# ECE 586: Vector Space Methods
# Lecture 12 Flip Video: Markov Chains

Henry D. Pfister
Duke University

# Chutes and Ladders

## Example (Chutes and Ladders or Snakes and Ladders)

is a boardgame where a single die is rolled to determine how far you move on a gameboard defined by a grid. Some locations contain ladders that let you skip ahead while others contains chutes (or snakes) that set you back. It is based on an ancient game from India that teaches morality by associating ladders with virtues and snakes with vices.

Q: What is the chance a player finishes the game in *m* or fewer moves?

# Markov Chains

## Definition

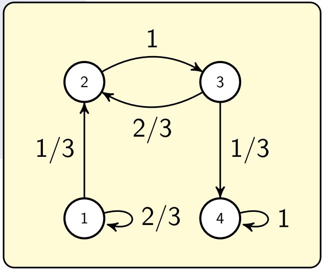
A finite-state Markov chain (FSMC) with $n$ states is a sequence of random variables $X_1, X_2, X_3, \ldots$ where each $X_i \in [n] \triangleq \{1, 2, \ldots, n\}$ and

$$\mathbb{P}\left(X_{t+1} = j \,|\, X_t = i, X_1, X_2, \ldots, X_{t-1}\right) = \mathbb{P}\left(X_{t+1} = j \,|\, X_t = i\right).$$

## Definition

If $\mathbb{P}(X_{t+1} = j \,|\, X_t = i)$ does not depend on $t$, then the Markov chain is called time invariant. For a time-invariant FSMC, let $P \in \mathbb{R}^{n \times n}$ denote the transition probability matrix with entries

$$[P]_{i,j} \triangleq P_{i,j} = \mathbb{P}(X_{t+1} = j \,|\, X_t = i).$$



Since each row of $P$ is a probability distribution, $P_{i,j} \geq 0$ and $\sum_{j=1}^{n} P_{i,j} = 1$. Matrices with this property are called stochastic

# Properties of Time-Invariant Markov Chains

## Multiple-Step Transition Probability

$$\begin{aligned}\mathbb{P}\left(X_{t+2}=j \,\middle|\, X_t=i\right) &= \sum_{k=1}^{n} \mathbb{P}\left(X_{t+2}=j \,\middle|\, X_{t+1}=k\right) \mathbb{P}\left(X_{t+1}=k \,\middle|\, X_t=i\right) \\ &= \sum_{k=1}^{n} P_{k,j} P_{i,k} = \sum_{k=1}^{n} P_{i,k} P_{k,j} = \left[P^2\right]_{i,j}\end{aligned}$$

By induction, it follows that $\mathbb{P}(X_{t+m}=j \,|\, X_t=i) = [P^m]_{i,j}$.

## Matrix-Vector Notation

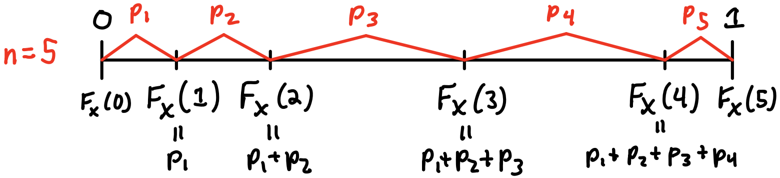
Using the notation $\underline{\pi}^{(t)} = (\pi_1^{(t)}, \ldots, \pi_n^{(t)})$ with $\pi_i^{(t)} \triangleq \mathbb{P}(X_t=i)$, we see that

$$\begin{aligned}\pi_j^{(t+1)} &= \sum_{i=1}^{n} \mathbb{P}(X_{t+1}=j, X_t=i) = \sum_{i=1}^{n} \mathbb{P}(X_{t+1}=j|X_t=i)\mathbb{P}(X_t=i) \\ &= \sum_{i=1}^{n} P_{i,j}\pi_i^{(t)} = \left[\underline{\pi}^{(t)} P\right]_j = \left[\underline{\pi}^{(t-1)} P^2\right]_j = \cdots = \left[\underline{\pi}^{(1)} P^t\right]_j.\end{aligned}$$

# Simulating a Markov Chain

Let $X \in [n]$ be a rv with pmf $\mathbb{P}(X = i) = p_i$ and cdf $F_X(j) = \sum_{i=1}^{j} p_i$

- Goal: Generate a sample of $X$ using a uniform rv $U \in [0, 1)$



- Idea: Find unique $p_j$ interval that contains $U$ and then return $j$
- Then, $X$ equals the unique $j$ satisfying $U \in [F_X(j-1), F_X(j))$
- Using this, one can simulate a Markov chain. If $X_t = x$, then $X_{t+1} = x'$ is generated according to $\mathbb{P}(X_{t+1} = x' | X_t = x)$ using the above method

# Hitting Time

## Definition

For a Markov chain starting from state $i$, the first hitting time of state $j$ is a random variable $T_{i,j}$ with distribution

$$\mathbb{P}\left(T_{i,j} = m\right) = \mathbb{P}\left(X_{m+1} = j, X_m \neq j, X_{m-1} \neq j, \ldots, X_2 \neq j \,|\, X_1 = i\right).$$

By convention, $\mathbb{P}\big(T_{i,j} = 0\big) = \delta_{i,j}$ where $\delta_{i,j}$ is Kronecker delta function and, if state $j$ is not reachable from state $i$, then $T_{i,j} = \infty$. Alternatively, we have

$$\mathbb{P}\left(T_{i,j} \leq m\right) = \mathbb{P}\left(\exists t \in \{1, 2, \ldots, m+1\}, X_t = j \,|\, X_1 = i\right).$$

## Definition

If $X_t$ can get stuck in state $j$ (i.e., $P_{j,j} = 1$), then state $j$ is called absorbing.

## Lemma

If $j$ is an absorbing state, then getting to state $j$ means staying in state $j$ and

$$\mathbb{P}\left(T_{i,j} \leq m\right) = \mathbb{P}\left(X_{m+1} = j \,|\, X_1 = i\right) = \left[P^m\right]_{i,j}.$$

# Stationary Probabilities of Recurrent Markov Chains

## Definition

A state distribution $\underline{\pi}$ is called stationary if it satisfies $\underline{\pi}P = \underline{\pi}$.

## Definition

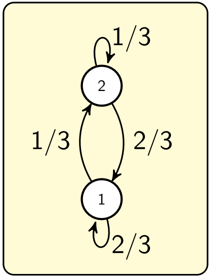
State $j$ is reachable from state $i$ if $[P^m]_{i,j} > 0$ for some $m \in \mathbb{N}$. Two states reachable from each other are communicating. A state is recurrent if it is expected to return to itself infinitely many times.

## Theorem (Perron)

If $P_{i,j} > 0$ for all $i, j \in [n]$, then every state is recurrent and the Markov chain has a unique stationary distribution. This also holds if all entries of $P^m$ are strictly positive for some $m \geq 1$. The stationary distribution $\underline{\pi}$ equals the expected fraction of time that the process spends in each state.

Can find stationary distribution using row reduction after rewriting $\underline{\pi}P = \underline{\pi}$ as

$$(I - P)^T \underline{\pi}^T = \underline{0}.$$



# Next Steps

- To continue studying after this video –
    - Try the required reading: Markov Chains Handout