L22: Markov Chains

Jeff M. Phillips

April 8, 2020

Markov Chain : Life Lessons

Markov Chain : Life Lessons

- [**L1**] *Only your current position matters going forward, don't worry about the past.*

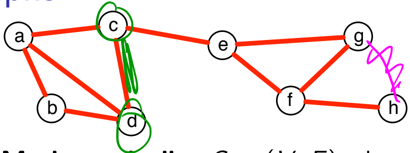
Markov Chain : Life Lessons

- [**L1**] *Only your current position matters going forward, don't worry about the past.*
- [**L2**] *You just need to worry about one step at a time; you will get there eventually (or you won't).*

Markov Chain : Life Lessons

- [**L1**] *Only your current position matters going forward, don't worry about the past.*
- [**L2**] *You just need to worry about one step at a time; you will get there eventually (or you won't).*
- [**L3**] *In the limit, everyone has perfect karma.*

Graphs



Mathematically: $G = (V, E)$ where

$V = \{a, b, c, d, e, f, g\}$ and

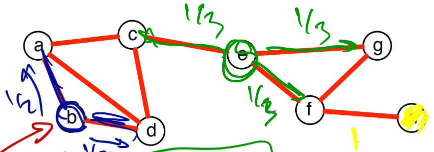
$$E = \Big\{\{a,b\},\{a,c\},\{a,d\},\{b,d\},\{c,d\},\{c,e\},\{e,f\},\{e,g\},\{f,g\},\{f,h\}\Big\}.$$

Matrix-Style: As a matrix with 1 if there is an edge, and 0 otherwise. (For a directed graph, it may not be symmetric).

adjacency

$$G = \begin{array}{c|cccccccc} & a & b & c & d & e & f & g & h \\ \hline a & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ b & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ c & 1 & 0 & 0 & \boxed{1} & 1 & 0 & 0 & 0 \\ d & 1 & 1 & \boxed{1} & 0 & 0 & 0 & 0 & 0 \\ e & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ f & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ g & 0 & 0 & 0 & 0 & 1 & 1 & 0 & \boxed{0} \\ h & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{array} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Markov Chain



$\frac{1}{10}$ th in a

$\frac{3}{10}$ in d

$\frac{6}{10}$ in f

(V, P, q): V node set, P probability transition matrix, q initial state.

e.g. $q^T = [0\ 1\ 0\ 0\ 0\ 0\ 0\ 0]$ or $q^T = [0.1\ 0\ 0\ 0.3\ 0\ 0.6\ 0\ 0]$.

$$
\begin{array}{cc}
 & \begin{array}{cccccccc} & & & & e & & & \end{array} \\
P = & \begin{pmatrix}
0 & 1/2 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\
1/3 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 \\
1/3 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 0 \\
1/3 & 1/2 & 1/3 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/3 & 0 & 0 & 1/3 & 1/2 & 0 \\
0 & 0 & 0 & 0 & 1/3 & 0 & 1/2 & 1 \\
0 & 0 & 0 & 0 & 1/3 & 1/3 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 0
\end{pmatrix}
\end{array}
$$

Transitions

$$P = \begin{pmatrix} 0 & 1/2 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 0 & 1/3 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/2 & 1 \\ 0 & 0 & 0 & 0 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 0 \end{pmatrix} \text{ and } q^T = \boxed{[0\ 1\ 0\ 0\ 0\ 0\ 0\ 0]}$$

in b wp. 1

$$q_1 = Pq = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}^T .$$

u d

Transitions

$$P = \begin{pmatrix} 0 & 1/2 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 0 & 1/3 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/2 & 1 \\ 0 & 0 & 0 & 0 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 0 \end{pmatrix} \text{ and } q^T = [0\ 1\ 0\ 0\ 0\ 0\ 0\ 0]$$

$$q_1 = Pq = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

$$q_2 = P\boxed{q_1} = \boxed{PP}q = \boxed{P^2}q = \begin{bmatrix} \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

Transitions

$$P \underset{e}{=} \begin{array}{c} \begin{matrix} 1/6 & 2/6 & \boxed{2/6} & 1/6 & & & & \end{matrix} \\ \begin{pmatrix} 0 & 1/2 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{1/3} & 0 & 0 & 1/3 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/2 & 1 \\ 0 & 0 & 0 & 0 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 0 \end{pmatrix} \end{array} \text{ and } q^T = [0\ 1\ 0\ 0\ 0\ 0\ 0\ 0]$$

$$q_1 = Pq = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

$$q_2 = Pq_1 = PPq = P^2 q = \begin{bmatrix} \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

$$q_3 = Pq_2 = \begin{bmatrix} \frac{1}{3} & \frac{1}{9} & \frac{1}{9} & \frac{1}{3} & \frac{1}{9} & 0 & 0 & 0 \end{bmatrix}^T.$$

Transitions

$$P = \begin{pmatrix} 0 & 1/2 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 0 & 1/3 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/2 & 1 \\ 0 & 0 & 0 & 0 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 0 \end{pmatrix} \text{ and } q^T = [0\ 1\ 0\ 0\ 0\ 0\ 0\ 0]$$

$$q_1 = Pq = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

$$q_2 = Pq_1 = PPq = P^2 q = \begin{bmatrix} \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

$$q_3 = Pq_2 = \begin{bmatrix} \frac{1}{3} & \frac{1}{9} & \frac{1}{9} & \frac{1}{3} & \frac{1}{9} & 0 & 0 & 0 \end{bmatrix}^T.$$

In the limit: $q_n = \boxed{P^n} q$ lim $n \to \infty$

not uniform

Transitions

$$P = \begin{pmatrix} 0 & 1/2 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 0 & 1/3 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/2 & 1 \\ 0 & 0 & 0 & 0 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 0 \end{pmatrix} \text{ and } q^T = [0\ 1\ 0\ 0\ 0\ 0\ 0\ 0]$$

$$q_1 = Pq = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

$$q_2 = Pq_1 = PPq = P^2q = \begin{bmatrix} \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

$$q_3 = Pq_2 = \begin{bmatrix} \frac{1}{3} & \frac{1}{9} & \frac{1}{9} & \frac{1}{3} & \frac{1}{9} & 0 & 0 & 0 \end{bmatrix}^T.$$

In the limit: $q_n = P^n q$

[**L1**] *Only your current position matters going forward, don't worry about the past.*

Two ways to think about Markov Chains

① Only consider 1 possible location at a time

($e.g.$ $q = [0, 1, 0, 0, \dots 0]$)

↳ Random Walk

② Probability distribution on states

e.g. $q = \left(\frac{1}{10}, 0, \frac{3}{10}, 0, 0, \frac{6}{10}, 0 \dots\right)$

Ergodic: MC is ergodic if
$\exists\, t$ so that for **all** $n > t$

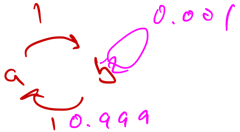
$$g_n(j) > 0.$$

True if not

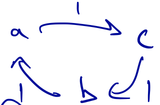
1. Cyclic
2. has Absorbing & transient states
3. disconnect.

Cyclic Examples

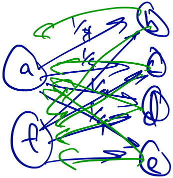
$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$



$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

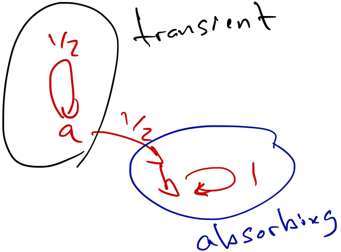


$$\begin{array}{c} \begin{pmatrix} 0 & 1/2 & 1/2 & 1/2 & 1/2 & 0 \\ 1/4 & 0 & 0 & 0 & 0 & 1/4 \\ 1/4 & 0 & 0 & 0 & 0 & 1/4 \\ 1/4 & 0 & 0 & 0 & 0 & 1/4 \\ 1/4 & 0 & 0 & 0 & 0 & 1/4 \\ 0 & 1/2 & 1/2 & 1/2 & 1/2 & 0 \end{pmatrix} \\ \begin{array}{cccccc} a & & & & & f \end{array} \end{array}$$



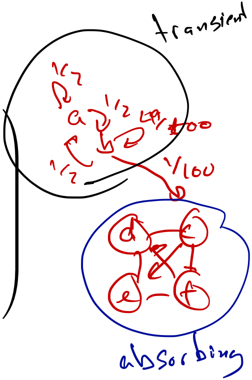
Absorbing and Transient Examples

$$\begin{pmatrix} 1/2 & 0 \\ 1/2 & 1 \end{pmatrix}$$



$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 49/100 & 0 & 0 & 0 & 0 \\ 0 & 1/100 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$



Unconnected Examples

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 1/2 & 1/3 & 0 \\ 0 & 0 & 1/3 & 0 & 1/3 & 0 \\ 0 & 0 & 1/3 & 1/2 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Limiting State

if MC is ergodic

Let $P^* = P^n$ as $n \to \infty$.
Let $q_* = P^* q$.

in 1 step P^* pushes the state to the final state

q_* vector, has weight for each node

$q_*(j)$ jth node

Limiting State

Let $P^* = P^n$ as $n \to \infty$.
Let $q_* = P^* q$.

> [**L2**] *You just need to worry about one step at a time; you will get there eventually (or you won't).*

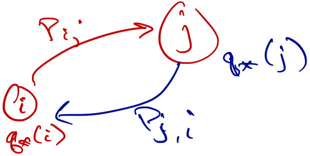
Delicate Balance

q* is an eigen-
vector of P.
↳ always the top
eigenvector

Let $P^* = P^n$ as $n \to \infty$.
Let $q_* = P^* q$.
Also $q_* = P\underbrace{\boxed{P^* q}}_{q_*}$ thus $q_* = Pq_*$.

So the probability of (being in a state i and leaving to j) is the same as (being in another state j and arriving in i)

$$P_{i,j}q_*(i) = P_{j,i}q_*(j)$$



Delicate Balance

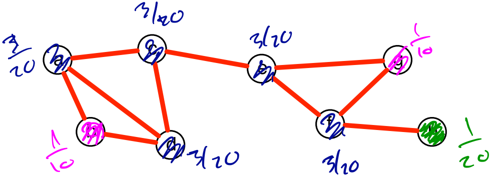
Let $P^* = P^n$ as $n \to \infty$.
Let $q_* = P^* q$.
Also $q_* = PP^* q$ thus $q_* = Pq_*$.

So the probability of (being in a state i and leaving to j) is the same as (being in another state j and arriving in i)

$$P_{i,j} q_*(i) = P_{j,i} q_*(j)$$

[**L3**] *In the limit, everyone has perfect karma.*

Limiting State



In our example

$$q_* = (0.15, 0.1, 0.15, 0.15, 0.15, 0.15, 0.1, 0.05)$$
$$= (\frac{3}{20}, \frac{1}{10}, \frac{3}{20}, \frac{3}{20}, \frac{3}{20}, \frac{3}{20}, \frac{1}{10}, \frac{1}{20})$$

Algorithms for computing q_*

① eig(P) $\Rightarrow$ top eigenvector v_1

$$q_* = \frac{v_1}{\|v_1\|_1} \qquad \|v\|_1 = \sum_{j=1}^{n} |v(j)|$$

② $q_* \approx \underbrace{\boxed{P \cdot P} \cdot (P \cdots (P}_{m} q_0)))$

for $i = 1$ to m

$q_i = P q_{i-1}$

return $q_* = q_m$

③ $q_* \approx (P^m) q_0$

$$P^m = \prod_{i=1}^{m} P = P^{m/2} \cdot P^{m/2}$$

$$P^{m/2} = P^{m/4} \cdot P^{m/4}$$

④ Random Walk

Maintain explicit state

$q = [0, 1, 0, \ldots 0] \Rightarrow q = b \quad q \in V$

Burn in Period
≈ 1000 steps

each step

$v_i \rightarrow$ to neighbor on edge v_{i+1}

collect the next
≈ 5000 steps

Collect $\{v_1, v_2, \ldots v_m\} = S$

$$q_*(j) = \left(\frac{\#\ v \in S \text{ s.t. } v = j}{m} \right)$$

Metropolis Algorithm

Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller in 1953

energy state g

$E(g)$

Probability particle in g

proportional to

$w(g) = e^{-E(g)}$

Metropolis Algorithm

Markov Chain

$(V, P, \mathcal{E})$

Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller in 1953

$V = \mathbb{R}^d$

P = algorithm (based on weights w)

$v_0 \in \mathbb{R}^d$

<u>Metropolis on V and w</u>

Initialize $v_0 = [0\ 0\ 0\ \ldots\ 1\ \ldots\ 0\ 0]^T$.

repeat

  Generate $u \sim K(v, \cdot)$

  if $(w(u) \geq w(v_i))$ **then**

    Set $v_{i+1} = u$

  else

    With probability $w(u)/w(v)$ set $v_{i+1} = u$ < 1

  else

    Set $v_{i+1} = v_i$

until "converged"

return $\boxed{V = \{v_1, v_2, \ldots, \}}$