$\boxed{\text{4 May 2022}}$ Mixing time and coupling

Recall Metropolis–Hastings... (MCMC)

Finite set $\mathcal{X}$ (think: $\mathcal{X}$ exponentially large
e.g. $\mathcal{X} = \{\text{functions } V(G) \to [q]\}$)

Distribution $\pi(x) = \dfrac{w(x)}{\sum_{y \in \mathcal{X}} w(y)}$ that we wish
to sample from.

<u>Strategy</u>: Run Markov chain with transition matrix
P such that $\pi P = \pi$ until it "mixes",
meaning state distrib gets close to π.
Then output the current state; it will be
an approximate sample from π.

<u>Metropolis–Hastings</u>: given a different Markov chain K
on $\mathcal{X}$ with $K_{xy} = K_{yx}$. ($\Rightarrow$ uniform distrib
is stationary for K). Want to modify K into
a Markov chain P such that

(a) $\pi P = \pi$

(b) Easy to simulate one state transition
of P.

$$P_{xy} = K_{xy} \cdot \frac{\min\{w(x), w(y)\}}{w(x)}.$$

P: Given current state x

1. Random sample "proposed next state" y with prob K_{xy}.

2. Transition to y with prob $\frac{\min\{w(x), w(y)\}}{w(x)}$.

Assumption is that both K and P will be implemented by subroutines that input state x and output state y with probability K_{xy} or P_{xy} respectively.

E.g. Metropolis-Hastings for sampling random proper q-coloring of graph G.

$$\mathbb{X} = \{\text{functions } V(G) \to [q]\}$$

$$w(x) = \begin{cases} 1 & \text{if } x(u) \neq x(v) \quad \forall \text{ edge } (u,v) \\ \emptyset & \text{if } x(u) = x(v) \quad \text{for some edge } (u,v) \end{cases}$$

For K, we will use a very simple subroutine.

In state x, pick a unif random vertex $v \in V(G)$ a unif random color $c \in [q]$ and modify the color of v to equal c.

i.e. output state $y: V(G) \to [q]$ s.t.

$$y(u) = \begin{cases} c & \text{if } u = v \\ x(u) & \text{if } u \neq v. \end{cases}$$

Def. Hamming distance $d(x,y)$ is

$$d(x,y) = \#\{v \mid x(v) \neq y(v)\}$$

$$K_{xy} = \begin{cases} 0 & \text{if } d(x,y) > 1 \\ \frac{1}{n \cdot q} & \text{if } d(x,y) = 1 \\ \frac{1}{q} & \text{if } d(x,y) = 0 \end{cases}$$

$K_{xy} = K_{yx}$ b/c Hamming dist is a symm function.

Metropolis-Hastings for this particular K says.

Given current state x. (Assume x a proper coloring.)

1. Sample y with probability K_{xy}. $\Rightarrow w(x) = 1$

Pick random $v \in V(G)$, $c \in [q]$,

try setting $y = \{x$ with v recolored as $c\}$.

2. Transition to y with probability $\frac{\min\{w(x), w(y)\}}{w(x)}$,

$$\text{Prob(trans)} = \min\{1, w(y)\} = w(y)$$
$$= \begin{cases} 1 & \text{if } y \text{ is proper coloring} \\ \emptyset & \text{if not.} \end{cases}$$

Test if color c matches the color of a neighbor of v.

If so, no transition. (Output x)

If not, output y.

After how many steps can we terminate the Markov chain, safe in the knowledge that its marginal distribution is near stationary?

Def. If π, π' are prob. distribs, their total variation distance is

$$\|\pi - \pi'\|_{TV} = \max_{S \subseteq \mathcal{X}} \{ \pi(S) - \pi'(S) \}$$

$$= \frac{1}{2} \| \pi - \pi' \|_1$$

For us, "near stationary" will mean ε-close to the stationary distrib in total variation distance.

Def. A coupling of two Markov chains $\mathcal{M}_1, \mathcal{M}_2$ is a probability distrib. over sequences of pairs (X_t, X'_t) such that the marginal distrib of $X_0, X_1, X_2, \dots$ is a random sample from $\mathcal{M}_1$

$X'_0, X'_1, X'_2, \dots$ — — — — $\mathcal{M}_2$.

Lemma. If $\{(X_t, X'_t)\}$ is a coupling of two Markov chains with same state set $\mathcal{X}$, same transition matrix P and with X_0

drawn from π, stationary distrib of P then the marginal distrib of X_t, called π_t, satisfies

$$\| \pi_t - \pi \|_{TV} \leq \Pr(X_t \neq X_t').$$

Proof. Since X_0' is drawn from π and $X_0', X_1', \ldots$ follows transition matrix P whose stationary distrib is π, it means $X_t' \sim \pi$ for all $t > 0$. (Induct on t.)

By definition of TV distance, to prove inequality in Lemma we must show

$$\pi_t(S) - \pi(S) \leq \Pr(X_t \neq X_t') \quad \forall S \subseteq \mathcal{X}.$$

$$\| $$

$$\Pr(X_t \in S) - \Pr(X_t' \in S)$$

If $X_t \in S$ then either $X_t' \in S$ or $X_t \neq X_t'$.

$\Downarrow$ (union bd)

$$\Pr(X_t \in S) \leq \Pr(X_t' \in S) + \Pr(X_t \neq X_t').$$

$$\Pr(X_t \in S) - \Pr(X_t' \in S) \leq \Pr(X_t \neq X_t')$$