Monte Carlo Methods

collection of tools for estimating values thru sampling & estimation

(ε, δ) approximation

A randomized alg gives an (ε,δ) approx for value V of the output X of alg satisfies

$$\Pr(|X-V| > \varepsilon|V|) \leq \delta$$

Example:

Sample indep random vars whose mean is quantity you want

Let $X_1, \ldots, X_m$ iid. indicator w/ $\mu = E(X_i)$

If $m \geq \dfrac{3\ln\frac{2}{\delta}}{\varepsilon^2 \mu}$ then from Chernoff bounds

$$\Pr\left(\left|\frac{1}{m}\sum X_i - \mu\right| \geq \varepsilon\mu\right) \leq \delta$$

DNF counting

want to know # satisfying assignments

$$(\bar{x}_1 \wedge x_2) \vee (x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge \bar{x}_3 \wedge x_4) \vee (x_3 \wedge \bar{x}_4)$$

Observation: satisfying such a formula is easy

counting # satisfying assignments is hard

if could do this, could solve 3SAT

CNF formula $f \xrightarrow{\text{De Morgan}} \bar{f}$ DNF

$\downarrow$

if $2^n \to f$ unsat $\leftarrow$ count # of solns

else sat

Problem actually #P complete $\longrightarrow$ strong intractability

#P counting analog of NP: problems of form

compute $f(x)$ where $f(x)$

solns to problem in NP

Approx counting of # solns of DNF formula φ

Obvious approach: sample random assignments,

estimate (fraction satisfying) 2^n

From (ε, δ) calculation need $m = \Omega\left(\frac{1}{\mu}\right)$

μ could be exponentially small e.g. $\frac{n^2}{2^n}$

$\varphi = C_1 \vee C_2 \vee \dots \vee C_t$

SC_i : set of assignments that satisfy clause i

want to approx $\left| \bigcup_i SC_i \right|$

Let $U = \{(i, \vec{x}) \mid 1 \le i \le t,\ \vec{x} \in SC_i\}$

$$|U| = \sum_{i=1}^{t} |SC_i|$$

$$X = \{(i, \vec{x}) \mid \vec{x} \in SC_i,\ \vec{x} \notin SC_j\ \ 1 \le j \le i-1\}$$

estimate $|X|$ by approximating $\frac{|X|}{|U|}$. To do this

sample elts $(i,\vec{x}) \in U$ and determine if $(i,\vec{x}) \in S$

Key observation: $\frac{|S|}{|U|} \geq \frac{1}{t}$, so $\frac{3t \ln(\frac{2}{\delta})}{\varepsilon^2}$ samples suffice

for (ε,δ) approx

How to pick random sample from U?

pick i w/ prob $\frac{|SC_i|}{\sum_j |SC_j|}$ set vars outside C_i u.a.r.

$\Pr((i,\vec{x}) \text{ selected}) = \frac{|SC_i|}{|U|} \cdot \frac{1}{|SC_i|} = \frac{1}{|U|}$

DNF counting illustrates fundamental connection

sampling $\longleftrightarrow$ counting

very often use "self-reducibility" from problem of size $i \longrightarrow$ size $i-1$

approx sampling $\longrightarrow$ approx counting

illustrate reduction for counting ISs

Counting Indep Sets

$G=(V,E)$ $\quad e_1,\dots,e_m$ arbitrary ordering of edges

$E_i = \{e_1,\dots,e_i\}$ $\quad G_i = (V,E_i)$ $\quad G_m = G$ $\quad G_0 = (V,\emptyset)$

$\Omega(G_i)$: set of ISs in G_i

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

to estimate $|\Omega(G)|$ need good estimate for

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

Alg for estimating r_i, given alg for generating random samples

1. $X := 0$
2. repeat for $M = \lceil \frac{3}{(\frac{\varepsilon}{2m})^2} \ln\left(\frac{2}{\delta/m}\right)$ trials
   - generate unif sample from $\Omega(G_{i-1})$
   - if sample is indep set in G_i, $X{+}{+}$
3. return $\tilde{r}_i := \frac{X}{M}$

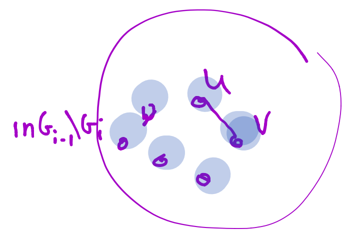
$\tilde{r}_i$ is $\left(\frac{\varepsilon}{2m}, \frac{\delta}{m}\right)$ approx for r_i

Fact: $r_i \geq \frac{1}{2}$

$\exists$ one edge (u,v) in $G_i \setminus G_{i-1}$

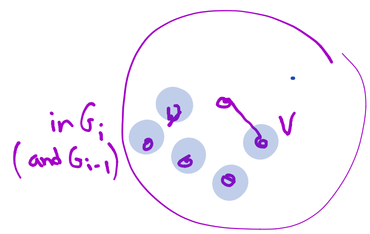
so any I in G_{i-1} that is not indep in G_i

contains both u & v



$$I \longrightarrow I \setminus v \quad \text{in } G_i \text{ \& } G_{i-1}$$

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



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

Our estimate: $2^n \prod_{i=1}^{m} \tilde{r}_i$ Truth: $2^n \prod_{i=1}^{m} r_i$

Claim: $\Pr\left(\left| \prod_{i=1}^{m} \frac{\tilde{r}_i}{r_i} - 1 \right| \leq \varepsilon \right) \geq 1 - \delta$

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

$$r_i\left(1 - \frac{\varepsilon}{2m}\right) \leq \tilde{r}_i \leq r_i\left(1 + \frac{\varepsilon}{2m}\right) \quad \text{w.p.} \geq 1 - \frac{\delta}{m}$$

$$1 - \frac{\varepsilon}{2m} \leq \frac{\tilde{r}_i}{r_i} \leq \left(1 + \frac{\varepsilon}{2m}\right) \qquad \text{"}$$

$$1-\varepsilon \le \left(1-\frac{\varepsilon}{2m}\right)^m \le \prod \frac{\tilde{r}_i}{r_i} \le \left(1+\frac{\varepsilon}{2m}\right)^m \le 1+\varepsilon \qquad \text{w.p.} \ge 1-\delta$$

we only need to figure out how to generate uniform sample from $\Omega(G_{i-1})$ $\forall i$

actually almost uniform sample suffices

How to sample elts from a universe Ω according to some distn π?

Cool idea: design MC whose state space is Ω that has stationary distn π

- simulate MC until it "mixes"
- use state at that time as sample

2 key questions:

1. how to design chain w/ right π?
2. how to bound mixing time?

Example: sampling indep sets uniformly from $G=(V,E)$

states: indep sets

X_t some indep set

choose vertex v u.a.r. from V

if $v \in X_t$ then $X_{t+1} = X_t \setminus v$

if $v \notin X_t$ & can be added w/o violating independence

then $X_{t+1} = X_t \cup v$

otherwise $X_{t+1} = X_t$

- irreducible
- if $\exists$ edge, then aperiodic $(u,v) \to P_{v,v} > 0$
- stationary distn uniform (chain doubly stochastic)
  $P_{I,I'}$ either $\frac{1}{n}$ or 0

general technique, given Ω and a connected graph on Ω to define transition probs so that will have stationary distn π

Metropolis Algorithm

Input: Ω, connected $G=(\Omega,E)$, π s.t. $\sum_{i\in\Omega}\pi_i=1$

Let Δ = max degree in graph

$$P_{xy}=\begin{cases}\frac{1}{2\Delta}\min\left(1,\frac{\pi_y}{\pi_x}\right) & x\neq y \quad y\in N(x)\\ 0 & x\neq y \quad y\notin N(x)\\ 1-\sum_{y\neq x}P_{xy} & x=y\end{cases}$$

nice fact: only depends on ratios

Claim

$\pi_x P_{xy}=\pi_y P_{yx} \quad \forall x\neq y \implies \pi$ is stationary distn

assume wlog $\pi_x\le\pi_y$

$$\pi_x\left(\frac{1}{2\Delta}\right)=\pi_y\left(\frac{1}{2\Delta}\frac{\pi_x}{\pi_y}\right)$$

Example:

suppose want to sample indep sets according to $\pi(I)=\frac{\lambda^{|I|}}{Z}$ $\qquad Z=\sum_{I'}\lambda^{|I'|}$

$P_{I,I'}=\frac{1}{2n}\min(1,\lambda)$ $\qquad$ (same graph as above) important connections to hardcore lattice gas model in statistical physics

Bounding mixing time

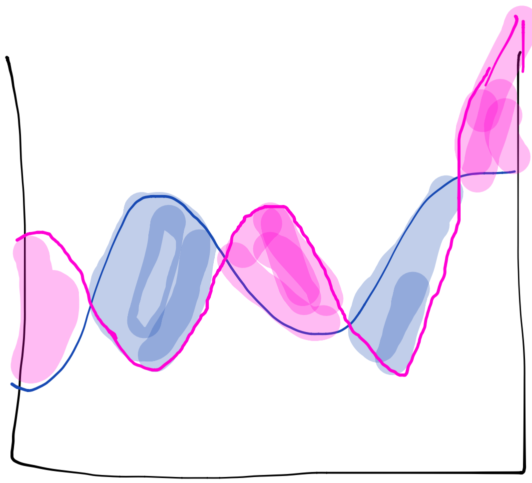
1. spectral gap, conductance, expansion ...

2. coupling

Coupling

Total variation distance between 2 distr's on same sample space Ω

$$\|D_1 - D_2\|_{TV} = \frac{1}{2}\sum_{x\in\Omega}|D_1(x) - D_2(x)| = \max_{A\subseteq\Omega}|D_1(A) - D_2(A)|$$



$\|D_1 - D_2\|_{TV}$ = pink area = blue area

Common defn of mixing time $\tau(\varepsilon)$

$$\tau(\varepsilon) = \min\{t \mid \|P^t, \pi\|_{TV} \leq \varepsilon\}$$

say MC is rapidly mixing if $\tau(\varepsilon)$ poly in $\log|\Omega|$ & $\log(\frac{1}{\varepsilon})$

(we know this is related to spectral gap)

Coupling: simple & elegant approach to bounding mixing times

Given a MC on Ω, a coupling is a MC on $\Omega \times \Omega$ defining stochastic process (X_t, Y_t) s.t.

1. each of X_t & Y_t in isolation is faithful copy of MC
2. If $X_t = Y_t$ then $X_{t+1} = Y_{t+1}$

Coupling Lemma Let $Z_t = (X_t, Y_t)$ be a coupling

Suppose $\exists\, T$ s.t. $\forall x, y$

$$\Pr(X_T \neq Y_T \mid X_0 = x, Y_0 = y) \leq \varepsilon$$

Then $\tau(\varepsilon) \leq T$.

Proof: Consider coupling w/ Y_0 chosen according to π, x_0 arbitrary

$$\forall A \subseteq \Omega \quad \Pr(X_T \in A) \geq \Pr(X_T = Y_T \cap Y_T \in A)$$

$$= 1 - \Pr(X_T \neq Y_T \cup Y_T \notin A)$$

$$\geq 1 - \Pr(Y_T \notin A) - \Pr(X_T \neq Y_T)$$

$$= \Pr(Y_T \in A) - \Pr(X_T \neq Y_T)$$

$$\downarrow Y_0 \sim \pi \qquad \downarrow \text{assumption of lemma}$$

$$= \pi_A - \varepsilon$$

Similarly $\Pr(X_T \notin A) \geq \pi_{\Omega \setminus A} - \varepsilon$

$$\Pr(X_T \in A) \leq \pi_A + \varepsilon$$

$\boxed{\text{Examples:}}$

① Random walk on hypercube $N = 2^n$ nodes

at each step choose random coordinate i, random bit $b \in \{0,1\}$

change i^{th} bit to b

coupling: at each step choose same i & b

one RW starts at π

how many steps till walks meet?

(2) indep sets of fixed size k

- here distances may $\uparrow$...

Chain: choose vertex $v \in X_t$ u.a.r. & a vertex $w \in V$ u.a.r.

if $w \notin X_t$ & $X_t - v + w$ indep, $X_{t+1} := X_t - v + w$

else $X_{t+1} = X_t$

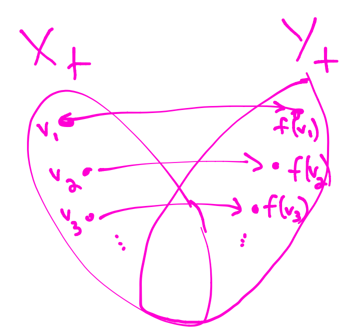
n # vertices

Δ max deg

irreducible, aperiodic, π uniform

Claim: MC rapidly mixing if $k \leq \frac{n}{(3\Delta+3)}$

coupling uses arbitrary bijection f between vertices of $X_t - Y_t$ & $Y_t - X_t$



$\underline{X_t}$

choose $v \in X_t$ u.a.r.

$w \in V$ u.a.r.

$\underline{Y_t}$ starts in π

$v \in Y_t$ use (v, w)

$v \notin Y_t$ use $(f(v), w)$

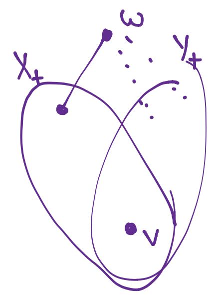
both behave like original chain

$d_t = |X_t - Y_t|$ can change by at most 1 each step

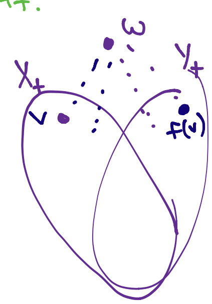
chains meet when $d_t = 0$

claim: d_t has negative drift so this doesn't take too long

$d_{t+1} = d_t + 1$



$$\Pr(d_{t+1} = d_t + 1 \mid d_t > 0) \le \Big(\underbrace{\frac{k - d_t}{k}}_{v \text{ in } X_t \cap Y_t}\Big)\Big(\underbrace{\frac{2 d_t (\Delta+1)}{n}}_{\substack{w \in \text{ neighborhood of one or other} \\ \text{or symm diff.}}}\Big)$$



$d_{t+1} = d_t - 1$

$$\Pr(d_{t+1} = d_t - 1 \mid d_t > 0) \ge \underbrace{\frac{d_t}{k}}_{v \in X_t \setminus Y_t} \cdot \underbrace{\frac{n - (k+d-2)(\Delta+1)}{n}}_{\substack{w \text{ not } \in X_t \cup Y_t \\ \text{or neighborhood of either}}}$$

$|X_t \cup Y_t| = k+d$

$$E(d_{t+1} \mid d_t) = \Pr(d_{t+1} = d_t + 1)(d_t + 1) + \Pr(d_{t+1} = d_t - 1)(d_t - 1)$$

$$\le d_t \underbrace{\left(1 - \frac{n - (3k-3)(\Delta+1)}{kn}\right)}_{\alpha < 1}$$

by Induction

$$E(d_t) \le d_0 \alpha^t \longrightarrow 0$$

$$\Pr(d_t \geq 1) \le E(d_t) \longrightarrow 0$$