ECE586 MDPs and Reinforcement Learning **University of Illinois at Urbana-Champaign**
Spring 2019 **Acknowledgment: R. Srikant's Notes and Discussions**

Lecture 2: Markov Chains-Part II, Steepest and Gradient Descent

Instructor: Dimitrios Katselis *Scribe: Andrew Chen and Zih-Siou Hung*

2.1 Periodicity, Limiting Distributions and Reversible Markov Chains

Definition 2.1. *We say that it is possible to return to state i in $n \geq 1$ steps if*

$$\mathbb{P}\left(X_n = i | X_0 = i\right) > 0. \tag{2.1}$$

Definition 2.2. ***(Period)*** *State i is said to have period k if it is only possible to return to state i in $n \geq 1$ steps, where n is a multiple of k. Formally, the period k of state i is defined as:*

$$k = \gcd\{n \geq 1 : P(X_n = i | X_0 = i) > 0\},$$

provided that this set is $\neq \emptyset$. Here, "gcd" denotes the greatest common divisor.

Note: Even though a state i has period k, it may not be possible to reach the state in k steps. For instance, suppose it is possible to return to state i in $6, 9, 12, \ldots$ steps. Then, the period is $k = 3$ although 3 does not appear in this list.

Example. In Figure 2.1 we have a Markov chain with two states. Both states $\{1\}$ and $\{2\}$ have period 2. Since the probability of transitioning from $\{1\}$ to $\{2\}$ is 1 and from $\{2\}$ to $\{1\}$ is also 1, we know that it takes exactly 2 steps to return to either of these states.

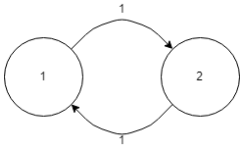

Figure 2.1: Markov chain with two states where each state has period 2.

Definition 2.3. ***(Aperiodic)*** *A state i is aperiodic if its period is 1. A chain is aperiodic if all states are aperiodic.*

Note: An irreducible Markov chain only needs one aperiodic state to imply that all states are aperiodic. This is a consequence of the following theorem:

Theorem 2.4. ***(Period is a Class Property)*** *If states i and j communicate, i.e., i is reachable from j and j is reachable from i, then they have the same period*

Example. In Figure 2.2, states $\{1\}$ and $\{2\}$ have period 1. We can see that the period of $\{2\}$ is 1 because we can return to $\{2\}$ in $\{2, 3, 4, 5, \ldots\}$ steps. Thus, the period of $\{2\}$ is $\gcd\{2, 3, 4, 5, \ldots\} = 1$. Therefore, the Markov chain is aperodic, since it is irreducible. We can verify that, indeed, state $\{1\}$ is aperiodic: in this case $k = \gcd\{1, 2, 3, \ldots\} = 1$.

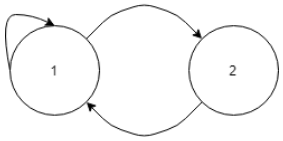

Figure 2.2: A two state irreducible and aperiodic Markov chain.

Example. In Figure 2.3, the Markov chain on the left is irreducible. The period of both states $\{1\}$ and $\{2\}$ is 2. If we add a self-loop to state $\{1\}$, which is shown on the top right chain, then state $\{1\}$ becomes aperiodic, which makes state $\{2\}$ also aperiodic. Similarly, if we add a self-loop to state $\{2\}$, the Markov chain becomes aperiodic. If we add self-loops to both states $\{1\}$ and $\{2\}$, then again the chain becomes aperiodic.

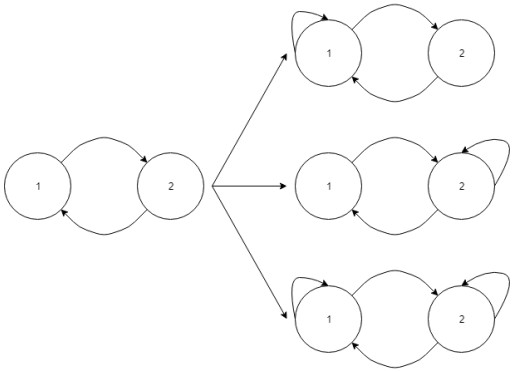

Figure 2.3: An example showing possible ways to turn an irreducible and periodic Markov chain into an aperiodic chain.

Combining aperiodicity and irreducibility, we obtain the following fundamental result for finite state Markov chains:

Theorem 2.5. *For an irreducible and aperiodic finite state Markov chain with unique stationary distribution π:*

$$\lim_{k\to\infty} p(k) = \pi, \quad \forall p(0).$$

Moreover, the following theorem provides a useful decomposition for irreducible Markov chains:

Theorem 2.6. ***(Cyclic structure)*** *For any irreducible Markov chain, one can find a unique partition of the state space $\mathcal{X}$ into d classes $C_0, C_1, \ldots, C_{d-1}$ such that for all k and all $i \in C_k$,*

$$\sum_{j \in C_{k+1}} P_{ij} = 1$$

where $C_d = C_0$ and d is maximal, i.e. there is no other such partition $C'_0, C'_1, \ldots, C'_{d'-1}$ with $d' > d$.

Here, $d \geq 1$ is the period of the chain. Also, $C_0, C_1, \ldots, C_{d-1}$ are called the cyclic classes.

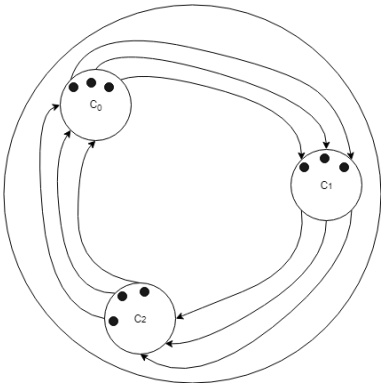

Figure 2.4: This diagram shows an example of cyclic structure for an irreducible Markov chain whose state space can be maximally decomposed into $d = 3$ cyclic classes C_0, C_1, C_2. The period of the chain is clearly 3.

Finally, we will introduce the concept of a reversible Markov chain.

Definition 2.7. ***(Reversible)*** *A Markov chain ($\mathcal{X} = \{1, \ldots, X\}$) with transition matrix P is said to be reversible if there exists a probability distribution π on $\mathcal{X}$ such that*

$$\pi_i P_{ij} = \pi_j P_{ji}, \quad \forall i, j \in \mathcal{X}. \tag{2.2}$$

Such a distribution π on $\mathcal{X}$ is also said to be reversible for the chain (or for P). (2.2) are called detailed balance equations.

Note: Assume that $X_0 \sim \pi$. The left-hand side of (2.2) can be thought of as the amount of probability mass flowing from state i to state j at time 1. Similarly, the right-hand side is the probability mass flowing from state j to state i at the same time instant. This corresponds to a strong form of equilibrium.

Consider a Markov chain with transition matrix P. If there exists a probability vector π satisfying the detailed balance equations (2.2), then π is a *stationary* distribution for this chain and the chain is reversible by definition. In addition, the name *reversible* comes from the fact that if $X_0 \sim \pi$, i.e., the Markov chain is stationary, then for any $n \geq 1$:

$$(X_0, X_1, \ldots, X_n) =_d (X_n, X_{n-1}, \ldots, X_0).$$

Here, $=_d$ denotes identical distribution. Therefore, forward and backward passes of the chain are statistically indistinguishable.

Reversed Chain: To elaborate a bit more on reversibility, consider a homogeneous Markov chain X with transition matrix P and suppose that the chain admits a stationary distribution π such that $\pi(i) > 0, \forall i \in \mathcal{X}$. Define a matrix $Q = [Q_{ij}]$ whose entries satisfy:

$$\pi_i Q_{ij} = \pi_j P_{ji}.$$

It can be easily seen that Q is stochastic:

$$\sum_{j\in\mathcal{X}} Q_{ij} = \sum_{j\in\mathcal{X}} \frac{\pi_j P_{ji}}{\pi_i} = \frac{\pi_i}{\pi_i} = 1, \quad \forall i \in \mathcal{X}.$$

Suppose now that X is initialized at π. Then, by Bayes' rule:

$$P(X_n = j|X_{n+1} = i) = \frac{P(X_{n+1} = i|X_n = j)P(X_n = j)}{P(X_{n+1} = i)} = Q_{ij}.$$

Thus, Q is the transition matrix of the initial chain when time is reversed, i.e., of the reversed chain. For a reversible Markov chain:

$$Q_{ij} = P_{ij} \text{ or } P(X_n = j|X_{n+1} = i) = P(X_{n+1} = j|X_n = i).$$

Therefore, the chain and the time-reversed chain are statistically indistinguishable.

2.2 Markov Chain Examples

We now look at some examples of Markov chains.

Example. Given a two state Markov chain with transition matrix $P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$, where $\alpha, \beta \in (0, 1)$, our goal is to find the stationary distribution π. We know that the stationary distribution π satisfies the global balance equation $\pi = \pi P$. Thus, we have:

$$\begin{bmatrix} \pi(1) & \pi(2) \end{bmatrix} = \begin{bmatrix} \pi(1) & \pi(2) \end{bmatrix} \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$

from which we get

$$\begin{bmatrix} \pi(1) & \pi(2) \end{bmatrix} = \begin{bmatrix} \pi(1)(1-\alpha) + \pi(2)\beta & \pi(1)\alpha + \pi(2)(1-\beta) \end{bmatrix}$$

Combining the two equations, we obtain:

$$\pi(1)\alpha = \pi(2)\beta.$$

Employing the constraint

$$\pi(1) + \pi(2) = 1,$$

we get

$$\pi = \begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{bmatrix}.$$

Example. (**Random walks on graphs**) A graph $G = (V, E)$ consists of a vertex set $V = \{v_1, v_2, \cdots v_k\}$ and an edge set $E = \{e_1, e_2, \cdots e_m\}$. Two vertices are neighbors if there is an edge connecting them. A random walk on a graph $G = (V, E)$ is a Markov chain with state space V and transition probabilities P_{ij} defined as follows:

$$P_{ij} = \begin{cases} \frac{1}{d_i}, & \text{if } v_i \text{ and } v_j \text{ are neighbors} \\ 0, & \text{otherwise} \end{cases},$$

where vertex v_i is the current state and d_i is the degree of v_i.

It turns out that random walks on graphs are reversible Markov chains with reversible or stationary distribution

$$\pi = \left(\frac{d_1}{\sum_{i=1}^k d_i}, \frac{d_2}{\sum_{i=1}^k d_i}, \cdots, \frac{d_k}{\sum_{i=1}^k d_i} \right).$$

We can see that the detailed balance equations hold for this choice of π:

$$\pi_i P_{ij} = \begin{cases} \frac{d_i}{\sum_{l=1}^k d_l} \cdot \frac{1}{d_i} = \frac{d_j}{\sum_{l=1}^k d_l} \cdot \frac{1}{d_j} = \pi_j P_{ji}, & \text{if } v_i \text{ and } v_j \text{ are neighbors} \\ 0 = \pi_j P_{ji}, & \text{otherwise} \end{cases}.$$

Example. (**Birth-and-Death Process**) Let X be a Markov chain with state space $\mathcal{X} = \{1, 2, \ldots, X\}$ and transition matrix P satisfying the following properties:

1. $P_{ij} > 0$, if $|i - j| = 1$ (note: some or all the self-loops P_{ii} are possibly absent)
2. $P_{ij} = 0$, if $|i - j| \geq 2$.

Such a Markov chain is often called a *birth-and-death process*. Any Markov chain of this kind is reversible. To construct a reversible distribution π for this chain, we start by setting $\tilde{\pi}_1 = a$ for some arbitrary $a > 0$. Imposing the detailed balance equation for $i = 1, j = 2$, we obtain:

$$\tilde{\pi}_2 = \frac{aP_{12}}{P_{21}}.$$

Applying the balance equation for $i = 2, j = 3$:

$$\tilde{\pi}_3 = \frac{\tilde{\pi}_2 P_{23}}{P_{32}} = \frac{aP_{12}P_{23}}{P_{21}P_{32}}.$$

Continuing in this way, we obtain:

$$\tilde{\pi}_i = \frac{a \prod_{m=1}^{i-1} P_{m,m+1}}{\prod_{m=1}^{i-1} P_{m+1,m}}$$

Therefore, π can be chosen to be:

$$\pi = \left(\frac{\tilde{\pi}_1}{\sum_{i=1}^X \tilde{\pi}_i}, \frac{\tilde{\pi}_2}{\sum_{i=1}^X \tilde{\pi}_i}, \cdots, \frac{\tilde{\pi}_X}{\sum_{i=1}^X \tilde{\pi}_i} \right).$$

It can be easily checked that this is a reversible distribution.

2.3 Steepest and Gradient Descent

In this section, we start introducing the steepest and gradient descent algorithms. Steepest descent is a special case of gradient descent. The purpose of these methods is to numerically solve the minimization problem:

$$\min_{x \in \mathbb{R}^n} f(x). \tag{2.3}$$

Steepest descent corresponds to the iterative scheme:

$$x_{k+1} = x_k - \varepsilon_k \nabla f(x_k),$$

where $\varepsilon_k \geq 0$ is the stepsize controlling how large the step should be at the kth iteration. The intuition behind this algorithm is the fact that the gradient $\nabla f(x_k)$ is the direction towards which the function f increases the most. Thus, if we move towards the opposite direction, i.e. towards $-\nabla f(x_k)$ by a reasonable amount, the value of the function should decrease. Note that the sequence of the stepsizes $\{\varepsilon_k\}$ affects the convergence of the algorithm as seen in Figure 2.5. If the stepsizes are too small, the convergence rate to a local minimum would be small. In contrast, if the stepsizes are too large, steepest descent may diverge since it may overshoot local minima (right dotted line in Figure 2.5). Thus, the stepsize sequence $\{\varepsilon_k\}$ should be judiciously chosen to balance between the goal of convergence to a local minimum and the goal of a reasonable convergence rate.

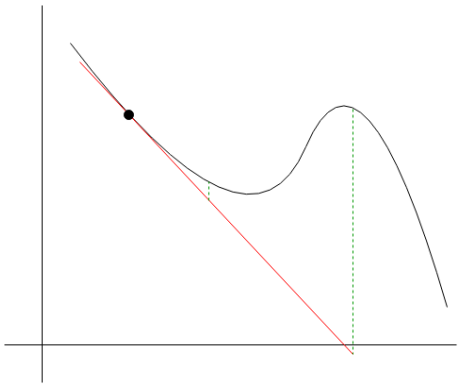

Figure 2.5: An example of the effect of $\{\varepsilon_k\}$ on the descent algorithm.