Markov Chain

random walk in state space

1 million people

Population migration

probability \( \frac{1}{2} \), Stay, probability \( \frac{1}{4} \) to one other state.

Arrow of time, coverage to stationary distribution

Entropy (randomness) increases
**Notation:**

- $X_t$: State at time $t$
- $p^{(t)}$: distribution of $X_t$ (population distribution)

<table>
<thead>
<tr>
<th>$p^{(t)}$</th>
<th>$p_1^{(t)}$</th>
<th>$p_2^{(t)}$</th>
<th>$p_3^{(t)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t=0$</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$t=1$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>$t=2$</td>
<td>$\frac{3}{8}$</td>
<td>$\frac{5}{16}$</td>
<td>$\frac{5}{16}$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$t \to \infty$</td>
<td>$\frac{1}{3} \pi_1, \frac{1}{3} \pi_2, \frac{1}{3} \pi_3$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Key / driving force:**

- Transition Probability

\[
K_{ij} = P(X_{t+1} = j \mid X_t = i)
\]

or \[
K(x, y) = P(X_{i+1} = y \mid X_i = x)
\]

**transition matrix:**

\[
\begin{pmatrix}
1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\
\frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{2}
\end{pmatrix}
\]
Master Equation

\[ p_j^{(t+1)} = p(X_{t+1} = j) \]

\[ = \sum_i p(X_{t+1} = j \& X_t = i) \]

\[ = \sum_i p(X_t = i) p(X_{t+1} = j / X_t = i) \]

\[ = \sum_i p_i^{(t)} k_{ij} \]

So the master equation is:

\[ p_j^{(t+1)} = \sum_i p_i^{(t)} k_{ij} \]

\[ p^{(t)} \xrightarrow{K} p^{(t+1)} \]

\[ p^{(t+1)} = p^{(t)} K \Rightarrow p^{(t)} = p^{(0)} k^t \]

Interpret the master equation
1  \[ \frac{1}{2} \] \[ \frac{3}{6} \] 1 \[ P_i^{(t)} = \text{# of people in } i \text{ at } t \]

Kij = fraction of people in i who will go to j

\[ P_i^{(t)} K_{ij} = \text{# of people in } i \text{ at } t \text{ who will go to } j \]

\[ \sum P_i^{(t)} K_{ij} = \text{# of people who will end up in } j \text{ at } t+1 \]

\[ \approx \text{equilibrium} \]

**Stationary } \pi \]

\[ \pi \nu = \pi K \]

\[ \pi \nu_j = \sum \pi \nu_i K_{ij} \rightarrow \text{Overall balance} \]

\[ \text{Detailed balance } \nu_{i,j} \]

\[ \nu_i K_{ij} = \nu_j K_{ji} \]

\[ \nu_i, K_{ij}, \nu_j \]
Page Rank: know $K$, want $\pi$

Start from $p^{(0)}$ (e.g. uniform)

$p^{(0)} \xrightarrow{K} p^{(1)} \xrightarrow{K} p^{(2)} \xrightarrow{\cdots} \xrightarrow{K} p^{(T)} \approx \pi$

$k$ is a stochastic matrix, the sum of each row $= 1$

Why recursive? more stable, accurate measure of popularity

MCMC: know $\pi$, want $K$

(starting distribution) $\Rightarrow$ (iterative algorithm) $\Rightarrow$

so that $\pi^T = \pi^T K$ (designed principle)

Start from $X_0 \sim p^{(0)}$

$X_0 \rightarrow X_1 \rightarrow X_2 \rightarrow \cdots \rightarrow X_t \rightarrow X_{t+1} \rightarrow \cdots \rightarrow X_T$

Approx: $X_T \sim \pi$
Metropolis algorithm

1950s Metropolis, Roseblatt, Teller

Basic idea

Base Chain: B

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\
2 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\
3 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\
\end{array}
\]

\[\tau_0\]

Think of \( \square \) as a visa approval/rejection office

Extra-rejection probability

good, balanced

bad, unbalanced
Algorithm:

1. If $\pi(x) B(x, y) > \pi(y) B(y, x)$
   \[ \text{prob(accept } x \rightarrow y) = \frac{\pi(y) B(y, x)}{\pi(x) B(x, y)} \]

2. If $\pi(x) B(x, y) \leq \pi(y) B(y, x)$
   \[ \text{prob(accept } x \rightarrow y) = 1 \]

3. Combine (1), (2):
   \[ \text{prob(accept } x \rightarrow y) = \min\left(1, \frac{\pi(y) B(y, x)}{\pi(x) B(x, y)}\right) \]
\[ X_t = x \]
\[ X_{\text{propose}} \sim B(x, y) = P(X_{\text{propose}} = y | X_t = x) \]

Let \( X_{t+1} = \begin{cases} X_{\text{propose}} & \text{with } P = \min(1, \frac{\pi(y) B(y, x)}{\pi(x) B(x, y)}) \\ X_t = x & \text{with } 1 - P \end{cases} \]

**Special Case:**
Metropolis algorithm: \( B(x, y) = B(y, x) \)

\[ P = \min(1, \frac{\pi(y)}{\pi(x)}) \]

Hastings Extension: general \( B \)

**Traveling Salesman Problem**

\[ \begin{array}{c}
0 \rightarrow 5 \\
1 \rightarrow 5 \\
5 \rightarrow 4 \\
4 \rightarrow 3 \\
3 \rightarrow 2 \\
2 \rightarrow 1 \\
1 \rightarrow 0 \\
\end{array} \]

\( X: \) path 0 → 1 → 2 → 3 → 4 → 5 → 0

Another \( X: \) 0 → 1 → 3 → 2 → 5 → 4 → 0
\[ \xi(x) = \text{total length of path } x. \]
\[ \Pi(x) = \frac{1}{Z} e^{-\xi(x)} \]
\[ Z = \sum_x e^{-\xi(x)} \]

Gibbs distribution / Boltzmann distribution

Searching thru such huge space is extremely difficult

Sampling and optimization are related here

State space (n factorial)

- Each state space is a permutation
- Each move is a random transposition

Other permutations

State space: path

Random transposition

State space: path
Base Chain

Randomly pick 2 cities & exchange their positions.

Propose change by random transposition.

accept change with

\[ p = \min(1, \frac{\pi(y)}{\pi(x)}) \]

\[ = \min(1, e^{\frac{\mathcal{E}(x) - \mathcal{E}(y)}{T}}) \]

Recall \( \pi(-) = \frac{1}{Z} e^{-\mathcal{E}()} \)

Here \( \frac{1}{Z} \) gets cancelled.

\[ X_0 \rightarrow X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow \ldots \]

\((1\,2\,3\,4\,5) \rightarrow (1\,4\,3\,2\,5) \rightarrow (1\,4\,2\,3\,5) \rightarrow (1\,4\,2\,3\,5) \rightarrow \ldots \)

\( \rightarrow X_{1000} \) we end up w/ a good permutation

Approx: \( X_{1000} \sim \pi(X) \)