Mixing time bounds for Arithmetic Chain on the n cycle

Andrew Heeszel

Introduction

Let M be a markov chain on the state space of $\mathbb{Z}_p^n$ omitting the zero vector. Let $G = (V, E)$ be the n-cycle. Each coordinate of all elements of our sample space will uniquely correspond to a vertex of G. Given an element of the sample space, the transition rules are as follows:

1. Uniformly select an edge of the n cycle.
2. Uniformly select a single vertex on the edge. With probability $\frac{1}{2}$, the selcected vertex will take the value of the sum of its current value and the value of its selected neighbor mod p. Otherwise the coordinate will take the value of the difference between the current value of the selected vertex and the value of the selected neighboring vertex mod p.

These rules imply that the markov chain is aperiodic, irriducible, reversible and converges to uniform distribution.

Lower mixing time bound

Proposition: *For the arithmetic chain on the n cycle it follows $t_{mix} \geq \frac{n^2}{6}$.*
Proof:

Let X_t be the arithmetic chain on the n cycle with a sample space Ω. Define $A := \{y \in \Omega \mid y_{\frac{n}{2}} \neq 0\}$, and let $x \in \Omega$ be the element where $x_1 = 1$, and $x_i = 0$ for all $i \neq 1$. Let $X_0 = x$ and define the stopping time
$\tau_A := \min\{t \geq 1 \mid X_t \in A\}$. Notice that $X_t \in A$ only if $\tau_A \leq t$, giving us the inequality $p^t(x, A) \leq P(\tau_A \leq t)$ for all t. Let Y_t denote the minimal distance between any nonzero vertex of X_t and vertex $\frac{n}{2}$. Notice that if $Y_t \neq 0$, then $Y_{t+1} = Y_t - 1$ only if an edge between a nonzero vertex of minimal distance to $\frac{n}{2}$ of X_t and its neighbor nearest to $\frac{n}{2}$ is selected, and the nearest nonzero vertex of X_t either adds or subtracts its value onto its neighboring vertex. On the n cycle there are no more than two nonzero vertices of minimal distance to $\frac{n}{2}$ for any $X_t \notin A$, and therefore $P(Y_{t+1} < Y_t \mid Y_t \neq 0) \leq \frac{2n}{2} = \frac{1}{n}$. Since at any time t, $X_t \in A$ if and only if $Y_t = 0$, we can equivalently define our stopping time $\tau_A := \min\{t \geq 1 \mid Y_t = 0\}$. Since for any $t < \tau_A$, Y_t only decreases in

value with a probability at most $\frac{1}{n}$, if we define the random variable B_t to have binomial distribution with parameters t and $\frac{1}{n}$, then $P(\tau_A \leq t) \leq P(B_t \geq \frac{n}{2})$, since at any time there is at most a $\frac{1}{n}$ probability of decreasing the value of Y_t by 1, and Y_t must decrease in value $\frac{n}{2}$ times. Therefore, since $\mathrm{E}[B_t] = \frac{t}{n}$ and $\mathrm{Var}(B_t) = \frac{t}{n}(1 - \frac{1}{n})$, we have by Chebyshev's inequality,
$p^{\frac{n^2}{6}}(x, A) \leq P(\tau_A \leq \frac{n^2}{6}) \leq P(B_{\frac{n^2}{6}} \geq \frac{n}{2}) \leq P(|B_{\frac{n^2}{6}} - \frac{n}{6}| \geq 2\frac{n}{6}) \leq$
$P(|B_{\frac{n^2}{6}} - \frac{n}{6}| \geq 2\sqrt{\frac{n}{6}(1 - \frac{1}{n})}) \leq \frac{1}{4}$. Since A is the set of all elements of our sample space with vertex $\frac{n}{2}$ fixed to be nonzero, we have that $\pi(A) = \dfrac{p^{(n-1)}(p-1)}{p^n - 1} \geq \frac{1}{2}$.
Therefore, $d(\frac{n^2}{6}) \geq |\pi(A) - p^{\frac{n^2}{6}}(x, A)| \geq \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$.

Upper mixing time bound

We will use a method of dirichlet form comparison to bound the spectral gap of the transition matrix P of the arithmetic chain on the n cycle modulo p. For two reversible transition matricies P and $\tilde{P}$, let Γ_{xy} be an E path from x to y using the edges from the transition matrix P, and define
$B := \max_{e \in E} \dfrac{1}{Q(e)} \Sigma_{x,y:\Gamma_{xy} \ni e} \tilde{Q}(x, y) \mid \Gamma_{xy} \mid$.
Given this definition we have the following thorem.

Theorem 1: Let P and $\tilde{P}$ be reversible transition matricies, and let f be any real valued function on Ω. Then for the dirichlet forms $\mathcal{E}(f)$ and $\tilde{\mathcal{E}}(f)$ and congestion ratio B as defined, $\tilde{\mathcal{E}}(f) \leq B\mathcal{E}(f)$.

This theorem gives the following corollary when comparing the dirichlet form of P to the transition matrix Π on the same state space consisting of row vectors of π.

Corollary 1: Let P be a reversible transition matrix on the state state Ω with a stationary distribution π. Suppose Γ_{xy} is a choice of an E path from x to y, and let
$B = max_{e \in E} \dfrac{1}{Q(e)} \Sigma_{x,y:\Gamma_{xy} \ni e} \pi(x)\pi(y) \mid \Gamma_{xy} \mid$. Then the spectral gap γ of P satisfies $\gamma \geq B^{-1}$.

To directly bound the mixing time of the Markov chain, we will use the following theorem.

Theorem 2: For an irriducible, aperiodic Markov chain with a reversible transition matrix P, stationary distribution π with minimal value π_{min}, and absolute spectral gap γ, let $t_{rel} = \frac{1}{\gamma}$. Then, $t_{mix}(\epsilon) \leq t_{rel} \log(\frac{1}{\epsilon \pi_{min}})$.

We will begin by defining a set of paths from any two states $x, y \in \Omega$, and bounding the number of paths used in a fixed edge $e \in E$.

Proposition: There exists a set of canonical paths from all $x, y \in \Omega$ with a maximum length of np an upper bound of αnp^{n+1} paths using a given edge $e \in E$ for some constant α.

Proof:

We will begin by describing a set of canonical paths from any two elements of our sample space. Fix $x, y \in \Omega$. Order the vertices of the n cycle such that for any vertex $i \neq n$, vertex $(i+1)$ is the counterclockwise neighbor to i. For convention let vertex $n+1$ be vertex 1, and vertex 0 be vertex n.

1. Begin at the first nonzero vertex of x. Repeat the following until and including the furthest counterclockwise nonzero vertex of y. Suppose we are at the kth vertex for some integer $1 \leq k \leq n$, note that by our path rules the current vertex k will always be nonzero.

   (a) If vertex $(k+1)$ is the counter-clockwise furthest nonzero vertex of y from the first nonzero vertex of x, add vertex k onto vertex $(k+1)$ until vertex $(k+1)$ agrees with y. Then add the nonzero vertex $(k+1)$ onto vertex k until vertex k agrees with y.

   (b) If vertex $(k+1)$ is nonzero for x, add vertex $(k+1)$ onto vertex k until vertex k agrees with y.

   (c) If vertex $(k+1)$ is zero for x, add vertex k onto vertex $(k+1)$. After add vertex (k+1) onto vertex k until vertex k agrees with y.

2. At this point all vertices in the interval between first nonzero vertex of x and counter-clockwise furthest nonzero vertex of y from the first nonzero vertex of x will agree with y, while all other vertices will agree with x. If the last nonzero vertex of x is not counter-clockwise further from the furthest vertex of of y from the lowest nonzero vertex of x, then the the path from x to y is completed. Otherwise begin at the the last nonzero vertex of x.

   (a) Beginning at some vertex l, if vertex $(l-1)$ is the counter-clockwise furthest nonzero vertex of y from the first nonzero vertex of x, add vertex $(l-1)$ onto vertex l until vertex l is 0. At this point the path from x to y is completeted.

   (b) At vertex l if vertex $(l-1)$ is nonzero, add vertex $(l-1)$ onto vertex l until vertex l is 0.

   (c) At vertex l if vertex $(l-1)$ is 0, add vertex l onto vertex $(l-1)$, and then add vertex $(l-1)$ onto vertex l until vertex l is 0.

Notice by our path rules fixing the value of the counter-clockwise furthest nonzero vertex of y from the first nonzero vertex of x and its immediate clockwise neighbor takes no more than $2p$ steps, while all other verticies take no more than p steps to change their values to match y, we can conclude for any E path Γ_{xy} from x to y that $\mid \Gamma_{xy} \mid \leq (n-2)p + 2p = np$.

We will now fix a given edge $e \in E$ and bound the number of paths Γ_{xy} from a state x to another state y that use the fixed edge. We will define our edge $e = \{z, w\}$, where $z = (z_1, z_2, ...z_n)$ and $w = (w_1, w_2, ..., w_n)$.

We will now fix two neighboring states z, $w \in \Omega$ and bound the number of paths using the edge connecting z and w. By our path rules, we can assume that all paths follow the edge in direction $z \to w$ where z and w are identical copies aside from w having a vertex i that is the sum of the ith vertex of z and a neighboring vertex of z modulo p. For notation we will order the vertices for each element of our sample space $z = (z_1, ..., z_n),\ w = (w_1, ..., w_n)$. Additionally all paths must either follow the path rules described in 1) or 2), and w_i must either be the sum of z_i and its clockwise or counterclockwise neighbor modulo p.

For paths following the rules in 1), suppose that w_i is the sum of z_{i-1} and z_i modulo p. Let k be the first nonzero vertex of x. Let x, $y \in \Omega$ be any states such that the path from x to y includes the edge $z \to w$. Our path rules imply that vertices $z_k, z_{k+1}, ..., z_{i-2}$ must agree with y while other other vertices of z aside from z_{i-1} and z_i must agree with x. Since the edge involves the addition from a counter-clockwise neighbor, this occurs only if $x_i = 0$. Therefore, of the $2n$ ambiguous vertices of x and y there exists at least $n-1$ fixed. This implies that conditioned on the location of k the number of paths using the rules in 1) from a clockwise neigbor is $O(p^{n+1})$. Therefore since there are n locations possible for k, the total number of paths must be $O(np^{n+1})$.

To find the numer of used in 1) from a counterclockwise neighbor, as before let $x, y \in \Omega$ be states of our sample space such that the path from x to y involves the edge $z \to w$. Let k be the first nonzero vertex of k. Since we are adding from a counterclockwise neighbor, it follows that $w_i = z_i + z_{i+1}$ modulo p. Our path rules then imply that $z_k, z_{k+1}, ..., z_{i-1}$ must agree with y, while all other vertices of z aside from z_i and z_{i+1} must agree with x. Suppose x_{i+1} is nonzero. This implies $x_{i+1} = z_{i+1}$, and therefore of there are $n-1$ vertices of x and y fixed of the $2n$ total. Therefore there exists no more than $\frac{p^{2n}}{p^{n-1}} = p^{n+1}$ paths using the edge $z \to w$. If $x_{i+1} = 0$, then additionally of the $2n$ total vertices of x and y, $n-1$ must be fixed. Therefore there exists no more than p^{n+1} paths using the edge $z \to w$. Thus, conditioned on the location of k, the number of paths using the edge $z \to w$ is $O(p^{n+1})$. Since there are n possible locations for vertex k, the number of paths in 1) when adding from a counter-clockwise neighbor must be $O(np^{n+1})$.

We will now bound the number of paths using the edge $z \to w$ when following the rules in 2). Suppose the path from state $x \in \Omega$ to $y \in \Omega$ uses the edge $z \to w$ by setting $w_i = z_i + z_{i+1} \mod p$. Let k be the last nonzero vertex of y. Our path rules then imply that all vertices of z along the path $i+2, i+3, ..., k$ must agree with y Additionally, all vertices of z along the path $k+1, k+2, ..., i-1$ must agree with x. If x_{i+1} is nonzero, then our movement rules imply vertex $x_{i+1} = z_{i+1}$, showing at least $n-1$ of the $2n$ vertices of x and y must be fixed, which additionally is the case if $x_{i+1} = 0$. Therefore there are no more than $2p^{n+1}$ paths using the rules in 2) conditioning on $w_i = z_i + z_{i+1} \mod p$ and fixing the location of k. Since there are no more than n locations k can take, it follows that the total number of paths following the rules in 2) where the edge $z \to w$ involves adding from a counterclockwise neighbor must be $O(np^{n+1})$.

Suppose the path from state $x \in \Omega$ to state $y \in \Omega$ uses the edge $z \to w$ using the rules in 2) and $w_i = z_{i-1} + z_i \mod p$. Let k be the last nonzero vertex of y. Our path rules imply that vertices of z along the path of the n cycle, $k+1, k+2, ..., i-2$ must agree with x while vertices of z along the path $i+1, i+2, ..., k$ must agree with y. If vertex $i-1$ is the last nonzero vertex of y, then $y_k = z_k$ and $n-1$ vertices must be fixed. Otherwise, if x_{i-1} is nonzero, then vertex $x_{i-1} = z_{i-1}$, or $x_{i-1} = 0$. This implies of the $2n$ vertices at least $n-1$ are fixed, implying the number of paths using the given edge conditioned on the location of vertex k must be $O(p^{n+1})$. Since vertex k has a maximum of n possible locations, it follows the total number of paths using the rules in 2) adding from a clockwise neighbor must be $O(np^{n+1})$, implying that the to the total number of paths from any two vertices $x, y \in \Omega$ using a fixed edge must be $O(np^{n+1})$.

Proposition: *If* $B := max\ _{e \in E} \dfrac{1}{Q(e)} \Sigma_{x,y:\Gamma_{x,y} \ni e} \pi(x)\pi(y)\Gamma_{x,y}$ *denotes the congestion ratio when comparing the transition matrix* P *with stationary distribution* π *of our studied chain to the chain on* Ω *with transition matrix* Π *consisting of row vectors of* π*, then it follows* $B \leq \beta n^3 p^2$*.* for some constant β.

Proof:

Since $|\Omega| = p^n - 1$, and our chain converges to uniform distribution, it follows $\pi(x) = \frac{1}{p^n-1}$ for all $x \in \Omega$. We have also shown there exists a set of canonical paths with maximum length np and having no more than αnp^{n+1} paths used per edge, for some constant α. And by our transition rules, for any two neighboring states $x, y \in \Omega$, $P(x,y) \geq \frac{1}{4n}$. These inequalities imply,

$$B = \max\ _{e \in E} \frac{1}{Q(e)} \Sigma_{x,y:\Gamma_{x,y} \ni e} \pi(x)\pi(y)\Gamma_{x,y} \leq \max\ _{e \in E} \frac{1}{\frac{1}{4n} \cdot \frac{1}{p^n-1}} \Sigma_{x,y} \frac{1}{p^n-1} \cdot \frac{1}{p^n-1} \cdot np$$

$$= \max\ _{e \in E} \frac{\alpha n^2 p}{p^n - 1} \Sigma_{x,y} 1 = \frac{\alpha n^2 p \cdot np^{n+1}}{p^n - 1} \leq \beta n^3 p^2.$$

Proposition: *For Arithmetic chain on the n-cycle it follows,*
$t_{mix} \leq \gamma n^4 p^2 \log(\frac{p}{\epsilon})$*, for some constant* γ independent of n and p.

Proof:

Since we have shown $B \leq \beta n^3 p^2$, for some constant β independent of n and p. Since the chain converges to uniform distribution, $\pi_{min} = \frac{1}{p^n - 1}$. By corollary 13.24, $B \geq t_{rel}$ when B is the congestion ratio of P and Π. Therefore by theorem 12.3,

$t_{mix}(\epsilon) \leq \beta n^3 p^2 \log(\frac{1}{\epsilon} \cdot \dfrac{1}{\frac{1}{p^n - 1}}) \leq \beta n^3 p^2 \log(\frac{1}{\epsilon} \cdot (p^n - 1)) \leq \beta n^3 p^2 \log((\frac{p}{\epsilon})^n) \leq \gamma n^4 p^2 \log(\frac{p}{\epsilon})$.