Approximate Uniform Sampling

Wednesday, 1 March 2023 10:08 AM

Def : (ϵ-uniform sample)

Let ω be (random) output of a sampling algorithm for a finite sample space Ω. The sampling algo. generates an ϵ-uniform sample of Ω if for any subset S of Ω

$$\left| Pr(\omega \in S) - \frac{|S|}{|\Omega|} \right| \leq \epsilon$$

FPAUS : Given input x, $\epsilon > 0$ generates ϵ-uniform sample of $\Omega(x)$ and time complexity poly in $\ln \epsilon^{-1}$ and size of input x.

Example Problem :- Independent sets counting

Want to show :- Given an FPAUS for Independent sets, we can construct an FPRAS for counting the number of independent sets.

G has m edges : $e_1, e_2, \ldots e_m$ (arbitrary ordering)

E_i = set of first i edges in E

$G_i = (V, E_i)$

$G = ?$ How to get G_{i-1} from G_i?

$\Omega(G_i)$ = set of independent sets in G_i

$$|\Omega(G)| = \frac{|\Omega(G_m)|}{|\Omega(G_{m-1})|} \times \frac{|\Omega(G_{m-1})|}{|\Omega(G_{m-2})|} \times \ldots \times \frac{|\Omega(G_1)|}{|\Omega(G_0)|} \times |\Omega(G_0)|$$

$\Omega(G_0) = 2^n$ (n is # of vertices. Why?)

Define ratio $r_i = \frac{|\Omega(G_i)|}{|\Omega(G_{i-1})|}$

We develop estimates $\tilde{r}_i$ and estimate $|\Omega(G)|$ as $2^n \prod_{i=1}^{m} \tilde{r}_i$

To get error in our estimate define

$$R = \prod_{i=1}^{m} \frac{\tilde{r}_i}{r_i}$$

We want $Pr(|R-1| < \epsilon) \geq 1-\delta$

Claim :- If $\forall i$, $1 \leq i \leq m$, $\tilde{r}_i$ is an $(\epsilon/2m, \delta/m)$-approximation for r_i, then

$$Pr(|R-1| \leq \epsilon) \geq 1-\delta$$

Proof :- For each $i \in [1, m]$

$$Pr\left(|\tilde{r}_i - r_i| \leq \frac{\epsilon}{2m} r_i\right) \geq 1 - \delta/m$$

i.e. $Pr\left(|\tilde{r}_i - r_i| > \frac{\epsilon}{2m} r_i\right) < \delta/m$

Apply union bound

$$|\tilde{r}_i - r_i| \leq \left(\frac{\epsilon}{2m}\right) r_i$$

$\boxed{\forall i}$ w.p atleast $1-\delta$

i.e. $1 - \frac{\epsilon}{2m} \leq \frac{\tilde{r}_i}{r_i} \leq 1 + \epsilon/2m$

$\forall i$ w.p atleast $1-\delta$

$$(1-\epsilon) \leq \left(1 - \frac{\epsilon}{2m}\right)^m \leq \prod_{i=1}^{m} \frac{\tilde{r}_i}{r_i} \leq \left(1 + \frac{\epsilon}{2m}\right)^m \leq (1+\epsilon)$$

Require $(\epsilon/2m, \delta/m)$-app. for r_i

Algo :-

I/p : Graphs G_{i-1}, G_i

O/p : $\tilde{r}_i$

$X \leftarrow 0$

Repeat for $M = \left\lceil \frac{129 m^2 \ln(2m/\delta)}{\epsilon^2} \right\rceil$ trials

a) Generate $(\epsilon/6m)$-uniform sample of $\Omega(G_{i-1})$

b) If its independent set in G_i, $X \leftarrow X+1$

Return $\tilde{r}_i \leftarrow X/M$

Claim :- For $m \geq 1$, $0 < \epsilon \leq 1$, Algo gives $(\epsilon/2m, \delta/m)$-app. for r_i

Proof :- Show r_i not to small. Why?

G_i & G_{i-1} differ in edge (u, v)

$\Omega(G_i) \subseteq \Omega(G_{i-1})$ Why?

An independent set in $\Omega(G_{i-1}) - \Omega(G_i)$ contains both u and v. Why?

For each $I \in \Omega(G_{i-1}) - \Omega(G_i)$

let $I - \{v\} \in \Omega(G_i)$

$$|\Omega(G_{i-1}) - \Omega(G_i)| \leq |\Omega(G_i)|$$

Why??

($\because$ $I' \in \Omega(G_i)$ is associated with no more than one $I' \cup \{v\} \in \Omega(G_{i-1}) - \Omega(G_i)$)

$$\therefore r_i = \frac{|\Omega(G_i)|}{|\Omega(G_{i-1})|} = \frac{|\Omega(G_i)|}{(|\Omega(G_i)| + |\Omega(G_{i-1}) - \Omega(G_i)|)} \geq 1/2$$

For M samples

Let $X_k = 1$ if k^{th} sample is in $\Omega(G_i)$

0 o/w

Now

$$\left| Pr(X_k = 1) - \frac{|\Omega(G_i)|}{|\Omega(G_{i-1})|} \right| \leq \frac{\epsilon}{6m}$$

Why??

$$\Rightarrow \left| E[X_k] - \frac{|\Omega(G_i)|}{|\Omega(G_{i-1})|} \right| \leq \frac{\epsilon}{6m}$$

$$\Rightarrow \left| E\left[\frac{\sum_{i=1}^{M} X_k}{M}\right] - \frac{|\Omega(G_i)|}{|\Omega(G_{i-1})|} \right| \leq \frac{\epsilon}{6m}$$

$$\Rightarrow |E[\tilde{r}_i] - r_i| \leq \frac{\epsilon}{6m}$$

Now $r_i \geq 1/2$

$$E[\tilde{r}_i] \geq r_i - \epsilon/6m \geq \frac{1}{2} - \frac{\epsilon}{6m} \geq 1/3$$

Chernoff Bound gives

if $M \geq \frac{3 \ln(2m/\delta)}{(\epsilon/12m)^2 (1/3)}$

then

$$Pr\left(\left|\frac{\tilde{r}_i}{E[\tilde{r}_i]} - 1\right| \geq \frac{\epsilon}{12m}\right) = Pr\left(|\tilde{r}_i - E[\tilde{r}_i]| \geq \frac{\epsilon}{12m} E[\tilde{r}_i]\right) \leq \delta/m$$

So w.p $1 - \delta/m$

$$1 - \frac{\epsilon}{12m} \leq \frac{\tilde{r}_i}{E[\tilde{r}_i]} \leq 1 + \frac{\epsilon}{12m} \quad \text{— (A)}$$

Also $|E[\tilde{r}_i] - r_i| \leq \epsilon/6m$,

$$1 - \frac{\epsilon}{6m r_i} \leq \frac{E[\tilde{r}_i]}{r_i} \leq 1 + \frac{\epsilon}{6m r_i}$$

Finally $r_i \geq 1/2$

$$\Rightarrow 1 - \frac{\epsilon}{3m} \leq \frac{E[\tilde{r}_i]}{r_i} \leq 1 + \frac{\epsilon}{3m} \quad \text{— (B)}$$

From (A) & (B)

$$1 - \frac{\epsilon}{2m} \leq \left(1 - \frac{\epsilon}{3m}\right)\left(1 - \frac{\epsilon}{12m}\right) \leq \frac{\tilde{r}_i}{r_i} \leq \left(1 + \epsilon/3m\right)\left(1 + \frac{\epsilon}{12m}\right) \leq (1 + \epsilon/2m)$$

Hence Proved.