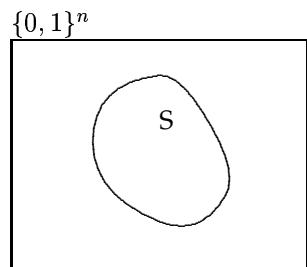


The typical situation is that we are sampling a large (superpolynomial) set (here labelled as S) that, while large is exponentially smaller than the total set (here $\{0, 1\}^n$)



Strategy: construct an artificial random walk on the set that quickly converges to a random point.

Generically, the Markov chains that we define will be on partially directed graphs with vertices and edges (V, E) .

Typically,

$$p(u \rightarrow v|u) = p_{uv} = \begin{cases} \frac{1}{d_u+1} & \text{if } v = u \text{ or } \exists \text{ edge } u \rightarrow v \\ 0 & \text{otherwise} \end{cases}$$

or

$$p_{uv} = \begin{cases} \frac{1}{2d_u} & \text{if } \exists \text{ edge } u \rightarrow v \\ \frac{1}{2} & \text{if } u = v \\ 0 & \text{otherwise} \end{cases}$$

Cover time of a graph, $\text{COV}(M)$, is the maximum over vertices u of the expected time to hit all vertices, starting from u .

Mixing time of a process, $\text{MIX}(M)$ is

$$\max_{\text{initial distributions } p \text{ on } v} \min_t [|pM^t - \pi| < \frac{1}{2}]$$

where π is the unique stationary distribution of M .

The mixing time is defined in terms of the L_1 norm, $|p - q| = \sum_v |p(v) - q(v)|$, also called the "variation distance." (Often you will see the variation distance defined to be half of this.)

It would also make sense to define mixing time in terms of convergence in other metrics, but this is what is standard. The difference in the rate of mixing in different measures is usually not large.

Note that the covering time is at least linear in the number of nodes in the graph.

Simple Example: The undirected, fully connected graph

Using the first walk rule, we have mixing in one step.

Using the second walk rule, we have mixing in 1 or 2 steps depending on the variation permitted for mixing (the value $\frac{1}{2}$ is somewhat arbitrary).

Premise: to sample from a set, we hope to create a well connected graph whose vertices are the elements of the set (e.g., the set of perfect matchings) such that mixing time for the random walk over the graph will be small. In particular, we will construct cases in which the number of states, and the cover time, will be exponential in our complexity parameter n , but the mixing time will be polynomial.

Generally $MIX(M) \ll COV(M)$ when M is "well connected." We want to mix in time $\text{poly}(\log c^n) = \text{poly}(n)$. A clear lower bound for walks on undirected graphs is: mixing time $\geq \frac{1}{2} \text{Diameter}(\text{graph})$. (The diameter of a graph is the longest shortest path between nodes.) (Fix a pair of furthest nodes. Their neighborhoods of radius $\frac{1}{2} \text{Diameter}(\text{graph})$ are disjoint, so one of them contains at most half the stationary distribution on the graph.)

In general, there are several ways to bound mixing time:

1. Coupling Method
2. Conductance Method
3. Eigenvalue Method

As a warm-up, consider a random walk on a connected undirected graph, such that $\forall u p_{uu} = \frac{1}{2}$. Show that the unique stationary distribution is that in which π_v is proportional to d_v , the degree of v .

1. Existence: let d be the vector (d_v) . Use $w \sim v$ to mean w is a neighbor of v .

$$(d * M)_v = \frac{d_v}{2} + \sum_{w \sim v} d_w \frac{1}{2d_w} = d_v$$

2. Uniqueness: Let p be arbitrary. Show

$$\exists k : |pM^k - \pi| < |p - \pi|$$

Define p' :

$$p' = pM$$

so that

$$p'_v = \frac{1}{2}p_v + \sum_{w \sim v} p_w \frac{1}{2d_w}$$

Define π' similarly as

$$\pi' = \pi M$$

By definition

$$\pi_v = \pi'_v = \frac{1}{2}\pi_v + \sum_{w \sim v} \pi_w \frac{1}{2d_w}$$

Subtracting the two we get

$$p'_v - \pi'_v = \frac{1}{2}(p_v - \pi_v) + \sum_{w \sim v} \frac{1}{2d_w}(p_w - \pi_w)$$

Taking the L_1 norm of the vector difference we see first that it is nonincreasing:

$$\begin{aligned} |p' - \pi'| &= \sum_v |p'_v - \pi'_v| \\ &= \sum_v \left| \frac{1}{2}(p_v - \pi_v) + \sum_{w \sim v} \frac{1}{2d_w}(p_w - \pi_w) \right| \\ &\leq \frac{1}{2} \sum_v \left[|p_v - \pi_v| + \sum_{w \sim v} \frac{1}{d_w} |p_w - \pi_w| \right] \\ &= \sum_v |p_v - \pi_v| = |p - \pi| \end{aligned}$$

Let

$$\begin{aligned} V^+ &= \{v : p_v > \pi_v\} \\ V^- &= \{v : p_v < \pi_v\} \end{aligned}$$

The inequality $|p' - \pi'| \leq |p - \pi|$ is strict if, and only if, $D(V^+, V^-) \leq 2$. (Here D is distance in the graph.)

For $D(V^+, V^-) > 2$, let $(V^+)' = V^+ \cup N(V^+)$

That is, $(V^+)'$ is the “ V^+ of p' .” Define $(V^-)'$ similarly. ($N(V^+)$ is the neighborhood of V , ie, $N(S) = \{w : \exists v \in S, v \sim w\}$.)

v is guaranteed to be in $(V^+)'$ if $(\{v\} \cup N(\{v\})) \cap V^- = \emptyset$ and if $(\{v\} \cup N(\{v\})) \cap V^+ \neq \emptyset$. Similarly for $(V^-)'$. Therefore $V^+ \subset (V^+)'$ and $V^- \subset (V^-)'$ so long as $D(V^+, V^-) > 2$. Since the graph is connected, we will reach $D(V^+, V^-) \leq 2$, triggering strict decrease of the variation distance.

Comment: The proof works also for other walks, such as the walk with uniform transition probabilities $\frac{1}{d_v+1}$ for each node. In that case, $\pi_v \sim d_v + 1$.

In general, figuring out the stationary distribution of a vertex of a Markov chain requires examining the entire transition matrix. However, there's a special class of chains, which includes many of those we're interested in, for which the distribution is easy to obtain.

Definition: A chain is **balanced** if

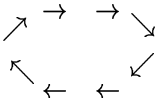
1. The transition probabilities are uniform on the out neighbors.
2. The number of in neighbors equals the number of out neighbors at every vertex.

Both of the types of walks on undirected graphs described at the beginning of the lecture are balanced. For a balanced chain, the stationary distribution is obtained as follows: let d_v be the in-degree (equivalently the out-degree) of v , and let p_{vv} (as before) be the probability that the walk stays at v . Then (as is easy to verify), π_v is proportional to $d_v/(1 - p_{vv})$.

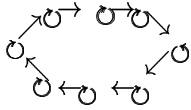
A **regular** chain is balanced and has the same degree, and same p_{vv} , at each vertex. So the stationary distribution of a regular chain is uniform.

Examples:

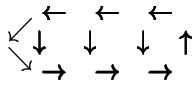
- A chain that is regular but does not mix.



- This graph has degree in = degree out = 2 everywhere. With transition probabilities of $\frac{1}{2}$ each, this chain does mix.



- A chain that is directed and unbalanced. One can imagine this chain with the ladder being very long.



Note that there are two kinds of junctions on the ladder:



and



The stable state for the vertex at the top and bottom of each rung is the same, but the stationary distribution for each vertex decreases exponentially with distance from the right hand side.

Another special type of Markov chain, commonly considered both in Physics and in Computer Science, is a *reversible* Markov chain. This is an ergodic chain satisfying the *detailed balance condition*: $\forall u, v : \pi_u p_{uv} = \pi_v p_{vu}$. We won't have more to say about these chains just now.