $= k$.

Hence G_1 has min cut size $= k$

& thus min degree $\geq k$ (as before).

let $m_1 = \#$ of edges in G_1.

$\#$ of vertices in $G_1 = n-1$.

Therefore, $m_1 \geq \frac{(n-1)k}{2}$

Thus,

$$\Pr(\mathcal{E}_2 \mid \mathcal{E}_1) \geq 1 - \frac{k}{(n-1)k/2} = 1 - \frac{2}{n-1}$$

Similarly,

$$\Pr(\mathcal{E}_i \mid \mathcal{E}_1 \cap \cdots \cap \mathcal{E}_{i-1}) \geq 1 - \frac{2}{n-(i-1)}$$

Plugging these back in,

$$\Pr\left(\begin{matrix}\text{algorithm outputs} \\ \text{cut } S^* \leftrightarrow S^*\end{matrix}\right) = \Pr(\mathcal{E}_1 \cap \cdots \cap \mathcal{E}_{n-2})$$

$$= \Pr(\mathcal{E}_1)\Pr(\mathcal{E}_2 \mid \mathcal{E}_1)\Pr(\mathcal{E}_3 \mid \mathcal{E}_1 \cap \mathcal{E}_2) \times \cdots \times \Pr(\mathcal{E}_{n-2} \mid \mathcal{E}_1 \cap \cdots \cap \mathcal{E}_{n-3})$$

$$\geq \left(1-\frac{2}{n}\right)\left(1-\frac{2}{n-1}\right)\left(1-\frac{2}{n-2}\right) \times \cdots \times \left(1-\frac{2}{3}\right)$$

$$= \left(\frac{n-\cancel{2}}{n}\right)\left(\frac{\cancel{n-3}}{n-1}\right)\left(\frac{\cancel{n-4}}{\cancel{n-2}}\right)\left(\frac{\cancel{n-5}}{\cancel{n-3}}\right) \times \cdots \times \left(\frac{\cancel{4}}{6}\right)\left(\frac{\cancel{3}}{\cancel{5}}\right)\left(\frac{2}{\cancel{4}}\right)\left(\frac{1}{\cancel{3}}\right)$$

$$= \frac{2}{(n)(n-1)} = \frac{1}{\binom{n}{2}}.$$

∎

Min cut is of size k.

How many cuts of size k can a graph have?

Say there are ℓ min cuts.

Let $S_1, S_2, \ldots, S_\ell$ denote them.

Let $\mathcal{E}_i$ = event that Karger's algorithm outputs $S_i \leftrightarrow \overline{S_i}$

We know that for all i,

$$\Pr(\mathcal{E}_i) \geq \frac{1}{\binom{n}{2}}$$

Only 1 cut is output so

$$\Pr(\mathcal{E}_i \cap \mathcal{E}_j) = 0$$

Thus,

$$\Pr(\mathcal{E}_1 \cup \cdots \cup \mathcal{E}_\ell) = \Pr(\mathcal{E}_1) + \Pr(\mathcal{E}_2) + \cdots + \Pr(\mathcal{E}_\ell) \geq \frac{\ell}{\binom{n}{2}}$$

≤ 1 since the prob. of an event is ≤ 1.

Therefore, $\ell \leq \binom{n}{2}$

So every graph has $\leq \binom{n}{2}$ cuts of minimum size.

Running time:

$O(n^2 \log n)$ runs of Karger's algorithm.
Each run takes $O(n^2)$ time.
$\Rightarrow O(n^4 \log n)$ total time.

Faster version:

Key idea: in early rounds we're likely to succeed.

$$\Pr(\mathcal{E}_1) \geq 1 - \frac{2}{n}$$

But in later rounds it's less likely,

$$\Pr(\mathcal{E}_{n-2}) \geq \frac{2}{3}$$

So instead of re-running the whole algorithm, repeat the later rounds.

$\Pr($alg. is successful for 1^{st} ℓ contractions$)$

$$= \Pr(\mathcal{E}_1 \cap \cdots \cap \mathcal{E}_\ell)$$

$$\geq \left(\frac{n-2}{n}\right)\left(\frac{n-3}{n-1}\right)\left(\frac{n-4}{n-2}\right) \times \cdots \times \left(\frac{n-\ell-2}{n-\ell}\right)$$

$$= \frac{1}{\binom{n}{2}} \left(\frac{n-\ell-1}{n-\ell-3}\right)\left(\frac{n-\ell-2}{n-\ell-4}\right) \times \cdots \times \left(\frac{3}{1}\right)$$

$$= \frac{\binom{n-\ell-1}{2}}{\binom{n}{2}}$$

So probability that the desired min cut survives down to j vertices means
$l = n-j-1$ rounds were successful

$$\Rightarrow \geq \frac{\binom{j}{2}}{\binom{n}{2}}$$

Faster algorithm:

From multigraph $G=(V,E)$,
if $|V| > 2$ then:
- Repeat twice:
  1. Run Karger's algorithm down to $\frac{n}{\sqrt{2}}$ vertices
  2. Recurse on the resulting graph
- Take the 2 outputs & return the best of these 2 cuts.

Running time:

for step 1. for step 2.

$$T(n) = O(n^2) + 2T\left(\frac{n}{\sqrt{2}}\right)$$

$$= O(n^2 \log n) \text{ by the Master theorem.}$$

What's the success probability?

Let $P(n)$ = prob. alg. succeeds on n vertices

$$P(n) \geq \underbrace{1 - \left(1 - \underset{\substack{\downarrow \\ \text{step 1.}}}{\tfrac{1}{2}}\, \underset{\substack{\downarrow \\ \text{step 2.}}}{P\left(\tfrac{n}{\sqrt{2}}\right)}\right)^2}$$

step 1. is $\geq 1/2$ since $\dfrac{\binom{n/\sqrt{2}}{2}}{\binom{n}{2}} \geq \dfrac{1}{2}$

So steps 1 & 2 both succeed $\geq \frac{1}{2} P\left(\frac{n}{\sqrt{2}}\right)$

something fails $\leq 1 - \frac{1}{2} P(n/\sqrt{2})$

in both tries something fails $\leq \left(1 - \frac{1}{2} P(n/\sqrt{2})\right)^2$

$\geq$ one try succeeds $\geq 1 -$

this solves to $\Omega\left(\frac{1}{\log n}\right)$ need to do inductive proof.

Run this new algorithm $O(\log^2 n)$ times to get success probability $\geq 1 - \frac{1}{n^c}$.

Then $O(n^2 \log^3 n)$ total running time.