# Markov Chains

Gonzalo Mateos
Dept. of ECE and Goergen Institute for Data Science
University of Rochester
gmateosb@ece.rochester.edu
http://www.ece.rochester.edu/~gmateosb/

September 23, 2020

# Markov chains

# Markov chains in discrete time

- Consider discrete-time index $n = 0, 1, 2, \ldots$
- Time-dependent random state $X_n$ takes values on a countable set
  - In general, states are $i = 0, \pm 1, \pm 2, \ldots$, i.e., here the state space is $\mathbb{Z}$
  - If $X_n = i$ we say "the process is in state $i$ at time $n$"
- Random process is $X_{\mathbb{N}}$, its history up to $n$ is $\mathbf{X}_n = [X_n, X_{n-1}, \ldots, X_0]^T$
- **Def:** process $X_{\mathbb{N}}$ is a Markov chain (MC) if for all $n \geq 1$, $i, j$, $\mathbf{x} \in \mathbb{Z}^n$

$$\mathrm{P}\left(X_{n+1} = j \,\middle|\, X_n = i, \mathbf{X}_{n-1} = \mathbf{x}\right) = \mathrm{P}\left(X_{n+1} = j \,\middle|\, X_n = i\right) = P_{ij}$$

- Future depends only on current state $X_n$ (memoryless, Markov property)

  $\Rightarrow$ Future conditionally independent of the past, given the present

# Observations

- Given $X_n$, history $\mathbf{X}_{n-1}$ irrelevant for future evolution of the process
- From the Markov property, can show that for arbitrary $m > 0$

$$\mathrm{P}\left(X_{n+m} = j \,\middle|\, X_n = i, \mathbf{X}_{n-1} = \mathbf{x}\right) = \mathrm{P}\left(X_{n+m} = j \,\middle|\, X_n = i\right)$$

- Transition probabilities $P_{ij}$ are constant (MC is time invariant)

$$\mathrm{P}\left(X_{n+1} = j \,\middle|\, X_n = i\right) = \mathrm{P}\left(X_1 = j \,\middle|\, X_0 = i\right) = P_{ij}$$

- Since $P_{ij}$'s are probabilities they are non-negative and sum up to 1

$$P_{ij} \geq 0, \qquad \sum_{j=0}^{\infty} P_{ij} = 1$$

$\Rightarrow$ Conditional probabilities satisfy the axioms

# Matrix representation

- Group the $P_{ij}$ in a transition probability "matrix" $\mathbf{P}$

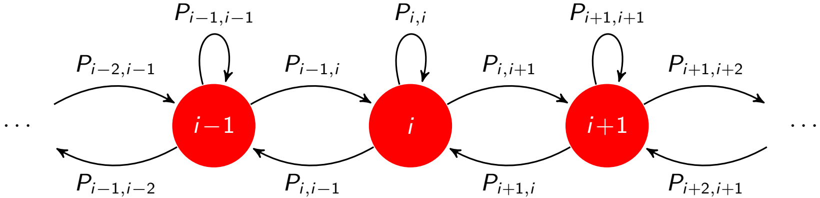
$$\mathbf{P} = \begin{pmatrix} P_{00} & P_{01} & P_{02} & \ldots & P_{0j} & \ldots \\ P_{10} & P_{11} & P_{12} & \ldots & P_{1j} & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{i0} & P_{i1} & P_{i2} & \ldots & P_{ij} & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$\Rightarrow$ Not really a matrix if number of states is infinite

- Row-wise sums should be equal to one, i.e., $\sum_{j=0}^{\infty} P_{ij} = 1$ for all $i$

# Graph representation

- A graph representation or state transition diagram is also used




- Useful when number of states is infinite, skip arrows if $P_{ij} = 0$
- Again, sum of per-state outgoing arrow weights should be one

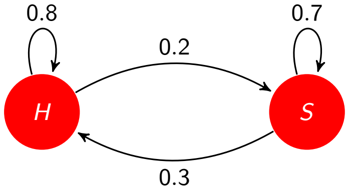
# Example: Happy - Sad

- I can be happy ($X_n = 0$) or sad ($X_n = 1$)

  $\Rightarrow$ My mood tomorrow is only affected by my mood today

- Model as Markov chain with transition probabilities

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}$$

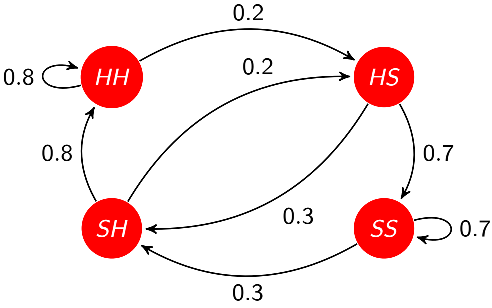



- Inertia $\Rightarrow$ happy or sad today, likely to stay happy or sad tomorrow
- But when sad, a little less likely so ($P_{00} > P_{11}$)

# Example: Happy - Sad with memory

- Happiness tomorrow affected by today's and yesterday's mood
  - $\Rightarrow$ Not a Markov chain with the previous state space
- Define double states HH (Happy-Happy), HS (Happy-Sad), SH, SS
- Only some transitions are possible
  - HH and SH can only become HH or HS
  - HS and SS can only become SH or SS

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.3 & 0.7 \\ 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.3 & 0.7 \end{pmatrix}$$

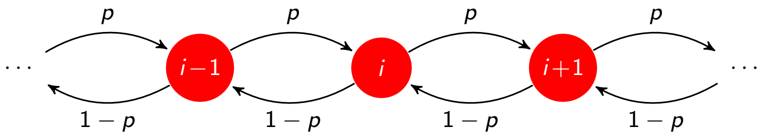



- **Key:** can capture longer time memory via state augmentation

# Random (drunkard's) walk

- Step to the right w.p. $p$, to the left w.p. $1-p$

  $\Rightarrow$ Not that drunk to stay on the same place



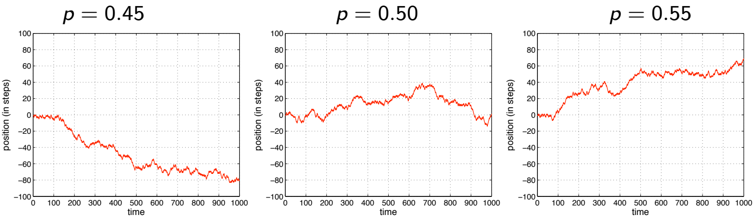

- States are $0, \pm 1, \pm 2, \dots$ (state space is $\mathbb{Z}$), infinite number of states
- Transition probabilities are

$$P_{i,i+1} = p, \qquad P_{i,i-1} = 1-p$$

- $P_{ij} = 0$ for all other transitions

## Random (drunkard's) walk (continued)

- Random walks behave differently if $p < 1/2$, $p = 1/2$ or $p > 1/2$




$\Rightarrow$ With $p > 1/2$ diverges to the right ($\nearrow$ almost surely)

$\Rightarrow$ With $p < 1/2$ diverges to the left ($\searrow$ almost surely)

$\Rightarrow$ With $p = 1/2$ always come back to visit origin (almost surely)

- Because number of states is infinite we can have all states transient
  - Transient states not revisited after some time (more later)

# Two dimensional random walk

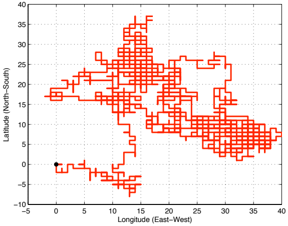
- Take a step in random direction E, W, S or N
  - $\Rightarrow$ E, W, S, N chosen with equal probability
- States are pairs of coordinates $(X_n, Y_n)$
  - $X_n = 0, \pm 1, \pm 2, \ldots$ and $Y_n = 0, \pm 1, \pm 2, \ldots$
- Transiton probs. $\neq 0$ only for adjacent points

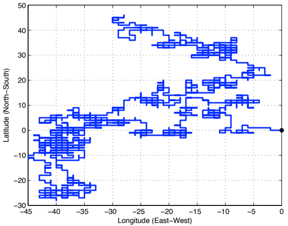
East: $P\left(X_{n+1} = i+1, Y_{n+1} = j \,\middle|\, X_n = i, Y_n = j\right) = \dfrac{1}{4}$

West: $P\left(X_{n+1} = i-1, Y_{n+1} = j \,\middle|\, X_n = i, Y_n = j\right) = \dfrac{1}{4}$

North: $P\left(X_{n+1} = i, Y_{n+1} = j+1 \,\middle|\, X_n = i, Y_n = j\right) = \dfrac{1}{4}$

South: $P\left(X_{n+1} = i, Y_{n+1} = j-1 \,\middle|\, X_n = i, Y_n = j\right) = \dfrac{1}{4}$



# More about random walks

- Some random facts of life for equiprobable random walks
- In one and two dimensions probability of returning to origin is 1
  - $\Rightarrow$ Will almost surely return home
- In more than two dimensions, probability of returning to origin is $< 1$
  - $\Rightarrow$ In three dimensions probability of returning to origin is 0.34
  - $\Rightarrow$ Then 0.19, 0.14, 0.10, 0.08, ...

# Another representation of a random walk

- Consider an i.i.d. sequence of RVs $Y_{\mathbb{N}} = Y_1, Y_2, \ldots, Y_n, \ldots$
- $Y_n$ takes the value $\pm 1$, $\mathrm{P}\left(Y_n = 1\right) = p$, $\mathrm{P}\left(Y_n = -1\right) = 1 - p$

- Define $X_0 = 0$ and the cumulative sum

$$X_n = \sum_{k=1}^{n} Y_k$$

$\Rightarrow$ The process $X_{\mathbb{N}}$ is a random walk (same we saw earlier)

$\Rightarrow$ $Y_{\mathbb{N}}$ are i.i.d. steps (increments) because $X_n = X_{n-1} + Y_n$

- Q: Can we formally establish the random walk is a Markov chain?
- A: Since $X_n = X_{n-1} + Y_n$, $n \geq 1$, and $Y_n$ independent of $\mathbf{X}_{n-1}$

$$\begin{aligned}\mathrm{P}\left(X_n = j \,\middle|\, X_{n-1} = i, \mathbf{X}_{n-2} = \mathbf{x}\right) &= \mathrm{P}\left(X_{n-1} + Y_n = j \,\middle|\, X_{n-1} = i, \mathbf{X}_{n-2} = \mathbf{x}\right)\\ &= \mathrm{P}\left(Y_1 = j - i\right) := P_{ij}\end{aligned}$$

# General result to identify Markov chains

**Theorem**

*Suppose $Y_{\mathbb{N}} = Y_1, Y_2, \ldots, Y_n, \ldots$ are i.i.d. and independent of $X_0$. Consider the random process $X_{\mathbb{N}} = X_1, X_2, \ldots, X_n, \ldots$ of the form*

$$X_n = f(X_{n-1}, Y_n), \quad n \geq 1$$

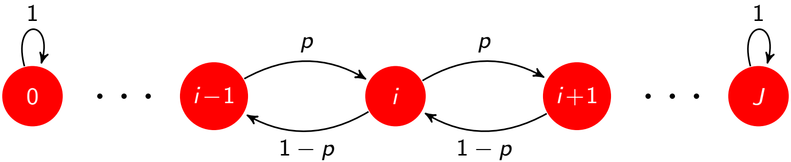
*Then $X_{\mathbb{N}}$ is a Markov chain with transition probabilities*

$$P_{ij} = \mathrm{P}\left(f(i, Y_1) = j\right)$$

- Useful result to identify Markov chains
  - $\Rightarrow$ Often simpler than checking the Markov property
- Proof similar to the random walk special case, i.e., $f(x, y) = x + y$

# Random walk with boundaries (gambling)

- As a random walk, but stop moving when $X_n = 0$ or $X_n = J$
  - Models a gambler that stops playing when ruined, $X_n = 0$
  - Or when reaches target gains $X_n = J$




- States are $0, 1, \ldots, J$, finite number of states
- Transition probabilities are

$$P_{i,i+1} = p, \quad P_{i,i-1} = 1 - p, \qquad P_{00} = 1, \quad P_{JJ} = 1$$

- $P_{ij} = 0$ for all other transitions
- States 0 and $J$ are called absorbing. Once there stay there forever

  $\Rightarrow$ The rest are transient states. Visits stop almost surely

# Chapter-Kolmogorov equations

# Multiple-step transition probabilities

- Q: What can be said about multiple transitions?
- Ex: Transition probabilities between two time slots

$$P_{ij}^{2} = \mathrm{P}\left(X_{m+2} = j \,\middle|\, X_m = i\right)$$

$\Rightarrow$ Caution: $P_{ij}^2$ is just notation, $P_{ij}^2 \neq P_{ij} \times P_{ij}$

- Ex: Probabilities of $X_{m+n}$ given $X_m$ $\Rightarrow$ $n$-step transition probabilities

$$P_{ij}^{n} = \mathrm{P}\left(X_{m+n} = j \,\middle|\, X_m = i\right)$$

- Relation between $n$-, $m$-, and $(m+n)$-step transition probabilities

  $\Rightarrow$ Write $P_{ij}^{m+n}$ in terms of $P_{ij}^{m}$ and $P_{ij}^{n}$

- All questions answered by Chapman-Kolmogorov's equations

# 2-step transition probabilities

- Start considering transition probabilities between two time slots

$$P_{ij}^2 = \mathrm{P}\left(X_{n+2} = j \,\middle|\, X_n = i\right)$$

- Using the law of total probability

$$P_{ij}^2 = \sum_{k=0}^{\infty} \mathrm{P}\left(X_{n+2} = j \,\middle|\, X_{n+1} = k, X_n = i\right) \mathrm{P}\left(X_{n+1} = k \,\middle|\, X_n = i\right)$$

- In the first probability, conditioning on $X_n = i$ is unnecessary. Thus

$$P_{ij}^2 = \sum_{k=0}^{\infty} \mathrm{P}\left(X_{n+2} = j \,\middle|\, X_{n+1} = k\right) \mathrm{P}\left(X_{n+1} = k \,\middle|\, X_n = i\right)$$

- Which by definition of transition probabilities yields

$$P_{ij}^2 = \sum_{k=0}^{\infty} P_{kj} P_{ik}$$

# Relating $n$-, $m$-, and $(m+n)$-step probabilities

- Same argument works (condition on $X_0$ w.l.o.g., time invariance)

$$P_{ij}^{m+n} = \mathrm{P}\left(X_{n+m} = j \,\middle|\, X_0 = i\right)$$

- Use law of total probability, drop unnecessary conditioning and use definitions of $n$-step and $m$-step transition probabilities

$$P_{ij}^{m+n} = \sum_{k=0}^{\infty} \mathrm{P}\left(X_{m+n} = j \,\middle|\, X_m = k, X_0 = i\right) \mathrm{P}\left(X_m = k \,\middle|\, X_0 = i\right)$$

$$P_{ij}^{m+n} = \sum_{k=0}^{\infty} \mathrm{P}\left(X_{m+n} = j \,\middle|\, X_m = k\right) \mathrm{P}\left(X_m = k \,\middle|\, X_0 = i\right)$$

$$P_{ij}^{m+n} = \sum_{k=0}^{\infty} P_{kj}^{n} P_{ik}^{m} \quad \text{for all } i, j \text{ and } n, m \geq 0$$

$\Rightarrow$ These are the Chapman-Kolmogorov equations

# Interpretation

- Chapman-Kolmogorov equations are intuitive. Recall

$$P_{ij}^{m+n} = \sum_{k=0}^{\infty} P_{ik}^{m} P_{kj}^{n}$$

- Between times 0 and $m+n$, time $m$ occurred
- At time $m$, the Markov chain is in some state $X_m = k$

  $\Rightarrow$ $P_{ik}^{m}$ is the probability of going from $X_0 = i$ to $X_m = k$

  $\Rightarrow$ $P_{kj}^{n}$ is the probability of going from $X_m = k$ to $X_{m+n} = j$

  $\Rightarrow$ Product $P_{ik}^{m} P_{kj}^{n}$ is then the probability of going from $X_0 = i$ to $X_{m+n} = j$ passing through $X_m = k$ at time $m$
- Since any $k$ might have occurred, just sum over all $k$

# Chapman-Kolmogorov equations in matrix form

- Define the following three matrices:
  - $\Rightarrow$ $\mathbf{P}^{(m)}$ with elements $P_{ij}^{m}$
  - $\Rightarrow$ $\mathbf{P}^{(n)}$ with elements $P_{ij}^{n}$
  - $\Rightarrow$ $\mathbf{P}^{(m+n)}$ with elements $P_{ij}^{m+n}$
- Matrix product $\mathbf{P}^{(m)}\mathbf{P}^{(n)}$ has $(i,j)$-th element $\sum_{k=0}^{\infty} P_{ik}^{m} P_{kj}^{n}$
- Chapman Kolmogorov in matrix form

$$\mathbf{P}^{(m+n)} = \mathbf{P}^{(m)}\mathbf{P}^{(n)}$$

- Matrix of $(m+n)$-step transitions is product of $m$-step and $n$-step

# Computing $n$-step transition probabilities

- For $m = n = 1$ (2-step transition probabilities) matrix form is

$$\mathbf{P}^{(2)} = \mathbf{P}\mathbf{P} = \mathbf{P}^2$$

- Proceed recursively backwards from $n$

$$\mathbf{P}^{(n)} = \mathbf{P}^{(n-1)}\mathbf{P} = \mathbf{P}^{(n-2)}\mathbf{P}\mathbf{P} = \ldots = \mathbf{P}^n$$

- Have proved the following

**Theorem**

*The matrix of n-step transition probabilities $\mathbf{P}^{(n)}$ is given by the n-th power of the transition probability matrix $\mathbf{P}$, i.e.,*

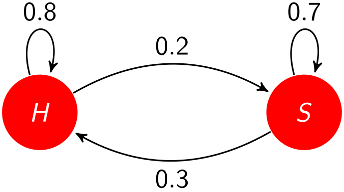
$$\mathbf{P}^{(n)} = \mathbf{P}^n$$

*Henceforth we write $\mathbf{P}^n$*

# Example: Happy-Sad

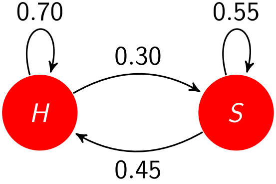
- Mood transitions in one day

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}$$




- Transition probabilities between today and the day after tomorrow?

$$\mathbf{P}^2 = \begin{pmatrix} 0.70 & 0.30 \\ 0.45 & 0.55 \end{pmatrix}$$



# Example: Happy-Sad (continued)

- ... After a week and after a month

$$\mathbf{P}^7 = \begin{pmatrix} 0.6031 & 0.3969 \\ 0.5953 & 0.4047 \end{pmatrix} \qquad \mathbf{P}^{30} = \begin{pmatrix} 0.6000 & 0.4000 \\ 0.6000 & 0.4000 \end{pmatrix}$$

- Matrices $\mathbf{P}^7$ and $\mathbf{P}^{30}$ almost identical $\Rightarrow \lim_{n\to\infty} \mathbf{P}^n$ exists

  $\Rightarrow$ Note that this is a regular limit

- After a month transition from H to H and from S to H w.p. 0.6

  $\Rightarrow$ State becomes independent of initial condition (H w.p. 0.6)

- Rationale: 1-step memory $\Rightarrow$ Initial condition eventually forgotten
  - More about this soon

# Unconditional probabilities

- All probabilities so far are conditional, i.e., $P_{ij}^n = \mathrm{P}\left(X_n = j \,\middle|\, X_0 = i\right)$

  $\Rightarrow$ May want unconditional probabilities $p_j(n) = \mathrm{P}\left(X_n = j\right)$

- Requires specification of initial conditions $p_i(0) = \mathrm{P}\left(X_0 = i\right)$

- Using law of total probability and definitions of $P_{ij}^n$ and $p_j(n)$

$$p_j(n) = \mathrm{P}\left(X_n = j\right) = \sum_{i=0}^{\infty} \mathrm{P}\left(X_n = j \,\middle|\, X_0 = i\right) \mathrm{P}\left(X_0 = i\right)$$

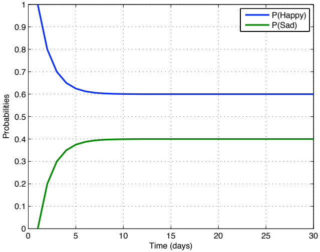
$$= \sum_{i=0}^{\infty} P_{ij}^n p_i(0)$$

- In matrix form (define vector $\mathbf{p}(n) = [p_1(n), p_2(n), \ldots]^T$)

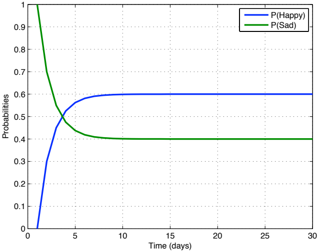
$$\mathbf{p}(n) = \left(\mathbf{P}^n\right)^T \mathbf{p}(0)$$

- Transition probability matrix $\Rightarrow$ $\mathbf{P} = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}$

$\mathbf{p}(0) = [1, 0]^T$




$\mathbf{p}(0) = [0, 1]^T$




- For large $n$ probabilities $\mathbf{p}(n)$ are independent of initial state $\mathbf{p}(0)$

# Gambler's ruin problem

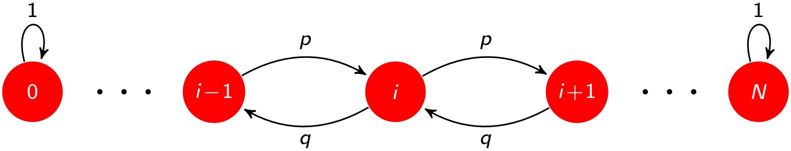
# Gambler's ruin problem

- You place \$1 bets
  (i) With probability $p$ you gain \$1, and
  (ii) With probability $q = 1 - p$ you loose your \$1 bet
- Start with an initial wealth of \$$i$
- Define bias factor $\alpha := q/p$
  - If $\alpha > 1$ more likely to loose than win (biased against gambler)
  - $\alpha < 1$ favors gambler (more likely to win than loose)
  - $\alpha = 1$ game is fair
- You keep playing until
  (a) You go broke (loose all your money)
  (b) You reach a wealth of \$$N$ (same as first lecture, HW1 for $N \to \infty$)
- Prob. $S_i$ of reaching \$$N$ before going broke for initial wealth \$$i$?
  - $S$ stands for success, or successful betting run (SBR)

# Gambler's Markov chain

- Model wealth as Markov chain $X_{\mathbb{N}}$. Transition probabilities

$$P_{i,i+1} = p, \quad P_{i,i-1} = q, \quad P_{00} = P_{NN} = 1$$




- Realizations $x_{\mathbb{N}}$. Initial state = Initial wealth = $i$
  - $\Rightarrow$ Sates 0 and $N$ are absorbing. Eventually end up in one of them
  - $\Rightarrow$ Remaining states are transient (visits eventually stop)
- Being absorbing states says something about the limit wealth

$$\lim_{n\to\infty} x_n = 0, \text{ or } \lim_{n\to\infty} x_n = N \quad \Rightarrow \quad S_i := \mathrm{P}\left(\lim_{n\to\infty} X_n = N \,\middle|\, X_0 = i\right)$$

# Recursive relations

- Total probability to relate $S_i$ with $S_{i+1}, S_{i-1}$ from adjacent states
  - $\Rightarrow$ Condition on first bet $X_1$, Markov chain homogeneous

$$S_i = S_{i+1}P_{i,i+1} + S_{i-1}P_{i,i-1} = S_{i+1}p + S_{i-1}q$$

- Recall $p + q = 1$ and reorder terms

$$p(S_{i+1} - S_i) = q(S_i - S_{i-1})$$

- Recall definition of bias $\alpha = q/p$

$$S_{i+1} - S_i = \alpha(S_i - S_{i-1})$$

# Recursive relations (continued)

- If current state is 0 then $S_i = S_0 = 0$. Can write

  $$S_2 - S_1 = \alpha(S_1 - S_0) = \alpha S_1$$

- Substitute this in the expression for $S_3 - S_2$

  $$S_3 - S_2 = \alpha(S_2 - S_1) = \alpha^2 S_1$$

- Apply recursively backwards from $S_i - S_{i-1}$

  $$S_i - S_{i-1} = \alpha(S_{i-1} - S_{i-2}) = \ldots = \alpha^{i-1} S_1$$

- Sum up all of the former to obtain

  $$S_i - S_1 = S_1 \left( \alpha + \alpha^2 + \ldots + \alpha^{i-1} \right)$$

- The latter can be written as a geometric series

  $$S_i = S_1 \left( 1 + \alpha + \alpha^2 + \ldots + \alpha^{i-1} \right)$$

# Probability of successful betting run

- Geometric series can be summed in closed form, assuming $\alpha \neq 1$

$$S_i = \left(\sum_{k=0}^{i-1} \alpha^k\right) S_1 = \frac{1-\alpha^i}{1-\alpha} S_1$$

- When in state $N$, $S_N = 1$ and so

$$1 = S_N = \frac{1-\alpha^N}{1-\alpha} S_1 \Rightarrow S_1 = \frac{1-\alpha}{1-\alpha^N}$$

- Substitute $S_1$ above into expression for probability of SBR $S_i$

$$S_i = \frac{1-\alpha^i}{1-\alpha^N}, \quad \alpha \neq 1$$

- For $\alpha = 1$ $\Rightarrow S_i = iS_1$, $1 = S_N = NS_1$, $\Rightarrow S_i = \frac{i}{N}$

# Analysis for large $N$

- Recall

$$S_i = \begin{cases} (1-\alpha^i)/(1-\alpha^N), & \alpha \neq 1, \\ i/N, & \alpha = 1 \end{cases}$$

- Consider exit bound $N$ arbitrarily large

(i) For $\alpha > 1$, $S_i \approx (\alpha^i - 1)/\alpha^N \to 0$

(ii) Likewise for $\alpha = 1$, $S_i = i/N \to 0$

- If win probability $p$ does not exceed loose probability $q$

  $\Rightarrow$ Will almost surely loose all money

(iii) For $\alpha < 1$, $S_i \to 1 - \alpha^i$

- If win probability $p$ exceeds loose probability $q$

  $\Rightarrow$ For sufficiently high initial wealth $i$, will most likely win

- This explains what we saw on first lecture and HW1

# Queues in communication systems

# Queues in communication systems

- General communication systems goal

  $\Rightarrow$ Move packets from generating sources to intended destinations

- Between arrival and departure we hold packets in a memory buffer

  $\Rightarrow$ Want to design buffers appropriately

# Non-concurrent queue

- Time slotted in intervals of duration $\Delta t$
  $\Rightarrow$ $n$-th slot between times $n\Delta t$ and $(n+1)\Delta t$
- Average arrival rate is $\bar{\lambda}$ packets per unit time
  $\Rightarrow$ Probability of packet arrival in $\Delta t$ is $\lambda = \bar{\lambda}\Delta t$
- Packets are transmitted (depart) at a rate of $\bar{\mu}$ packets per unit time
  $\Rightarrow$ Probability of packet departure in $\Delta t$ is $\mu = \bar{\mu}\Delta t$
- Assume no simultaneous arrival and departure (no concurrence)
  $\Rightarrow$ Reasonable for small $\Delta t$ ($\mu$ and $\lambda$ likely to be small)

# Queue evolution equations

- $Q_n$ denotes number of packets in queue (backlog) in $n$-th time slot
- $\mathbb{A}_n$ = nr. of packet arrivals, $\mathbb{D}_n$ = nr. of departures (during $n$-th slot)

- If the queue is empty $Q_n = 0$ then there are no departures

  $\Rightarrow$ Queue length at time $n+1$ can be written as

$$Q_{n+1} = Q_n + \mathbb{A}_n, \quad \text{if } Q_n = 0$$

- If $Q_n > 0$, departures and arrivals may happen

$$Q_{n+1} = Q_n + \mathbb{A}_n - \mathbb{D}_n, \quad \text{if } Q_n > 0$$

- $\mathbb{A}_n \in \{0, 1\}$, $\mathbb{D}_n \in \{0, 1\}$ and either $\mathbb{A}_n = 1$ or $\mathbb{D}_n = 1$ but not both

  $\Rightarrow$ Arrival and departure probabilities are

$$\mathrm{P}\left(\mathbb{A}_n = 1\right) = \lambda, \qquad \mathrm{P}\left(\mathbb{D}_n = 1\right) = \mu$$

# Queue evolution probabilities

- Future queue lengths depend on current length only
- Probability of queue length increasing

$$\mathrm{P}\left(Q_{n+1} = i+1 \,\middle|\, Q_n = i\right) = \mathrm{P}\left(\mathbb{A}_n = 1\right) = \lambda, \qquad \text{for all } i$$

- Queue length might decrease only if $Q_n > 0$. Probability is

$$\mathrm{P}\left(Q_{n+1} = i-1 \,\middle|\, Q_n = i\right) = \mathrm{P}\left(\mathbb{D}_n = 1\right) = \mu, \qquad \text{for all } i > 0$$

- Queue length stays the same if it neither increases nor decreases

$$\mathrm{P}\left(Q_{n+1} = i \,\middle|\, Q_n = i\right) = 1 - \lambda - \mu, \qquad \text{for all } i > 0$$
$$\mathrm{P}\left(Q_{n+1} = 0 \,\middle|\, Q_n = 0\right) = 1 - \lambda$$

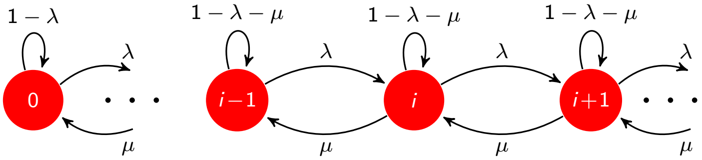
$\Rightarrow$ No departures when $Q_n = 0$ explain second equation

# Queue as a Markov chain

- MC with states $0, 1, 2, \ldots$. Identify states with queue lengths
- Transition probabilities for $i \neq 0$ are

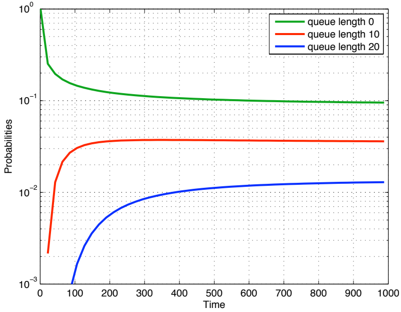
$$P_{i,i-1} = \mu, \qquad P_{i,i} = 1 - \lambda - \mu, \qquad P_{i,i+1} = \lambda$$

- For $i = 0$: $P_{00} = 1 - \lambda$ and $P_{01} = \lambda$



# Numerical example: Probability propagation

- Build matrix $\mathbf{P}$ truncating at maximum queue length $L = 100$
  - $\Rightarrow$ Arrival rate $\lambda = 0.3$. Departure rate $\mu = 0.33$
  - $\Rightarrow$ Initial distribution $\mathbf{p}(0) = [1, 0, 0, \ldots]^T$ (queue empty)




- Propagate probabilities $(\mathbf{P}^n)^T \mathbf{p}(0)$
- Probabilities obtained are

$$\mathrm{P}\left(Q_n = i \,\middle|\, Q_0 = 0\right) = p_i(n)$$

- A few $i$'s (0, 10, 20) shown
- Probability of empty queue $\approx 0.1$
- Occupancy decreases with $i$

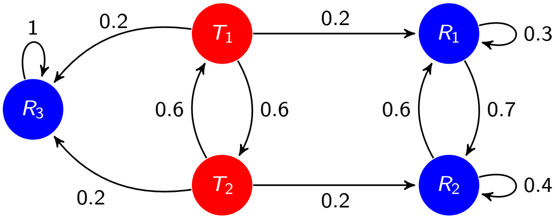
# Classes of states

# Transient and recurrent states

- States of a MC can be recurrent or transient

- Transient states might be visited early on but visits eventually stop
- Almost surely, $X_n \neq i$ for $n$ sufficiently large (qualifications needed)

- Visits to recurrent states keep happening forever. Fix arbitrary $m$
- Almost surely, $X_n = i$ for some $n \geq m$ (qualifications needed)



# Definitions

- Let $f_i$ be the probability that starting at $i$, MC ever reenters state $i$

$$f_i := \mathrm{P}\left(\bigcup_{n=1}^{\infty} X_n = i \,\middle|\, X_0 = i\right) = \mathrm{P}\left(\bigcup_{n=m+1}^{\infty} X_n = i \,\middle|\, X_m = i\right)$$

- State $i$ is recurrent if $f_i = 1$

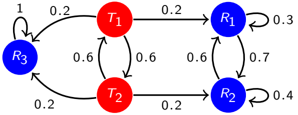
  $\Rightarrow$ Process reenters $i$ again and again (a.s.). Infinitely often

- State $i$ is transient if $f_i < 1$

  $\Rightarrow$ Positive probability $1 - f_i > 0$ of never coming back to $i$

# Recurrent states example

- State $R_3$ is recurrent because it is absorbing $\mathrm{P}\left(X_1 = R_3 \,\middle|\, X_0 = R_3\right) = 1$



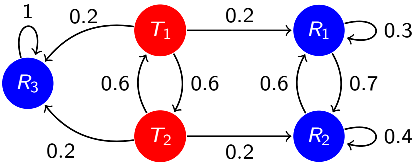

- State $R_1$ is recurrent because

$$\mathrm{P}\left(X_1 = R_1 \,\middle|\, X_0 = R_1\right) = 0.3$$
$$\mathrm{P}\left(X_2 = R_1, X_1 \neq R_1 \,\middle|\, X_0 = R_1\right) = (0.7)(0.6)$$
$$\mathrm{P}\left(X_3 = R_1, X_2 \neq R_1, X_1 \neq R_1 \,\middle|\, X_0 = R_1\right) = (0.7)(0.4)(0.6)$$
$$\vdots$$
$$\mathrm{P}\left(X_n = R_1, X_{n-1} \neq R_1, \ldots, X_1 \neq R_1 \,\middle|\, X_0 = R_1\right) = (0.7)(0.4)^{n-2}(0.6)$$

- Sum up: $$f_i = \sum_{n=1}^{\infty} \mathrm{P}\left(X_n = R_1, X_{n-1} \neq R_1, \ldots, X_1 \neq R_1 \,\middle|\, X_0 = R_1\right)$$
$$= 0.3 + 0.7\left(\sum_{n=2}^{\infty} 0.4^{n-2}\right)0.6 = 0.3 + 0.7\left(\frac{1}{1-0.4}\right)0.6 = 1$$

# Transient state example

- States $T_1$ and $T_2$ are transient
- Probability of returning to $T_1$ is $f_{T_1} = (0.6)^2 = 0.36$
  - $\Rightarrow$ Might come back to $T_1$ only if it goes to $T_2$ (w.p. 0.6)
  - $\Rightarrow$ Will come back only if it moves back from $T_2$ to $T_1$ (w.p. 0.6)




- Likewise, $f_{T_2} = (0.6)^2 = 0.36$

# Expected number of visits to states

- Define $N_i$ as the number of visits to state $i$ given that $X_0 = i$

$$N_i := \sum_{n=1}^{\infty} \mathbb{I}\left\{X_n = i \,\middle|\, X_0 = i\right\}$$

- If $X_n = i$, this is the last visit to $i$ w.p. $1 - f_i$
- Prob. revisiting state $i$ exactly $n$ times is ($n$ visits $\times$ no more visits)

$$\mathrm{P}\left(N_i = n\right) = f_i^n(1 - f_i)$$

$\Rightarrow$ Number of visits $N_i + 1$ is geometric with parameter $1 - f_i$

- Expected number of visits is

$$\mathbb{E}\left[N_i\right] + 1 = \frac{1}{1 - f_i} \;\Rightarrow\; \mathbb{E}\left[N_i\right] = \frac{f_i}{1 - f_i}$$

$\Rightarrow$ For recurrent states $N_i = \infty$ a.s. and $\mathbb{E}\left[N_i\right] = \infty$ $(f_i = 1)$

# Alternative transience/recurrence characterization

- Another way of writing $\mathbb{E}[N_i]$

$$\mathbb{E}[N_i] = \sum_{n=1}^{\infty} \mathbb{E}\Big[\mathbb{I}\left\{X_n = i \,\middle|\, X_0 = i\right\}\Big] = \sum_{n=1}^{\infty} P_{ii}^n$$

- Recall that: for transient states $\mathbb{E}[N_i] = f_i/(1-f_i) < \infty$
  for recurrent states $\mathbb{E}[N_i] = \infty$

Theorem

- *State $i$ is transient if and only if $\sum_{n=1}^{\infty} P_{ii}^n < \infty$*
- *State $i$ is recurrent if and only if $\sum_{n=1}^{\infty} P_{ii}^n = \infty$*

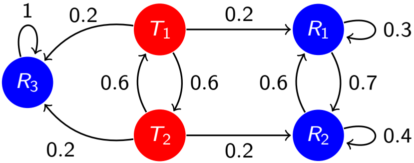
- Number of future visits to transient states is finite

  $\Rightarrow$ If number of states is finite some states have to be recurrent

- **Def:** State $j$ is accessible from state $i$ if $P_{ij}^n > 0$ for some $n \geq 0$

  $\Rightarrow$ It is possible to enter $j$ if MC initialized at $X_0 = i$

- Since $P_{ii}^0 = \mathrm{P}\left(X_0 = i \,\middle|\, X_0 = i\right) = 1$, state $i$ is accessible from itself




- All states accessible from $T_1$ and $T_2$
- Only $R_1$ and $R_2$ accessible from $R_1$ or $R_2$
- None other than $R_3$ accessible from itself

# Communication

- **Def:** States $i$ and $j$ are said to communicate ($i \leftrightarrow j$) if

  $\Rightarrow$ $j$ is accessible from $i$, i.e., $P_{ij}^n > 0$ for some $n$; and

  $\Rightarrow$ $i$ is accessible from $j$, i.e., $P_{ji}^m > 0$ for some $m$

- Communication is an equivalence relation

- Reflexivity: $i \leftrightarrow i$
  - Holds because $P_{ii}^0 = 1$
- Symmetry: If $i \leftrightarrow j$ then $j \leftrightarrow i$
  - If $i \leftrightarrow j$ then $P_{ij}^n > 0$ and $P_{ji}^m > 0$ from where $j \leftrightarrow i$
- Transitivity: If $i \leftrightarrow j$ and $j \leftrightarrow k$, then $i \leftrightarrow k$
  - Just notice that $P_{ik}^{n+m} \geq P_{ij}^n P_{jk}^m > 0$

- Partitions set of states into disjoint classes (as all equivalences do)

  $\Rightarrow$ What are these classes?

# Recurrence and communication

**Theorem**

*If state $i$ is recurrent and $i \leftrightarrow j$, then $j$ is recurrent*

**Proof.**

- If $i \leftrightarrow j$ then there are $l, m$ such that $P_{ji}^l > 0$ and $P_{ij}^m > 0$
- Then, for any $n$ we have

$$P_{jj}^{l+n+m} \geq P_{ji}^l P_{ii}^n P_{ij}^m$$

- Sum for all $n$. Note that since $i$ is recurrent $\sum_{n=1}^{\infty} P_{ii}^n = \infty$

$$\sum_{n=1}^{\infty} P_{jj}^{l+n+m} \geq \sum_{n=1}^{\infty} P_{ji}^l P_{ii}^n P_{ij}^m = P_{ji}^l \left( \sum_{n=1}^{\infty} P_{ii}^n \right) P_{ij}^m = \infty$$

$\Rightarrow$ Which implies $j$ is recurrent

$\square$

# Recurrence and transience are class properties

Corollary

*If state $i$ is transient and $i \leftrightarrow j$, then $j$ is transient*

Proof.

- If $j$ were recurrent, then $i$ would be recurrent from previous theorem

$\square$

- Recurrence is shared by elements of a communication class

  $\Rightarrow$ We say that recurrence is a class property

- Likewise, transience is also a class property

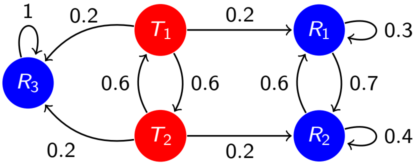
- MC states are separated in classes of transient and recurrent states

# Irreducible Markov chains

- A MC is called irreducible if it has only one class
  - All states communicate with each other
  - If MC also has finite number of states the single class is recurrent
  - If MC infinite, class might be transient
- When it has multiple classes (not irreducible)
  - Classes of transient states $\mathcal{T}_1, \mathcal{T}_2, \ldots$
  - Classes of recurrent states $\mathcal{R}_1, \mathcal{R}_2, \ldots$
- If MC initialized in a recurrent class $\mathcal{R}_k$, stays within the class
- If MC starts in transient class $\mathcal{T}_k$, then it might
  (a) Stay on $\mathcal{T}_k$ (only if $|\mathcal{T}_k| = \infty$)
  (b) End up in another transient class $\mathcal{T}_r$ (only if $|\mathcal{T}_r| = \infty$)
  (c) End up in a recurrent class $\mathcal{R}_l$
- For large time index $n$, MC restricted to one class

  $\Rightarrow$ Can be separated into irreducible components

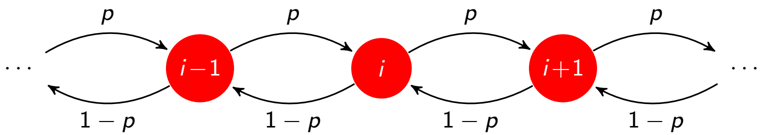
# Communication classes example




- Three classes
  - $\Rightarrow \mathcal{T} := \{T_1, T_2\}$, class with transient states
  - $\Rightarrow \mathcal{R}_1 := \{R_1, R_2\}$, class with recurrent states
  - $\Rightarrow \mathcal{R}_2 := \{R_3\}$, class with recurrent state
- For large $n$ suffices to study the irreducible components $\mathcal{R}_1$ and $\mathcal{R}_2$

# Example: Random walk

- Step right with probability $p$, left with probability $q = 1 - p$




- All states communicate $\Rightarrow$ States either all transient or all recurrent
- To see which, consider initially $X_0 = 0$ and note for any $n \geq 1$

$$P_{00}^{2n} = \binom{2n}{n} p^n q^n = \frac{(2n)!}{n!n!} p^n q^n$$

$\Rightarrow$ Back to 0 in $2n$ steps $\Leftrightarrow$ $n$ steps right and $n$ steps left

# Example: Random walk (continued)

- Stirling's formula $n! \approx n^n \sqrt{n} e^{-n} \sqrt{2\pi}$

  $\Rightarrow$ Approximate probability $P_{00}^{2n}$ of returning home as

$$P_{00}^{2n} = \frac{(2n)!}{n!n!} p^n q^n \approx \frac{(4pq)^n}{\sqrt{n\pi}}$$

- Symmetric random walk ($p = q = 1/2$)

$$\sum_{n=1}^{\infty} P_{00}^{2n} = \sum_{n=1}^{\infty} \frac{1}{\sqrt{n\pi}} = \infty$$

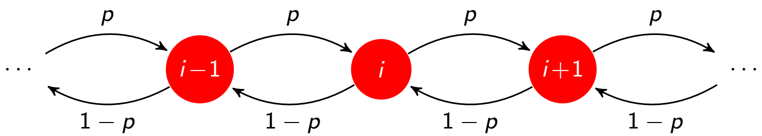
  $\Rightarrow$ State 0 (hence all states) are recurrent

- Biased random walk ($p > 1/2$ or $p < 1/2$), then $pq < 1/4$ and

$$\sum_{n=1}^{\infty} P_{00}^{2n} = \sum_{n=1}^{\infty} \frac{(4pq)^n}{\sqrt{n\pi}} < \infty$$

  $\Rightarrow$ State 0 (hence all states) are transient

- Alternative proof of transience of right-biased random walk ($p > 1/2$)




- Write current position of random walker as $X_n = \sum_{k=1}^{n} Y_k$

  $\Rightarrow$ $Y_k$ are the i.i.d. steps: $\mathbb{E}[Y_k] = 2p - 1$, $\text{var}[Y_k] = 4p(1-p)$

- From Central Limit Theorem ($\Phi(x)$ is cdf of standard Normal)

$$\mathrm{P}\left(\frac{\sum_{k=1}^{n} Y_k - n(2p-1)}{\sqrt{n4p(1-p)}} \leq a\right) \to \Phi(a)$$

# Example: Right-biased random walk (continued)

- Choose $a = \frac{\sqrt{n}(1-2p)}{\sqrt{4p(1-p)}} < 0$, use Chernoff bound $\Phi(a) \leq \exp(-a^2/2)$

$$\mathrm{P}\left(X_n \leq 0\right) = \mathrm{P}\left(\sum_{k=1}^{n} Y_k \leq 0\right) \to \Phi\left(\frac{\sqrt{n}(1-2p)}{\sqrt{4p(1-p)}}\right) < e^{-\frac{n(1-2p)^2}{8p(1-p)}} \to 0$$

- Since $P_{00}^n \leq \mathrm{P}\left(X_n \leq 0\right)$, sum over $n$

$$\sum_{n=1}^{\infty} P_{00}^n \leq \sum_{n=1}^{\infty} \mathrm{P}\left(X_n \leq 0\right) < \sum_{n=1}^{\infty} e^{-\frac{n(1-2p)^2}{8p(1-p)}} < \infty$$

- This establishes state 0 is transient

  $\Rightarrow$ Since all states communicate, all states are transient

# Take-home messages

- States of a MC can be transient or recurrent
- A MC can be partitioned into classes of communicating states
  - $\Rightarrow$ Class members are either all transient or all recurrent
  - $\Rightarrow$ Recurrence and transience are class properties
  - $\Rightarrow$ A finite MC has at least one recurrent class
- A MC with only one class is irreducible
  - $\Rightarrow$ If reducible it can be separated into irreducible components

# Glossary

- Markov chain
- State space
- Markov property
- Transition probability matrix
- State transition diagram
- State augmentation
- Random walk
- $n$-step transition probabilities
- Chapman-Kolmogorov eqs.
- Initial distribution
- Gambler's ruin problem
- Communication system
- Non-concurrent queue
- Queue evolution model
- Recurrent and transient states
- Accessibility
- Communication
- Equivalence relation
- Communication classes
- Class property
- Irreducible Markov chain
- Irreducible components