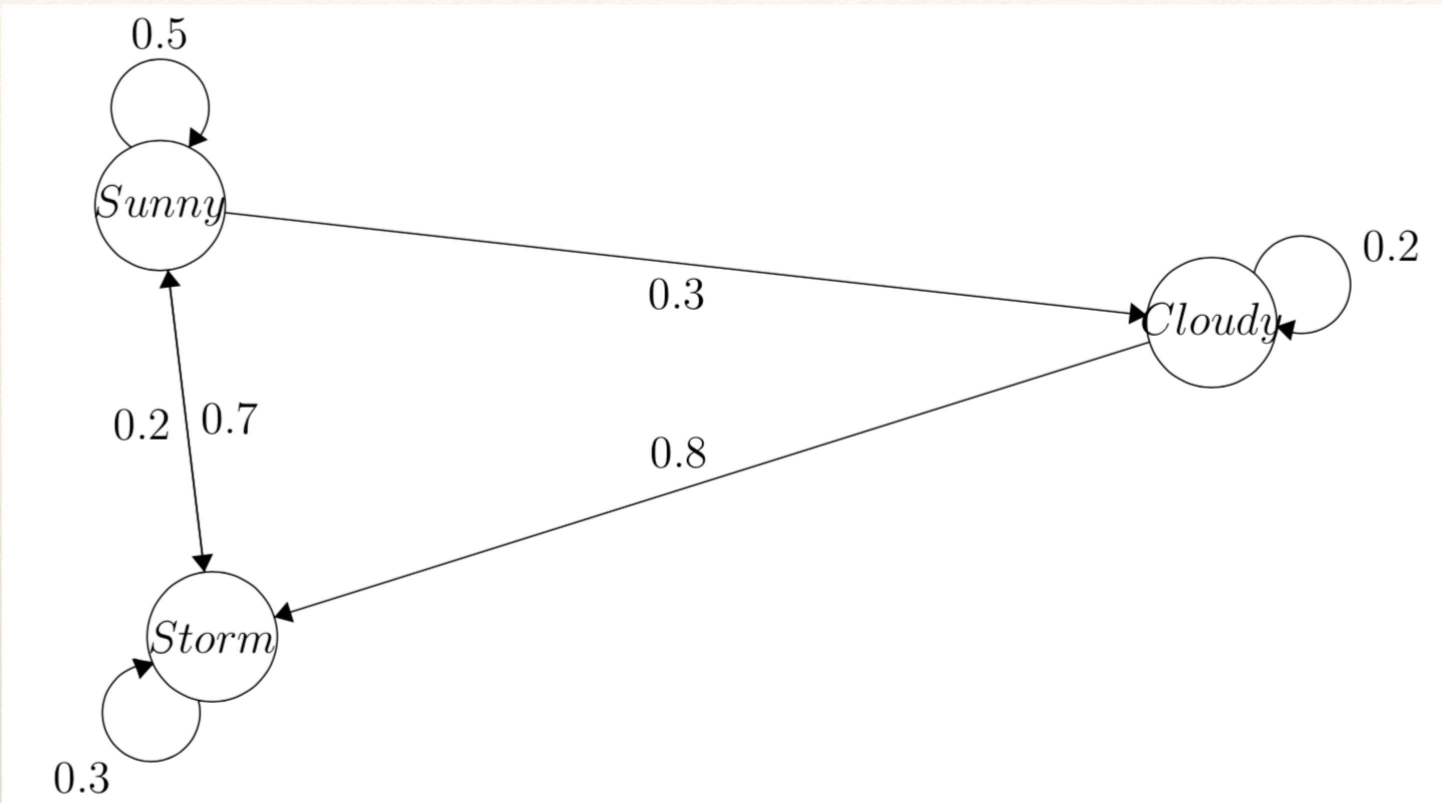
Directed Reading Program 2019 Spring

Markov Chains & Random Walks

Zifan Yu
Department of Mathematics,
University of Maryland
Mentored by Pranav Jayanti

Weather model:



	Sunny	Cloudy	Storm
Sunny	0.5	0.3	0.2
Cloudy	0	0.2	0.8
Storm	0.7	0	0.3

- Questions to consider:

  Given the probability distribution of the weather today is

  [a, b, c]

  - How do we predict the weather for tomorrow, if for each day, the probabilities of weather changes are all the same?
  - Is it possible that after a thousand years, the chances of weather for each day remain unchanged?

Markov Chains - what is it?

- Formally, a Markov chain is defined to be a sequence of random variables $(X_n)_{n\geq 0}$, taking values in a set of states, which we denote by S, with initial distribution λ and transition matrix P, if
  - X_0 has distribution $\lambda = \{\lambda_i \,|\, i \in S\}$
  - Transition matrix $P = (p_{ij})_{i,j\in S}$, and the Markov property holds:

    $$P(X_n = i_n \,|\, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_n = i_n \,|\, X_{n-1} = i_{n-1}) = p_{i_{n-1}i_n}$$

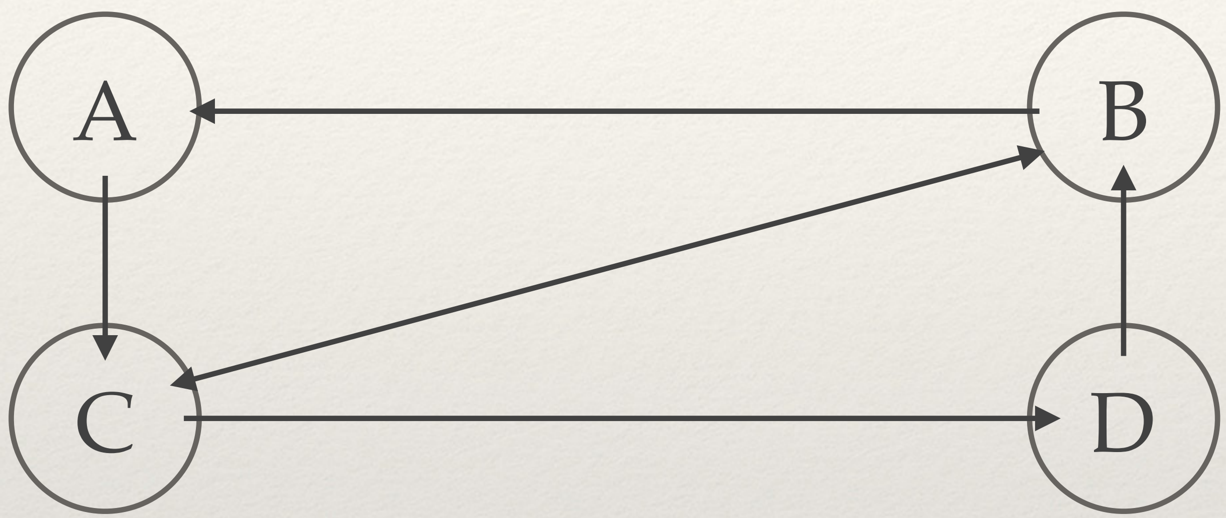
- Probability distributions

$$P(X_n = j) = (\lambda P^n)_j$$

$$P_i(X_n = j) = P(X_{n+m} = j \,|\, X_m = j) = p_{ij}^{(n)}$$

Markov Chains-communicating classes and irreducibility

We say that a state i communicate with state j if one can get to i from j, as well as from j to i with only finite many evolution times. We denote this relation as i <—>j.



Note: i —> j if and only if $p_{ik_1}, \cdots, p_{k_{n-1}j} > 0$. Also it requires the sequence $k_1, \ldots, k_{n-1}$ to be finite.

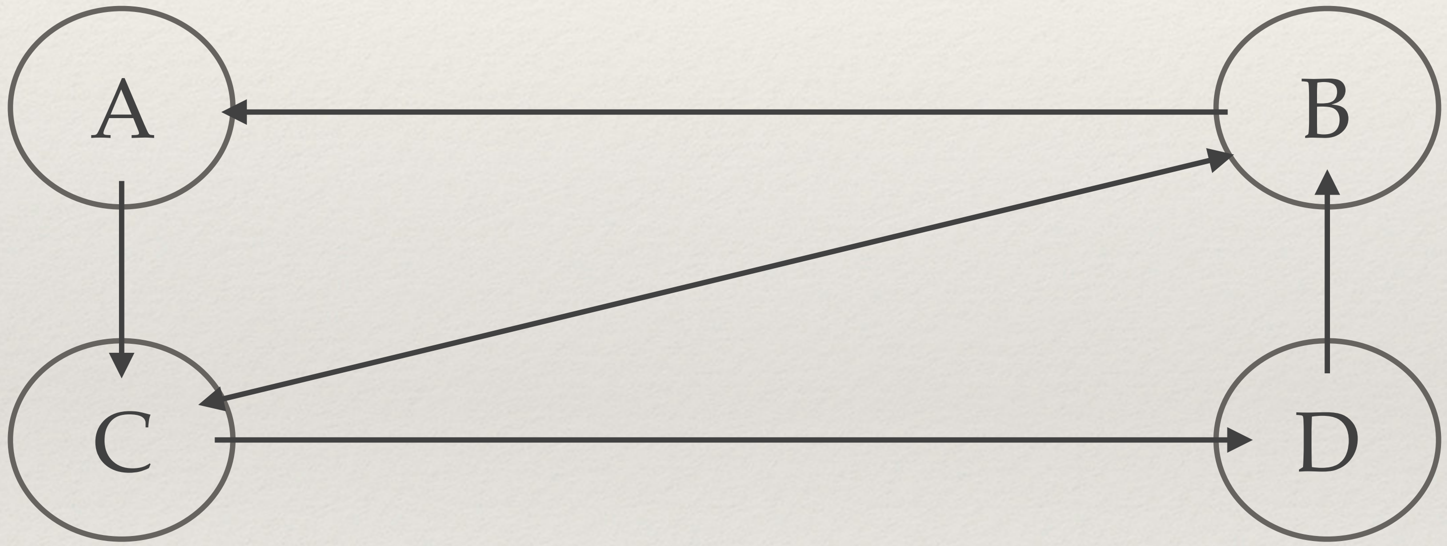
Also note that i<—>j means this relation is

(1) symmetric: if i —> j then j —> i;
(2) reflective: i <—> i;
(3) transitive: i <—> j and j <—> k imply i <—> k.

Markov Chains-communicating classes and irreducibility

The sets of states with states having such relation jointly are called communicating classes.

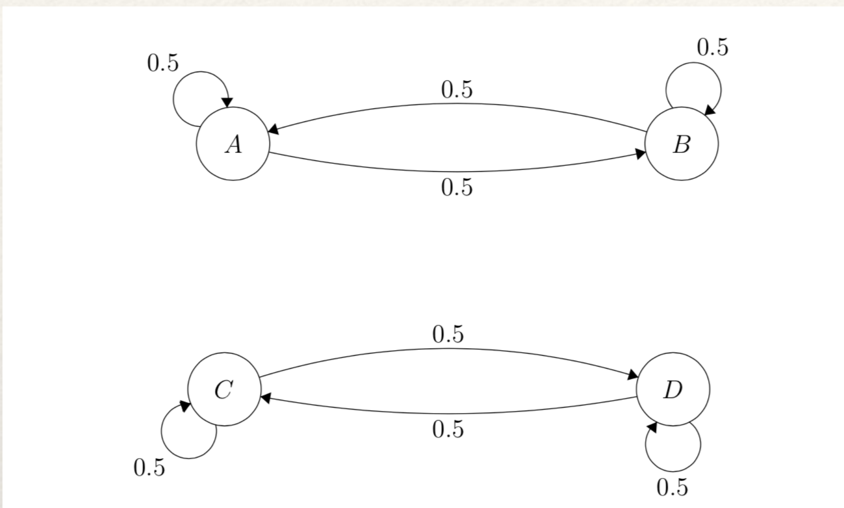
Therefore we can partition the set S, into communicating classes with respect to this equivalence relation.



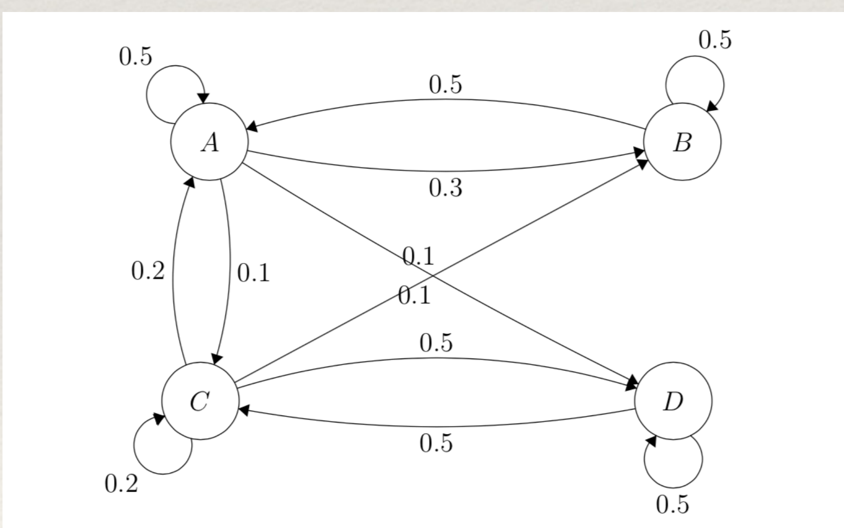
Definition : A Markov chain is irreducible if its set of states S is a single communicating class.

Markov Chains-communicating classes and irreducibility

Illustration of irreducible and reducible Markov chains:



$$\begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$



$$\begin{bmatrix} 0.5 & 0.3 & 0.1 & 0.1 \\ 0.5 & 0.5 & 0 & 0 \\ 0.2 & 0.1 & 0.2 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$

Note: Irreducibility of a Markov chain prepares us to study the equilibrium state of this chain.

Markov Chains-aperiodicity of Markov chains

- Definition: A state i is called <u>aperiodic</u>, if there exists a positive integer N, such that $p_{ii}^{(n)} > 0$ for all $n \geq N$.
- Theorem: If P is irreducible, and has an aperiodic state i, then for all states j and k, $p_{jk}^{(n)} > 0$ for all sufficiently large n. (therefore all states are aperiodic)

Sketch of the proof:

$$p_{jk}^{(r+n+s)} = \sum_{i_1,\ldots,i_n} p_{ji_1}^{(r)} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} p_{i_n k}^{(s)} \geq p_{ji}^{(r)} p_{ii}^{(n)} p_{ik}^{(s)} > 0$$

- Definition: We call a Markov chain <u>aperiodic</u> if all its states are aperiodic .

Now, recall the question: after sufficiently large evolution times, will the distribution of states reach an equilibrium?

Markov Chains-Invariant distributions

- A measure on a Markov chain is any vector $\lambda = \{\lambda_i \geq 0 \,|\, i \in S\}$
- In addition, λ is a distribution if $\sum_{i \in S} \lambda_i = 1$
- We say a measure λ is invariant if $\lambda = \lambda P$.

- Theorem: Suppose that $(X_n)_{n \geq 0}$ is a Markov chain with transition matrix P and initial distribution λ. If P is both irreducible and aperiodic, and has an invariant distribution π , then

$$P(X_n = j) = (\lambda P^n)_j \to \pi_j \text{ as } n \to \infty \text{ for all j.}$$

In particular,

$$p_{ij}^{(n)} \to \pi_j \text{ for all i,j.}$$

Markov Chains-Invariant distributions

(picture credit to Seattle Refined)

(picture credit to smithsonian.com)

(picture credit to BBC NEWS)

Markov Chains-Invariant distributions

By assuming that the finite-state Markov chain is irreducible and aperiodic, we can apply the **Perron-Frobenius Theorem.**

- The Perron-Frobenius Theorem:

  Let A be a positive square matrix. Then

  - A has one largest eigenvalue $\rho(A)$ in absolute value and it has an positive eigenvector.
  - $\rho(A)$ has geometric multiplicity 1.
  - $\rho(A)$ has algebraic multiplicity 1.

  Note: Also hold for nonnegative A s.t A^m is positive after some power m.

By applying the Perron-Frobenius Theorem to P,

- $\pi P = \pi \Leftrightarrow \rho(P) = 1$ with unique positive left eigenvector π .
- All other eigenvalues are of absolute values < 1.

Markov Chains-Invariant distributions

$$\lim qP^n$$

$$= \lim_{n\to\infty} \underset{1\times n}{[q_1\ q_2 \cdots q_N]} \underset{n\times n}{\underset{U}{\begin{bmatrix} 1 & * & * & \cdots \\ 1 & * & * & \cdots \\ \vdots & \vdots & \vdots & \\ 1 & * & * & \cdots & * \end{bmatrix}}} \underset{n\times n}{\underset{D^n}{\begin{bmatrix} 1 & & & \\ & \lambda_2^n & & \\ & & \ddots & \\ & & & \lambda_N^n \end{bmatrix}}} \underset{n\times n}{\underset{V}{\begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_N \\ * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{bmatrix}}}$$

$$= [1\ 0 \cdots 0] \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_N \\ * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{bmatrix}$$

$$= [\pi_1\ \pi_2 \cdots \pi_N]$$

Markov Chains - Recurrence and transience.

- Let $(X_n)_{n\geq 0}$ be a Markov chain with transition matrix P. Then a state $i \in S$ is recurrent if

$$P_i(X_n = i \text{ for infinitely many n}) = 1$$

- We say that i is transient if

$$P_i(X_n = i \text{ for infinitely many n}) = 0$$

Now we are ready to see one implementation of the abstract Markov chains- -the random walks.

Simple random walks-one dimension

We start by studying simple random walk on the integer latices. At each time step, the random walker flips a fair coin to decide its next move.

Let S_n denote the position at time n, x be the position it starts at. At each time step j,

$$X_j = \begin{cases} 1, \text{ if Head appears on the j-th throw;} \\ -1, \text{ otherwise.} \end{cases}$$

we have

$$S_n = x + X_1 + \ldots + X_n$$

$$P(X_j = 1) = P(X_j = -1) = 1/2$$

Questions:

- On average, how far is the walker from the starting point ?
- Does the walker keeps returning to the origin or does it eventually leave forever?

Simple random walks-one dimension

It’s easy to check that

$$E(S_n) = x + E(X_1) + \ldots + E(X_n) = x + 0 + \ldots + 0 = x;$$

and since (assume the walker starts from 0)

$$Var(X) = E(X^2) - E(X)^2 = E(X^2) = 1$$

we have

$$Var(S_n) = 0 + Var(X_1) + \ldots + Var(X_n) = n$$

$$\sigma_{S_n} = \sqrt{n}$$ (typical distance from the origin)

- What does this inform to us?

  In one dimension, there are at most $\sqrt{n}$ integers that are within typical distance with the mean distance.

  So the chance of lying on a particular integer should shrink as a a constant times $n^{-\frac{1}{2}}$.

$$P(S_n = j) \sim \frac{C}{\sqrt{n}}$$

Simple random walks-one dimension

We may notice that after an odd number of steps, the walker must end at an odd integer; similarly in order to get to an even integer, we need even steps.

So we claim that the return probability

$$P(S_{2n} = 0) = \binom{2n}{n}(\frac{1}{2})^n(\frac{1}{2})^n = \frac{(2n)!}{n!n!}(\frac{1}{2})^{2n}$$

Stirling's formula states that as $n \to \infty$,

$$n! \sim \sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}.$$

Then $$P(S_{2n} = 0) = \frac{(2n)!}{n!n!}(\frac{1}{2})^{2n} \sim \frac{\sqrt{2}}{\sqrt{2\pi}n^{1/2}} = \frac{C_0}{n^{1/2}}.$$

Simple random walks-one dimension

- Define V to be a random variable that denotes the number of time the walker returns to 0, then

$$V = \sum_{n=0}^{\infty} I\{S_{2n} = 0\}$$

(where I{A} is an indicator function)

- Consider the mean of the number of visits

$$E(V) = \sum_{n=0}^{\infty} E(I\{S_{2n} = 0\}) = 1 + \sum_{n=1}^{\infty} P(S_{2n} = 0) = 1 + \sum_{n=1}^{\infty} \frac{\sqrt{2}}{\sqrt{2\pi}} n^{-\frac{1}{2}}$$
$$= 1 + \frac{\sqrt{2}}{\sqrt{2\pi}} \sum_{n=1}^{\infty} n^{-\frac{1}{2}} = \infty$$

(Recall that the sum $\sum_{n=1}^{\infty} n^{-\frac{1}{2}}$ diverges since $\frac{1}{2} < 1$.)

If we let **q** = P(the walker ever return to 0), then we can show that **q** = 1 by supposing **q** < 1, and draw contradiction that E(V) will actually be finite.

Simple random walks-higher dimensions

- What will happen if the random walker takes action in higher dimensions, say Z^d?
  - In each direction, the random walks will be performed as in one dimension
  - In 2n steps, we expect (2n/d) steps to be taken in each of the d-directions

$$P(\text{any particular integer}) \sim \frac{c_d}{n^{d/2}}$$

  - Return to origin:

Since $P(S_n = 0) \sim \frac{c_d}{n^{d/2}}$

$$E(V) = \sum_{2n=0}^{\infty} P(S_n = 0) \sim \sum_{n=0}^{\infty} \frac{c_d}{n^{d/2}} = \begin{cases} < \infty, d \geq 3 \\ = \infty, d = 1,2 \end{cases}$$

- The results correspond to the facts that if the Markov chain is a simple symmetric on Z^2, all states are recurrent; if it's on Z^d, $d \geq 3$, all states are transient.

References:

- Cameron, M. (n.d.). *Discrete time Markov chains.*
- Lawler, G. F. (2011). *Random walk and the heat equation*. Providence, RI: American Mathematical Soc.
- Cairns, H. (2014). *A short proof of Perron's theorem.*