Finite Markov Chains

\( \Omega \) is a finite set.

\( X_0, X_1, \ldots, X_t, \ldots \)

is a sequence of random variables on \( \Omega \).

\[
\Pr [X_{t+1} = y | X_0, X_1, \ldots, X_t = \omega] = \Pr [X_{t+1} = y | X_t = \omega]
\]

Markov Property
\[ P(x, y) = P(x, y) \]

Prob. of going from state \( x \) to state \( y \) in one step.

\[ \sum_y P(x, y) = 1 \]

\( P \) is non-negative

(\( \omega \) stochastic)
Example 1

Pad 1

Pad 2

States are \( \mathcal{S} = \{ 1, 2 \} \)

\[
P = \begin{bmatrix}
1 & p \\
q & 1-q
\end{bmatrix}
\]

On pad 1 go to pad 2 with prob. \( p \), else stay put.
\[ P(x, y) = P_r[x \rightarrow y \text{ in one step}] \]

\[ P^b(x, y) = P_r[x \rightarrow y \text{ in } b \text{ steps}] \]

**Induction on \( b \).**

**True for \( b = 1 \).**

**Assume true for some \( b > 1 \).**

\[ P^{b+1}(x, y) = \sum_{z \in E} P^b(x, z) P(z, y) \]
\[ p^{t+1}(x, y) = \sum_{z \in \mathcal{Z}} p^t(x, z) p^t(z, y) \]

Going from \( x \) to \( y \) in \( t+1 \) steps

Different \( z \) define disjoint events

Start in state \( x \) with probability \( \mu(x) \)

Probability in \( y \) after \( t \) steps:

\[ \sum_{x} \mu(x) p^t(x, y) = (\mu p^t)_y \]
In many chains of interest, as $t \to \infty$, the probability of being here, starting at $t_0$, can be represented as:

$$\begin{pmatrix}
\pi_1 & \pi_2 & \cdots & \pi_n \\
\pi_1 & \pi_2 & \cdots & \pi_n \\
\pi_1 & \pi_2 & \cdots & \pi_n \\
\cdots & \cdots & \cdots & \cdots
\end{pmatrix}$$

all rows look same.
The probabilities $\pi_1, \ldots, \pi_n$ are called the **steady state distribution**.
Irreducibility and Aperiodicity

$P$ is irreducible if $\forall x, y$ there exists a $t = t(x,y)$ such that $P^t(x,y) > 0$

$\Omega$: directed graph $(\Omega, A)$

Irreducible $\iff$ strongly connected

$\exists$ directed path $\alpha \rightarrow y$ in $\Omega$, $\forall x, y$
Define a relation \( \approx \)

\[ a \approx b \iff \exists \text{ path } a \rightarrow b \quad \text{ and path } b \rightarrow a \]

\( \approx \) is an equivalence relation.

Equivalence classes are called "strong component." Denote \( D_1, D_2, \ldots, D_k \).

Define \( \Gamma = (\xi_1, \xi_2, \ldots, k, \xi), B \) such that

\[(i,j) \in B \iff \exists \ x \in D_i \quad \text{ and } \ y \in D_j \quad \text{ and path } x \rightarrow y \in D\]
\[ \Gamma \text{ in acyclic.} \]

\[ \Rightarrow \text{ all } w_i \text{ in same component.} \]

\[ \Gamma \text{ is DAG: } \]
Periodicity

Start at 1, at even times, on left
at odd times, on right
no steady state.

Fix this: by putting a loop at each vertex.

\[ P \rightarrow \frac{I + P}{2} \]

\[ P(x_n) \geq \frac{1}{2} \forall n \]
\[ T(n) = \exists \delta \in R : \rho^n(x, x) \geq 0 \delta \]

Period of \( n \) = \( \text{gcd } T(n) \)

**Lemma**

If \( \rho \) is irreducible then

\[ \text{gcd } T(x) = \text{gcd } T(y), \quad \forall x, y \]

= period of chain
Proof

For two stolen \( x, y \).

\( \exists r, l \) such that

\[ P'(x, y) > 0 \quad \text{and} \quad P'(l, x) > 0 \]

\[ m = r + l \in T(x) \cap T(y) \]

\[ T(x) \leq T(y) - m \]

\[ \sum x_0, x_1, \ldots, x_m \]

\( \gcd T(y) \) divides everything in \( T(x) \).

\[ x_i = y_i - x_0 \]

\[ y_0 \]
Proposition 1.7

If $P$ is aperiodic and irreducible, then there exists $r > 0$ such that $P^r(x,y) > 0$ for all $x, y$.

Proof

Fact: If $a_1, a_2, \ldots, a_k$ are positive integers with $\gcd(1, a_i) = 1$ for $i = 1, \ldots, k$, then there exist non-negative integers $n \geq n_0$ such that

$$n = a_1 \theta_1 + a_2 \theta_2 + \ldots + a_k \theta_k$$

(Frobenius number)

where $\theta_1, \theta_2, \ldots, \theta_k$ are non-negative integers.
Another way of phrasing:

\[ A \leq \xi \downarrow \omega \ldots \]

1. \( \text{gcd} A = 1 \)

2. \( A \) is closed under addition

\[ \Rightarrow \exists n_0 \text{ s.t. } n \geq n_0 \Rightarrow n \in \mathcal{A}. \]

\( \mathcal{T}(x) \)'s fit the claim.

\[ \rho^n(x, y) > 0 \]

\[ n \geq n_0 = \max(n_1, n_2, \ldots, n_{\omega}) \]
\[ \rho^t(x, y) > 0 \text{ for } t \geq r + n_0 \]
C.S. example of Markov Chains

\[ G = (V, E) \] of maximum degree \( \Delta \)

Suppose \( k > \Delta \)

Then \( \exists \) a \( k \)-coloring of \( G \).

[ Give each vertex a color from \( 1, 2, \ldots, k \) such that adjacent vertices get different color.]

Suppose I want to count \# ways \( \nu \)-coloring \( G \). (1) Difficult to do exactly
(2) Approximately?
\[ \Omega = \{ k \text{-colorings of } G \} \]
\[ \Omega \neq \emptyset \text{ if } k > \Delta \]

Suppose I could choose a random coloring. — (Markov chains will do this)

Choose large number
\[ v, \text{ } v \text{ not adjacent to } w \]

Can estimate ratio
\[ \frac{\text{colorings } c(v) = c(w)}{\text{# colorings } c(v) \neq c(w)} \]
Generating a random member \( \mathcal{S} \)

Define Markov chain on \( \mathcal{S} \).

Given a coloring \( w \in \mathcal{S} \)

1) choose random vertex
2) randomly recolor.

\( k \geq \Delta + 2 \implies \) (i) chain is "ergodic"

Run chain for a "long time" and take coloring

\( k > 2\Delta \theta(n\log n) \) is enough