Property Testing

Monday, July 14, 2014 2:26 PM

$S(f,x) = |\{i \mid f(x) \neq f(x^{(i)})\}|$

$x^{(i)} = \begin{cases} x_j & j \neq i \\ \overline{x_j} & j = i \end{cases}$

$S(f) = \max_x S(f,x)$, $as(f) = \frac{1}{2^n} \sum_x S(f,x)$

$\{x \mid f(x) = 1\} \subset \{0,1\}^n$

properties sbspace of the domain

$x \in \{0,1\}^{n^2}$ $P \subseteq \{0,1\}^{n^2}$ $P = \{$directed bipartite graph$\}$

$x \in \{0,1\}^{2^n}$ $P \subseteq \{0,1\}^{2^n}$ $P = \{$linear boolean function on n vars$\}$

$G_1 = K_{n/2} \cup K_{n/2}$. $P = \{$isomorphic copies of $G_1\}$

given a graph G. Is $G \in P$?

want to minimize "Query". don't care about time.

1) Deterministic algo (Decision tree).

2) Randomized algo.

Pr(I get the correct answer) $\geq 2/3$

---

pick $i, j \in [n]$, $i \neq j$

probability of getting an edge $= \begin{cases} 1 \\ 1/2 \end{cases}$

$P \subseteq \{0,1\}^n$. $d(x,y) = |\{i \mid x_i \neq y_i\}|$ Hamming distance.

$d(x,P) = \min_{y \in P} d(x,y)$

Property Testing

completeness

with probability 1.

$P \subseteq \{0,1\}^n$. given $x \in \{0,1\}^n$. want algo A. if $x \in P$, then $A(x) =$ yes. if $d(x,P)$ is big, then $A(x) =$ no.

Soundness

$P =$ collection of graphs. want A st if $x \in P$ then $A(x) = 1$. if $d(x,P) \geq \epsilon n^2$, $A(x) = 0$

$x \in \{0,1\}^{n^2}$ (with probability $\geq 2/3$)

choose random k pts? two

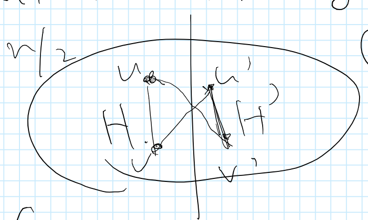
1-sided — 2-sided. cliques with close size.

$\Omega(n)$. $O(n\,\mathrm{poly}(\log n))$.

---

use any G_1. $\Omega(n)$. $O(n\,\mathrm{poly}(\log n))$. correct bound for $\forall G$.

G_1 random graph. $H \mid I$ — 2-sided $\tilde{\Theta}(\sqrt{n})$ lower bound



G_1 H' copy of H.

pick $\sqrt{n}$ pts

G_2. clone graph.

1-sided: certificate complexity.

D_N: distribution on graphs that are far

D_Y: — — — — — — — that are in P.

set of query

$\forall Q$. $|Q| = q$ $\quad D_N|_Q = D_Y|_Q$, then q is a lower bound.

$\forall x \in \{0,1\}^q$. $\Pr[\text{Answer} = x]$. distribution of answers

---

Alice | Bob
G_1 | G_2

communicate.

check that whether $G_1 \cong G_2$, min # of communications

pick random subset, check parity.

$P = \{$ Isomorphic copies of $G_1 \}$.

$d(G_2, P) \le \epsilon n^2 \longrightarrow$ yes.

$d(G_2, P) \ge \delta n^2 \longrightarrow$ no $\quad 0 < \epsilon < \delta < 1$

Thursday

P property of graph. Given a graph G (adj matrix)

$\nearrow G \in P$

$\searrow d(G, P) \ge \epsilon n^2$. Dense graph Model

Does G have a triangle?

$\nearrow$ G triangle free accept

$\searrow$ G far from triangle free. reject

1-sided.

Random pick (u, v, w) check whether form a ~~triangle~~ triangle.

$\Pr((u,v,w) \text{ is a triangle}) = \dfrac{\#\text{triangles}}{\binom{n}{3}} \ge \dfrac{1}{\sqrt{n}}$

Constant query $\Rightarrow \#\Delta = \Theta(n^3)$ if G is ϵ-far.

Regularity lemma. (Szemeredi)

Can you get $o(n^2)$ $\sqrt{n}$ query algo for the Δ-freeness.

$o(\sqrt{n})$.

lower bound for Δ-free $\Omega(1/\epsilon)$.

Is G planar?

$\nearrow$ If G planar accept

$\searrow$ If G is far from planar, reject

2-sided. random query

1-sided error query complexity is $\Omega(n^2)$.

graph w/ ϵn^2 edges. not ~~[illegible]~~ planar.

take random edge. pr(exists) = ε. pick $\frac{c \cdot n}{\varepsilon}$ times.
E(#edges) = $c \cdot n$.

Is G connected?

Sparse graph Model

Assume that G is ~~d-regular~~ max degree d graph. Give adjacency list.

Queries: Give the k^{th} neighbor of vertex v.

G connected Acc.

G far from connected. Rej

↑ One need to remove/add εn number of edges to make it connected.

In sparse graph Model the Δ-freeness has query $O(\sqrt{n})$
take random walk. ? random walk contains a triangle

Is s-t connected in $\vec{G}$? (path $s \to t$).

Orientation model.

$\vec{G}$ - directed graph. G underlying undirected graph.

G is known. Queries: the orientation of edges.

$\vec{G}$ is far from P. if you have to reverse more than εm edges.

1). $s \to t$ connected → constant $2^{2^{2^{1/\varepsilon}}}$ → (- Fischer - M - Newmann.
2) is there path $s \to t$ and $t \to s$
3). is there a path $s \to$ everything?
4). Strongly connected.

Only 1) known. If G is ε-far from $s \to t$, then ∃ a constant certificate.

5) Eulerian-ness. $\Theta(\sqrt{n})$. in degree = outdegree. $\forall$ vertex.

Fischer-Lachish-M-

Distribution on $\{1, \dots n\}$. D.

$f: [m] \longrightarrow \{1, \dots n\}$. $m >>> n$.

$\mu(i) = \frac{|f^{-1}(i)|}{m}$

Def $D: \{0, \dots n\} \longrightarrow [0,1]$. $\Sigma D(i) = 1$. prob distribution on $\{0, \dots n\}$.

Is this distribution uniform?

$\forall i \in \{0, \dots n\}$ $D(i) = 1/n$

Sampling: get i with probability $D(i)$.

D is uniform. $\Leftrightarrow$ $\forall i$ $D(i) = 1/n$.

D is ε-far from uniform.

$\Sigma |D(i) - \frac{1}{n}| \geq \varepsilon$. L_1 distance.

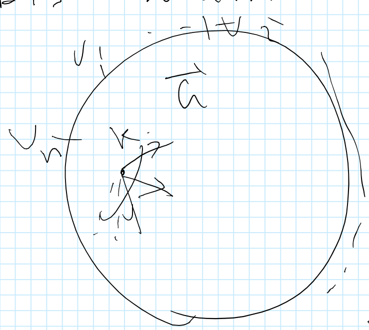
If $\leq \sqrt{n}$ queries no collision in both case.

upper bound.

Query complexity is $\Theta(\sqrt{n})$.

Batu-Fischer-Fort-Rubin-White.

Distribution not by sampling.



t-step

Probability that I am at vertex v.

① Pick a vertex randomly $1/n$.

② pick an outgoing edge at random.

then move to the neighbor.

for all vertex v. probability that I'm at v.

whether this is uniform?

Given G. known. G uniform as stationary distribution

Does G has unif as a stationary distribution.

If G ϵ-far from having unif as a stationary distribution.

If G is d-regular, then we have tight bound of $\Theta(\sqrt{n})$

aperiodic then all in degree need to be the same.

Markov chain.

p_{ij} transition probability $i \to j$.

$T = (p_{ij})$ $\quad \sum_j p_{ij} = 1$. $\quad q = (q_1, \cdots q_n)$ $\quad q_i$ probability of

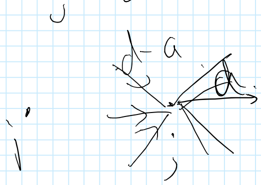
getting i initially

time $t \to t+1$. make a transition,

$(1, \cdots 1) T = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ $\quad$ Column Sum $= 1$.

j neighbors of i, then $p_{ij} = \begin{cases} \frac{1}{\deg^+ i} & i \to j \\ 0 & o/w \end{cases}$

$\Rightarrow \sum_j p_{ij} = 1 \Rightarrow \sum_{i \to j} \frac{1}{\deg^+ i} = 1$.



$f : \{\pm 1\}^n \longrightarrow \{\pm 1\}$

$s(f) = \max_x s(f, x)$. $\quad as(f) = \mathbb{E}_x (s(f, x))$.

$f(x_1 \cdots x_n) = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S(x)$. $\quad \chi_S(x) = \prod_{i \in S} x_i$.

$\langle f, g \rangle = \mathbb{E}_x (f(x) g(x))$. $\quad \chi_S$ ONB. $\quad \hat{f}(S) = \langle f, \chi_S \rangle$.

$\deg(f) = \max \{|S| \mid \hat{f}(S) \neq 0\}$.

$s(f)^c \leq \deg(f) \leq s(f)^d$ $\quad$ Conjecture: $\exists d$.

↑ Sensitivity Conjecture

Known: $a(f) \leq s(f) \leq \deg(f)$.

If $f: \{\pm 1\}^n \longrightarrow \mathbb{R}$. $\mathbb{E}(f(x)^2) = \sum_S \hat{f}(S)^2$.

If $f: \{\pm 1\}^n \longrightarrow \{\pm 1\}$ then $\sum_{S \subset [n]} \hat{f}(S)^2 = 1$. (Fourier Spectral distribution)

$p_1 \cdots p_n$ is a probability distribution if $p_i \geq 0$ and $\sum p_i = 1$.

$p_1 = p_2 = \cdots = p_n = \frac{1}{n}$.

$p_1 = 1$. $p_2 = \cdots = p_n = 0$.

$p_1 = p_2 = \cdots = p_r = \frac{1}{r}$. $p_{r+1} = \cdots = p_n = 0$.

Entropy $= \sum p_i \log_2 \frac{1}{p_i} = H(\bar{P})$.

Fourier Entropy Conjecture:

$H(\hat{f}^2) \leq C \cdot as(f)$. where C is some constant.

$H(\hat{f}^2) \leq C \cdot s(f)$.

$as(f) = \sum |S| \hat{f}(S)^2$.

this is defined for $\forall f: \{\pm 1\}^n \longrightarrow \mathbb{R}$ s.t. $\mathbb{E}(f(x)^2) = 1$.

and this conjecture is not true for all non boolean function

Counter example.

$\hat{f}(S) = \frac{1}{\sqrt{n}}$ for $\forall |S| = 1$.

$= 0$ for $\forall |S| \neq 1$.

$as(f) = 1$. $H(\hat{f}^2) = \sum \frac{1}{n} \log_2 n = \log_2 n$.

For non boolean range function we know

$H(\hat{f}^2) \leq O(\log n) \cdot as(f)$

If f is symmetric, FEI-Conjecture is true.

O'donell $\quad S_n$ can also be $S_{i_1} \times \cdots \times S_{i_\ell}$.

not known for graph properties / monotone

Relaxation: $as(f) \rightarrow s(f) \rightarrow degree(f)$.

C-Kulkarni - Lokam - Santhanam.

$H(\hat{f}^2) \leq c \deg(f) \leq c \cdot D(f)$.

$\bar{d}(f) = \mathbb{E}_x [D(f, x)] = \sum \frac{c_i}{2^{c_i}}$

If $C_1, \cdots C_\ell = \mathcal{C}$ are facets $L(\mathcal{C}) = \sum \frac{c_i}{2^{c_i}}$

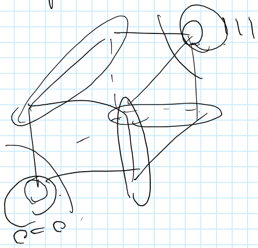
$L(f) = \min_{\mathcal{C}} L(\mathcal{C})$.

Conj $H(\hat{f}^2) \leq L(f)$.

we know $H(\hat{f}^2) \leq \bar{d}(f)$. and $\exists f$. $L(f) < \bar{d}(f)$

Q: Is there more than constant mult gap?

Example: f Not all equal



S-partition. $L(f) \leq \bar{d}(f)$.

$f: \{\pm1\}^k \longrightarrow \{\pm1\}$.

$g: \{\pm1\}^k \longrightarrow \{\pm1\}$. $f \otimes g: \{\pm1\}^{k\ell} \longrightarrow \{\pm1\}$. such that.

$f \otimes g (x_1, \cdots x_{k\ell}) = f(g(x_1, \cdots x_\ell), g(x_{\ell+1}, \cdots x_{2\ell}), \cdots)$

$\bar{d}(f \otimes f) = \overline{2d}(f)$? gap not increased by tensor product.

$\bar{a} = a_1, \cdots a_m$. $R_p(\bar{a}) = (\mathbb{E}\, a_i^p)^{1/p}$.

$As(f) = R_1(s(f, x)) = \sum_S |S| \hat{f}(S)^2 = R_1(|S|) = \mathbb{E}[|S|]$

$f * L(x) = \mathbb{E} f(x+y) L(x) = nf(x) - \sum f(x \oplus e_i)$.

$\mathbb{E}\, \cancel{f*}f(x)^2 = \sum_S \widehat{L * f}(S)^2 = \sum L(S)^2 \hat{f}(S)^2$

$= \sum |S|^2 \hat{f}(S)^2$.