Consider the following Markov chain:

$\Omega = \{0,1\}^n$ = n-bit vectors

From $X_t \in \Omega$,

1. Choose i uniformly at random (u.a.r.) from $\{1,\ldots,n\}$
2. Choose b u.a.r. from $\{0,1\}$
3. Set $X_{t+1}(j) = \begin{cases} X_t(j) & \text{for } j \neq i \\ b & \text{for } j = i \end{cases}$

This is an ergodic MC:

irreducible: from any $X \in \Omega$ can reach $\bar{0}$ (& can go back)

aperiodic: with prob. $\frac{1}{2}$, $X_{t+1} = X_t$ since can choose the same bit $X_t(i)$.

Note, P is symmetric, $P(x,y) = P(y,x) = \frac{1}{2n}$ if $|x \oplus y| = 1$.

$\Rightarrow$ unique stationary is $\pi = \text{Uniform}(\Omega)$.

How fast does it converge?

We know that for $\lim_{t\to\infty} \Pr(X_t = j) = \pi(j)$.

$$\lim_{t\to\infty} X_t \to \pi$$

Recall, for μ & π on Ω,

$$d_{TV}(\mu, \pi) = \frac{1}{2}\sum_{x\in\Omega} |\mu(x) - \pi(x)|$$

For $X_0 \in \Omega$, $\epsilon > 0$

$$T_{mix}^{X_0}(\epsilon) = \min\{t : d_{TV}(X_t, \pi) \leq \epsilon\}$$

= time it to get within total variation dist. $\leq \epsilon$ of π

$$T_{mix}(\epsilon) = \max_{X_0} T_{mix}^{X_0}(\epsilon) = \text{time from worst } X_0$$

$$T_{mix} = T_{mix}(1/4)$$

We'll show that: $T_{mix}(\epsilon) \leq T_{mix} \cdot \log(1/\epsilon)$

We'll soon see: $T_{mix} = O(n \log n)$

& $T_{mix}(\epsilon) = O(n \log(n/\epsilon))$.

Note for any $c < 1/2$, $T_{mix}(\epsilon) = O(T_{mix}(c) \log(1/\epsilon))$

this constant depends on c.

Top-to-random shuffle:

For cards $\{1, \dots, n\}$, let $\Omega = S_n$ = all $n!$ permutations of $\{1, \dots, n\}$

For $\sigma \in \Omega$, let $\sigma(1)$ = top card
$\vdots$
$\sigma(n)$ = bottom card

Thus for $\sigma = (5, 3, 2, 4, 1)$ think of
$$\begin{matrix} 5 \\ 3 \\ 2 \\ 4 \\ 1 \end{matrix}$$

The MC takes the top card & chooses a random position for the top card, & repeats.

Thus, for $\sigma = (5, 3, 2, 4, 1)$ if we choose position 3 then we move:

$$\begin{matrix} 5 \\ 3 \\ 2 \\ 4 \\ 1 \end{matrix} \longrightarrow \begin{matrix} 3 \\ 2 \\ 5 \\ 4 \\ 1 \end{matrix}$$

For $X_t \in \Omega$,

1. Choose i u.a.r. from $\{1, \dots, n\}$
2. Set $X_{t+1}(j) = \begin{cases} X_t(j+1) & \text{for } j < i \\ X_t(1) & \text{for } j = i \\ X_t(j) & \text{for } j > i \end{cases}$

Ergodic:

irreducible: for $\sigma \rightsquigarrow \tau$,
can get $\tau(n)$ in position, then induct...

aperiodic: can send top-card to itself (position 1)
So all states have a self-loop
(thus it's aperiodic, see HW1 Problem 1)

What's the unique stationary distribution?

it's not reversible (might have $\sigma \to \tau$ but not $\tau \to \sigma$)
So can't easily check.

Let $\pi = \text{uniform}(\Omega)$.

Note, for $\sigma \in \Omega$, there are n states with transition into σ
& each has prob. $1/n$

thus $(\pi P)(\sigma) = \sum_{\tau: \tau \to \sigma} \pi(\tau) P(\tau, \sigma) = \frac{n}{n!} \times \frac{1}{n} = \frac{1}{n!} = \pi(\sigma)$

& we'll see $T_{mix} = O(n \log n)$.

Random matchings:

$G=(V,E)$, let $\mathcal{M}$ = collection of all matchings of G.
let $\pi = \text{uniform}(\mathcal{M})$.
How to sample from π?

- choose random edge $e \in E$.
- modify current matching X_t at e & move there if it's still a matching, otherwise stay at X_t.

From $X_t \in \Omega$,

1. Choose $e=(u,v)$ u.a.r. from E.
2. Let $X' = \begin{cases} X_t \cup e & \text{if } e \notin X_t \\ X_t \setminus e & \text{if } e \in X_t \end{cases}$
3. If $X' \in \mathcal{M}$,

   then $X_{t+1} = \begin{cases} X' & \text{w.prob. } 1/2 \\ X_t & \text{w.prob. } 1/2 \end{cases}$

   else $X_{t+1} = X_t$.

Ergodic:

irreducible – for any $X \in M$, can go to $\emptyset$ & then can go to any $Y \in M$.

aperiodic – self-loop on every state.

What's the stationary distribution?

P is symmetric,

if $X = Y \cup e$

then $P(X,Y) = \frac{1}{2m} = P(Y,X)$

choose edge e & move w. prob. ½.

thus, $\pi = \text{uniform}(\Omega)$.

What's the mixing time?

We'll prove that it's poly(n) for any G.

More general problem:

Positive weights $\omega: \mathcal{M} \to \mathbb{R}_+$

<u>Goal</u>: Sample from distribution $\mu \propto \omega$

i.e., for $X \in \mathcal{M}$, $\mu(X) = \dfrac{\omega(X)}{Z}$

where $Z = \sum_{Y \in \mathcal{M}} \omega(Y)$

Z is called the partition function

μ is Gibbs or Boltzmann distribution.

Can we modify the previous chain?

Note, if $\omega(X) = 1$ for all $X \in \mathcal{M}$

then $\mu = \text{uniform}(\mathcal{M})$

& $Z = |\mathcal{M}|$.

Metropolis filter: (also called Metropolis-Hastings chain)

from $X_t \in \Omega$, propose new state X'

move with prob. $\min\left\{1, \dfrac{\omega(X')}{\omega(X_t)}\right\}$

This is reversible wrt μ:

Suppose without loss of generality,

$$\omega(X) \leq \omega(Y)$$

& look at $X \to Y$ & $Y \to X$

$$P(X,Y) = 1 \quad \& \quad P(Y,X) = \frac{\omega(X)}{\omega(Y)}$$

then:

$$\mu(X)P(X,Y) = \frac{\omega(X)}{Z} \times 1 = \frac{\omega(X)}{Z}$$

$$\mu(Y)P(Y,X) = \frac{\omega(\cancel{Y})}{Z} \times \frac{\omega(X)}{\omega(\cancel{Y})} = \frac{\omega(X)}{Z}$$

Here is the chain:

From $X_t \in \Omega$,

1. Choose $e=(u,v)$ u.a.r. from E
2. Let $X' = \begin{cases} X_t \cup e & \text{if } e \notin X_t \\ X_t \setminus e & \text{if } e \in X_t \end{cases}$
3. If $X' \in \mathcal{M}$,

   then $X_{t+1} = \begin{cases} X' & \text{w.p. } \min\left\{1, \frac{\omega(X')}{\omega(X_t)}\right\} \\ X_t & \text{o/w} \end{cases}$

   else $X_{t+1} = X_t$.

Ergodic:

irreducible – all $X \in \mathcal{M}$ can go to $\emptyset$ & it's reversible.

aperiodic – for maximal matching $X \in \mathcal{M}$, then $P(X,X) > 0$.

What's π?

reversible wrt $\mu(X) \propto \omega(X)$.

(assuming $E \supsetneq X$)

Original paper:

Metropolis, Rosenbluth, Rosenbluth, Teller, Teller '53

J. Chem. Phys.

later generalized by Hastings '70.

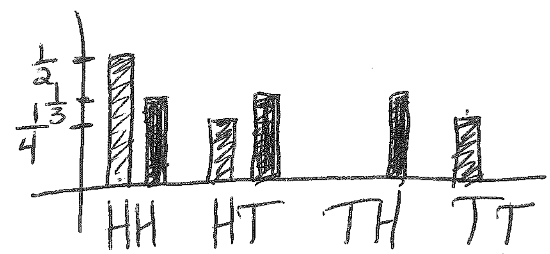
but most of work done for this paper

by husband/wife Rosenbluth & Rosenbluth.

Total variation distance:

Example: $\Omega = \{HH, HT, TH, TT\}$

$\mu = (.5, .25, 0, .25)$, $\nu = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0\right)$



$$\text{area b/w } \mu \,\&\, \nu = \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{4} - 0\right) = \frac{1}{6} + \frac{1}{4} = \frac{5}{12}$$

$$= \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{3} - 0\right) = \frac{1}{12} + \frac{1}{3} = \frac{5}{12}$$

$$= \frac{1}{2}\left[\left|\frac{1}{2} - \frac{1}{3}\right| + \left|\frac{1}{4} - \frac{1}{3}\right| + \left|\frac{1}{3}\right| + \left|\frac{1}{4}\right|\right] = \frac{5}{12}$$

$$d_{TV}(\mu, \nu) = \frac{1}{2} \sum_{X \in \Omega} |\mu(x) - \nu(X)|$$

$\frac{1}{2}$(area between) $\longrightarrow$

$$= \sup_{S \subset \Omega} \mu(S) - \nu(S)$$

$\cancel{S = \{X : \mu(x) > \nu(x)\}}$

$S = \{X : \mu(X) > \nu(X)\}$

μ over ν $\longrightarrow$

$$= \sup_{T \subset \Omega} \nu(T) - \mu(T)$$

ν over μ $\longrightarrow$