L22: Markov Chains

Jeff M. Phillips

April 9, 2018

Markov Chain : Life Lessons (ergodic)

Markov Chain : Life Lessons

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

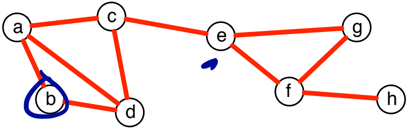
Markov Chain : Life Lessons

- [**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

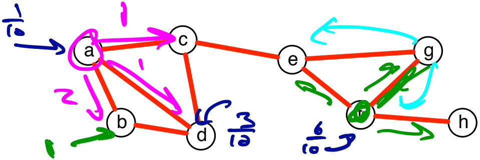
$V = \{a, b, c, d, e, f, g\}$ and

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

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

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

Markov Chain



(V, P, q): V node set, P probability transition matrix, $\boxed{q \text{ initial state.}}$

e.g. $\boxed{q^T = [0\ 1\ 0\ 0\ 0\ 0\ 0\ 0]}$ or $\boxed{q^T = [\underset{a}{0.1}\ 0\ 0\ \underset{d}{0.3}\ 0\ \underset{f}{0.6}\ 0\ 0]}$.

$$\begin{array}{c} \\ 1/2 \\ 1/4 \\ 1/4 \\ \\ \\ \\ \\ \end{array} P = \begin{pmatrix} 0 & 1/2 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 0 & 1/3 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & \boxed{1/2} & 1 \\ 0 & 0 & 0 & 0 & 1/3 & \boxed{1/3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 0 \end{pmatrix} \begin{array}{c} \\ \\ \\ \\ \\ f \\ g \\ \\ \end{array}$$

$$\begin{array}{cccccccc} & & & & & f & g & \end{array}$$

Transitions

prob transitions

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

state

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

Transitions

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

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

$$q_2 = Pq_1 = \underbrace{PPq}_{\text{mult } 2x} = \boxed{P^2} q = \begin{bmatrix} \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

Transitions

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

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

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

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

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: $q_n = P^n q$

$q_* \leftarrow$ ergodic

Transitions

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

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

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

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

$$P^n = P \cdot P^{n-1} = P^{n-1} \cdot P = \underbrace{P \cdot P \cdots P}_{n \text{ times}}$$

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

Markov

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

Two ways to think about MC

(1) Random Walk.
& always [0, 0 ... 1, 0, 0]
exactly on one vertex

maintain 1 state

Metropolis Algorithm

(2) Probability Distribution
& can be dense.

maintain dense &

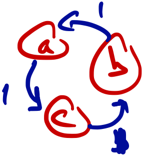
n states (2) takes $O(n)$ space
(1) takes $O(\log n)$ space

Cyclic Examples

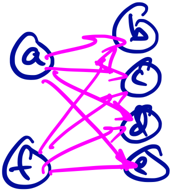
: MC alternate between states at fixed epoch of size $p \geq 2$

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

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



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



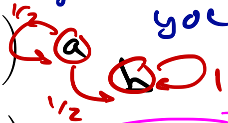
bipartite graph

Absorbing and Transient Examples

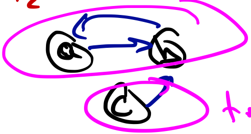
Set of states that you can never leave

set of states that if you leave you can't go back.

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



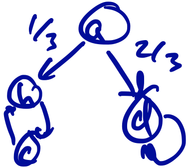
$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$



transient

$$\begin{array}{cccccc} a & b & & & & \\ \multicolumn{6}{c}{\left(\begin{array}{cccccc} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 49/100 & 0 & 0 & 0 & 0 \\ 0 & 1/100 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \end{array}\right)} \\ & & c & d & e & f \end{array}$$

absorbing



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

MC is ergodic if it

is not ① cyclic, ② having absorbing and transient states, and ③ unconnected.

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

Let $q_* = P^* q$.

independent of q

for any initial q

Limiting State

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

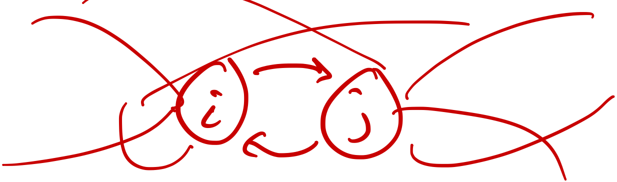
> [**L2**] *You just need to worry about one step at a time; you will get there eventually (or you won't).*

Delicate Balance

Let $P^* = P^n$ as $n \to \infty$.
Let $q_* = P^* q$.
Also $q_* = P\underline{P^* q}$ thus $q_* = P\underline{q_*}$.

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.*

How q_* relates to P

$\to$ 1st eigenvector of P.

$[V, L] = \text{eig}(P)$ $\Leftarrow$ matlab

eigen vectors values

$$\frac{v_1}{\text{sum}(v_1)} = q_*$$

2nd eigenvalue λ_2 (set $\lambda_1 = 1$)

Tells how fast converge.

smaller λ_2 $\to$ faster convergence.

$$\frac{\lambda_2}{\lambda_1}$$

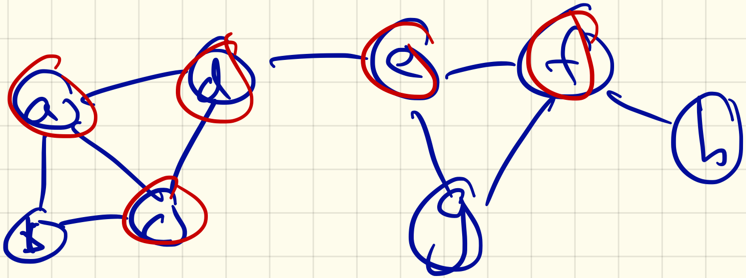
Example $\lambda_2 = 0.875$ $\lambda_2 \approx 0.9$ super fast

$\lambda_2 = 0.99$ slow fast

Example Graph

$$q_* = (0.15,\ 0.1,\ 0.15,\ 0.15,\ 0.15,\ 0.15,\ 0.1,\ 0.05)$$

a c d e f



Metropolis Algorithm

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

Set of states V $\{[\omega],\ \mathbb{R}^d,\ [0,1]^d\}$

$\rightarrow$ can evaluate $w(v)$ for $v \in V$

$\leftarrow$ one state.

$$\boxed{e^{-\text{Energy}(v)} = w(v)} \geq 0$$

$$\sum_{v \in V} w(v) \text{ or } \int_{v \in V} w(v)\,dv = W$$

Don't know W

$$\text{Prob}[v \in V] \approx w(v)$$

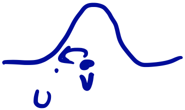
Metropolis Algorithm

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

Metropolis on V and w

exactly 1 start

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



repeat

Generate $u \sim K(v, \cdot)$ $= e^{-\|u, v\|^2}$

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

Set $v_{i+1} = u$

else

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

else

Set $v_{i+1} = v_i$

until "converged"

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

$\underbrace{\cancel{v_1 \ldots v_{1000}}}_{\text{burnin phase}},\ \overbrace{v_{1001} \ldots v_{6000}}^{\text{average}}$

v_n as $n \to \infty$

$Pr[v_n] = \frac{w(v_n)}{\Sigma w}$