18.821 PROJECT 2: ASYMMETRIC PROCESSES

MATTHEW BEVERIDGE, BARIŞ EKIM, JUSTIN LIM

Abstract. We examine asymmetric processes on two different types of graph: the closed loop and the open-ended line. We show that these processes can be likened to discrete Markov chains, and evaluate them as such. For the closed graph, we derive the stationary distribution and the mixing time. For the open graph, we derive the stationary distribution for a length 3 line, provide conditions for existence of a unique stationary distribution for arbitrarily long lines, and determine the mixing time. Finally, we conjecture a relation between the entry/exit parameters of the open graph and clusters in the stationary distribution.

1. Introduction

A *stochastic process* is a set of random variables $\{X_t, t \in T\}$ defined on a set D where T is a set of indices representing time, X_t the *state* of the process at time t, and D the set of states. A stochastic process could be discrete or continuous, depending on T being discrete or continuous; i.e, a stochastic process represents a random phenomenon evolving with time (number of events occurring, or measurements recorded over time).

A *Markov chain* is a stochastic process that occupies a state in a countable or finite state space S at any given point in time. Let the Markov chain currently be at state i, and moves to state j with probability $0 \leq P_{ij} \leq 1$, and $\sum P_{ij} = 1$. We call the matrix $P = \{P_{ij}\}_{ij}$ the *transition matrix.*

Clearly the Markov chain is a random process in time, hence a stochastic process, but the main property of the Markov chain is that it's *memoryless*; that is, the next state only depends of the current state. Formally,

$$\Pr[X_{t+1} = j | X_0 = i_0, X_1 = i_1 ... X_i = i] = \Pr[X_{t+1} = j | X_t = i]$$

Let $q^t = [q_1^t, q_2^t, ...q_n^t]$ be a row vector such that q_i^t is the probability that the Markov chain is at state i at time t. Then, q^t is called the state probability vector, or the distribution of the Markov chain at time t. The main property is

$$q^{t+1} = q^t \cdot P = q^0 \cdot P^t$$

where P is the aforementioned transition matrix.

Let π be such a distribution. If for such π,

$$\pi^t = \pi^t \cdot P$$

we say that π is the *stationary distribution*, or equivalently the *equilibrium distribution.* Clearly for such π we have

$$\pi^{t+1} = \pi^t \cdot P$$

Key words and phrases. asymmetric processes.

When a chain arrives at a stationary distribution π, it stays at that distribution, (i.e. the probability of being at any vertex tends to a limit independent of the initial vertex).

In this paper, we examine the stationary distribution π for asymmetric stochastic processes.

2. Closed case

In the closed case, the k particles are on a closed loop with N positions. Each position can be occupied by at most one particle, so there are $\binom{N}{k}$ states of the system. Given an *asymmetry parameter* q, we perform the following procedure indefinitely:

(1) Choose one of the k particles uniformly at random.
(2) Flip a biased coin with probability of heads equal to q.
(3) If heads, move the chosen particle one position clockwise (if that position is vacant). If tails, move the chosen particle one position counterclockwise (if that position is vacant).

2.1. **Stationary distributions on the closed loop.** Firstly, we consider the case where $k = 1$, so there is one particle and hence N states.

Theorem 2.1. *For a closed loop with N positions and $k = 1$ particle, the stationary distribution is the uniform distribution on the N states, i.e. each state is equally likely with probability $\frac{1}{N}$.*

Proof. Firstly, $\pi = (\pi_1, \ldots, \pi_N)$ is a stationary distribution if and only if $\sum_i \pi_i = 1$ and all the following system of equations hold:

$$\begin{aligned} (1-q)\pi_N + q\pi_2 &= \pi_1 \\ (1-q)\pi_1 + q\pi_3 &= \pi_2 \\ &\cdots \\ (1-q)\pi_{N-1} + q\pi_1 &= \pi_N \end{aligned}$$

Call this system $(*)$. Note that the uniform distribution $\pi^* = (\frac{1}{N}, \ldots, \frac{1}{N})$ indeed satisfies this system. If $q \in \{0, 1\}$, it is clear that π^* is the only solution. Hence assume $q \neq 0, 1$. We claim that this is still the only solution. Squaring every equation in $(*)$ yields the system

$$\begin{aligned} (1-q)^2\pi_N^2 + q^2\pi_2^2 + 2q(1-q)\pi_2\pi_N &= \pi_1^2 \\ (1-q)^2\pi_1^2 + q^2\pi_3^2 + 2q(1-q)\pi_3\pi_1 &= \pi_2^2 \\ &\cdots \\ (1-q)^2\pi_{N-1}^2 + q^2\pi_1^2 + 2q(1-q)\pi_1\pi_{N-1} &= \pi_N^2 \end{aligned}$$

Summing all of them, we get

$$(1-q)^2\left(\sum_i \pi_i^2\right) + q^2\left(\sum_i \pi_i^2\right) + 2q(1-q)\left(\sum_i \pi_i\pi_{i-2}\right) = \left(\sum_i \pi_i^2\right)$$

$$\iff 2q(1-q)\left(\sum_i \pi_i^2\right) = 2q(1-q)\left(\sum_i \pi_i\pi_{i-2}\right)$$

$$\iff \sum_i \pi_i^2 = \sum_i \pi_i \pi_{i-2}$$

where indices are taken modulo N. However,

$$\sum_i (\pi_i - \pi_{i-2})^2 = 2\left(\sum_i \pi_i^2 - \sum_i \pi_i \pi_{i-2}\right) = 0$$

It follows that $\pi_i = \pi_{i-2}$ for all i. This implies that $\pi_N = \pi_2$, so that $\pi_1 = (1-q)\pi_N + q\pi_2 = \pi_2$. This means that $\pi_1 = \pi_2 = \cdots = \pi_N$. Since they sum to 1, $\pi_i = \frac{1}{N}$ for all i, and we're done. $\square$

More generally, the statement of Theorem 2.1 holds for any number of particles k. Consider the $\binom{N}{k} \times \binom{N}{k}$ transition matrix A for this asymmetric process, where a_{ij} is the probability of transitioning from state i to state j. Clearly, the rows of A sum to 1. We prove that A is *doubly stochastic*: its columns also sum to 1.

Lemma 2.2. *Let A be the $\binom{N}{k} \times \binom{N}{k}$ transition matrix for the asymmetric process on k particles and N positions on a closed loop. Then the columns of A sum to 1.*

Proof. Consider some state U. Consider any one of the k particles in U, say u. There are two possible states U_1, U_2 that can transition to U by moving u_i:

- U_1: The state prior to an attempt to move u clockwise,
- U_2: The state prior to an attempt to move u counter-clockwise.

Figure 1 illustrates these states. Note that U_1, U_2 do not have to be different from U.

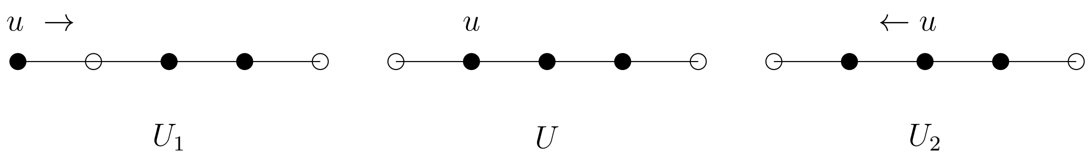


FIGURE 1. Illustration of U, U_1, and U_2. Note that $U \equiv U_2$.

Notice that $Pr(\text{transition to} U|U_1) = q$, $Pr(\text{transition to} U|U_2) = 1-q$. Since we pick u with probability $\frac{1}{k}$ in each case,

$$Pr(\text{transition to } U \text{ by moving } u) = \frac{1}{k}\cdot q + \frac{1}{k}\cdot(1-q) = \frac{1}{k}$$

Summing over all u, i.e. over all k particles, we find that

$$\sum_i a_{iU} = \sum_i Pr(\text{transition to } U|i) = \sum_{u\in U} Pr(\text{transition to } U \text{ by moving } u_i) = 1$$

which proves the Lemma. $\square$

Theorem 2.3. *For a closed loop with N positions and k particles, the stationary distribution is the uniform distribution on the $\binom{N}{k}$ states.*

Proof. As before, let A be the transition matrix for this asymmetric process. A vector π is a stationary distribution iff $\pi_i \geq 0$ for all i, $\sum_i \pi_i = 1$, and $\pi A = \pi$. Let v be the state with the greatest probability. π must satisfy the system:

$$\sum_i a_{iu}\pi_i = \pi_u \text{ for all states } u \tag{2.1}$$

By our Lemma, $\sum_i a_{iv} = 1$. Therefore,

$$\pi_v = \left(\sum_i a_{iv}\right) \pi_v \geq \sum_i a_{iv}\pi_i = \pi_v$$

It follows that each π_i is equal to π_v, i.e. π is uniform. It remains to note that the uniform distribution π clearly satisfies Equation 2.1, since

$$\sum_i a_{iu} \binom{N}{k}^{-1} = \binom{N}{k}^{-1}$$

which concludes our proof. □

2.2. **Mixing time on the closed loop.** Let $S = \binom{N}{k}$ be the number of states. We now know that the stationary distribution of the asymmetric process on a closed loop is the uniform distribution π^*. If $p = (p_1, \ldots, p_S)$ is our starting distribution, then the state distribution after t iterations of the process is pA^t. We're interested in the *mixing time* of the process: how quickly does p coverge to π^*?

Suppose $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_S|$ are the eigenvalues of A, corresponding to eigenvectors $v_1, \ldots, v_S$. Note that we're working with magnitudes because the eigenvalues and eigenvectors can be complex. The following result shows that 1 is the eigenvalue with largest magnitude.

Lemma 2.4. $\lambda_1 = 1$.

Proof. We use a similar argument as in Theorem 2.3. If v is an eigenvector of A with eigenvalue λ, and m is the index of the greatest $|v_m|$, then

$$|v_m| = \left(\sum_i a_{iv}\right)|v_m| \geq \left|\left(\sum_i a_{iv}\right) v_m\right| \geq \left|\sum_i a_{im}v_i\right| = |\lambda v_m| \Rightarrow |\lambda| \leq 1$$

By the same theorem, 1 is an eigenvalue, so $\lambda_1 = 1$, as desired. □

We can write the starting distribution p as a linear combination of the eigenvectors of A, which gives us the following:

$$p = \sum_{i=1}^{S} c_i v_i \Rightarrow pA^t = \sum_{i=1}^{S} c_i v_i A^t = \sum_{i=1}^{S} c_i v_i \lambda_i^t \tag{2.2}$$

Because Lemma 2.4 tells us that $|\lambda_i| \leq 1$, every λ_i^t goes to 0 except for λ_1. Hence p must converge to $v_1 = \pi^*$, and the rate at which this occurs is therefore controlled by the magnitude $|\lambda_2|$: the smaller it is, the quicker pA^t converges to π^*.

Example 2.5. Figures 2 and 3 illustrate the second largest magnitude of the eigenvalues of the transition matrix (i.e. $|\lambda_2|$), for the simple case with $k = 1$ particle. Note the following:

- For fixed q, $|\lambda_2|$ increases as N increases. As the number of positions N increases, it takes longer for the asymmetric process to converge to the uniform distribution. Furthermore, the pattern with which $|\lambda_2|$ increases seems to depend on the parity of N. The blue and orange lines in Figure 2 illustrate this.
- For fixed N, the fastest mixing time occurs when $q = 0.5$, which is when clockwise and counter-clockwise movements are equally likely. When q is biased in one direction, the mixing time is higher.

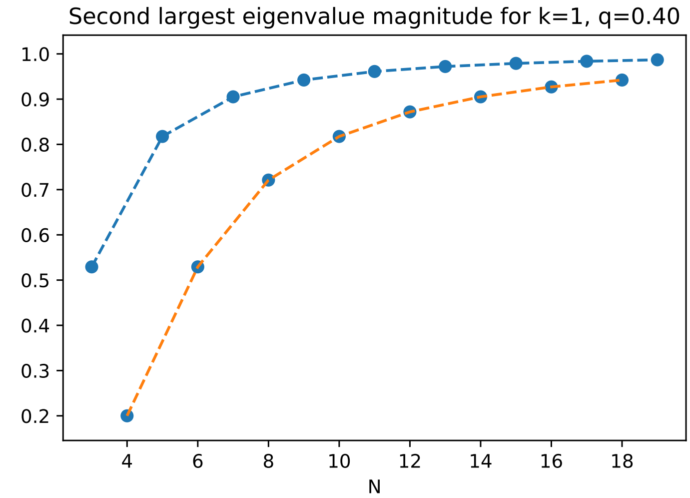


FIGURE 2. Second largest eigenvalue magnitude as N varies, for the case with $k = 1$ particle and asymmetric parameter $q = 0.4$.

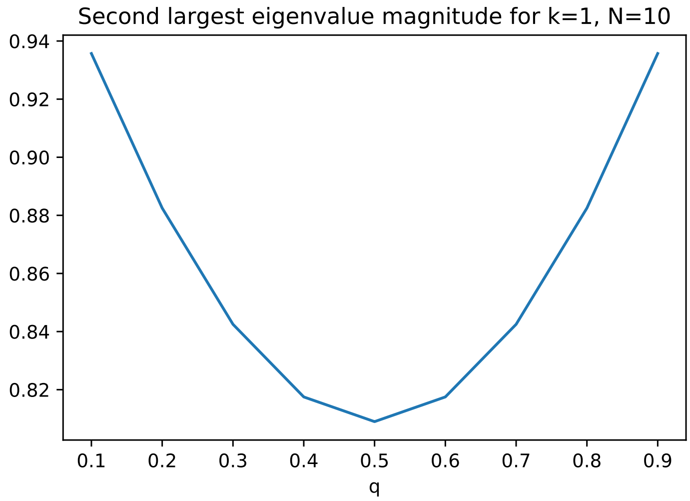


FIGURE 3. Second largest eigenvalue magnitude as q varies, for the case with $k = 1$ particle and $N = 10$

Example 2.6. Here we compute the mixing time for $k = 1$ particles on $N = 3$ positions. The transition matrix is

$$A = \begin{bmatrix} 0 & q & 1-q \\ 1-q & 0 & q \\ q & 1-q & 0 \end{bmatrix}$$

The eigenvalues of A are the roots of the polynomial

$$\lambda^3 - q^3 - (1-q)^3 - 3\lambda q(1-q) = 0 \iff \lambda^3 - 1 = (\lambda - 1) \cdot 3q(1-q)$$
$$\iff \lambda^2 + \lambda + 1 = 3q(1-q) \text{ if } \lambda \neq 1$$

The roots of this quadratic equation are $-\frac{1}{2} \pm \frac{\sqrt{3}}{2}(2q-1)i$, so here $|\lambda_2|^2 = \frac{1}{4}(1+3(2q-1)^2) < 1$. The minimum possible mixing time is attained at $q = \frac{1}{2}$, where $|\lambda_2| = \frac{1}{2}$.

3. Open Case

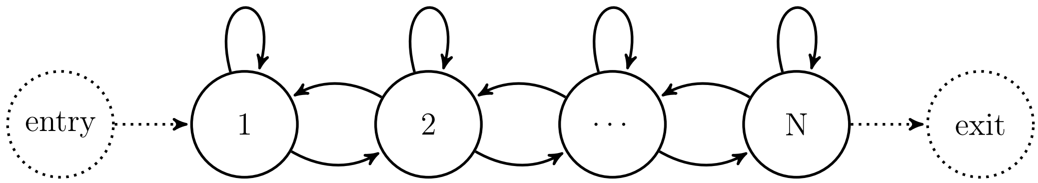


Figure 4. Open Markov chain with N nodes.

In the open case, we consider the chain of N positions on a line as shown in Figure 4. Particles can enter on the left and exit on the right, meaning the number of particles on the line is variable and there are 2^N possible states of the Markov chain. As in the closed case, we choose an *asymmetry parameter* $q \in [0, 1]$ but now additionally choose *entry* and *exit* parameters $\alpha, \beta \in [0, 1]$. The process behaves as follows:

1) Choose with uniform probability $\frac{1}{N+2}$ one of the N positions on the line, *entry*, or *exit*.
   - If one of the N positions on the line is chosen, and there is a particle in that position, move it to the left with probability q (if there is a position there and it is vacant) and to the right with probability $1 - q$ (if there is a position there and it is vacant).
     Note: the *entry* and *exit* positions are not occupiable and only serve as a representation of adding and removing particles.
   - If *entry* is chosen, add a particle to the leftmost position with probability α (if it is vacant).
   - If *exit* is chosen, remove the particle at the rightmost position with probability β (if it is occupied).
2) Repeat.

We examine the stationary distribution and mixing time for this process in relation to q, α, and β.

3.1. Stationary Distribution on Open Loop.

To begin, we will examine the case where $N = 3$ and expand from there. In this case, there are 8 possible states:

State 1:	○○○		
States 2, 3, 4:	●○○	○●○	○○●
States 5, 6, 7:	●●○	●○●	○●●
State 8:	●●●		

Following the rules outlined above we know that State 1 can only go to State 2 or itself, State 2 can go to State 3 or itself, State 3 can go to States 2, 4, 5, or itself, and so on. This is shown explicitly in Figure 5. Subsequently, we derive the transition matrix $\mathbf{P}$ for this outlined Markov chain.

Definition 3.1. The *transition matrix* $\mathbf{P}$ of a Markov chain is a matrix with entries $\mathbf{P}_{i,j}$ representing the probability of going from state i to state j.

$$\mathbf{P} = \begin{array}{cc} & \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} & \begin{Vmatrix} 1-\frac{\alpha}{N+2} & \frac{\alpha}{N+2} & & & & & & \\ & 1-\frac{1-q}{N+2} & \frac{1-q}{N+2} & & & & & \\ & \frac{q}{N+2} & 1-\frac{1+\alpha}{N+2} & \frac{1-q}{N+2} & \frac{\alpha}{N+2} & & & \\ \frac{\beta}{N+2} & & \frac{q}{N+2} & 1-\frac{q+\alpha+\beta}{N+2} & & \frac{\alpha}{N+2} & & \\ & & & & 1-\frac{1-q}{N+2} & \frac{1-q}{N+2} & & \\ & \frac{\beta}{N+2} & & & \frac{q}{N+2} & 1-\frac{1+\beta}{N+2} & \frac{1-q}{N+2} & \\ & & \frac{\beta}{N+2} & & & \frac{q}{N+2} & 1-\frac{q+\alpha+\beta}{N+2} & \frac{\alpha}{N+2} \\ & & & & \frac{\beta}{N+2} & & & 1-\frac{\beta}{N+2} \end{Vmatrix} \end{array}$$

Using the transition matrix we are able to get the probability distribution π_t for which state the chain will be in at any time step t.

$$\pi_t = \pi_0 \mathbf{P}^t$$

Where π_0 is the starting distribution and π_t is the distribution at time t. For example, if we start in State 2, then $\pi_0 = [0, 1, 0, 0, 0, 0, 0, 0]$. After some number of iterations, the distribution will reach an equilibrium. This distribution is called the stationary distribution π. If the Markov chain is is time-homogeneous, irreducible, and all its states are positive recurrent, then there exists a *unique pi*. Essentially this means if each state is eventually reached then there is an equilibrium distribution independent of which state the process starts in.

Definition 3.2. A Markov chain is *time-homogeneous* if the transition probabilities from state to state stay the same between steps.

Definition 3.3. A Markov chain is *irreducible* if for each state s_i in the set of possible states S, it is possible to reach all states $s_j \in S$ including (i.e. each state is reachable from each state). Explicitly,

$$\mathbb{P}[s_i \to s_j] > 0 \qquad\qquad \forall s_i, s_j \in S$$

Definition 3.4. A state is *positive recurrent* if the expected time to return to that state is finite.

Definition 3.5. The *stationary distribution* is the proportion of time a Markov chain spends in each possible state over the long run, regardless of the starting state. In other words, π is the vector such that:

$$\pi = \pi \mathbf{P}$$

It follows that π is akin to an eigenvector of $\mathbf{P}$.

The structure of our open Markov chain example means that all three of these conditions are satisfied (depending on the values of q, α, β), leading to Theorem 3.6.

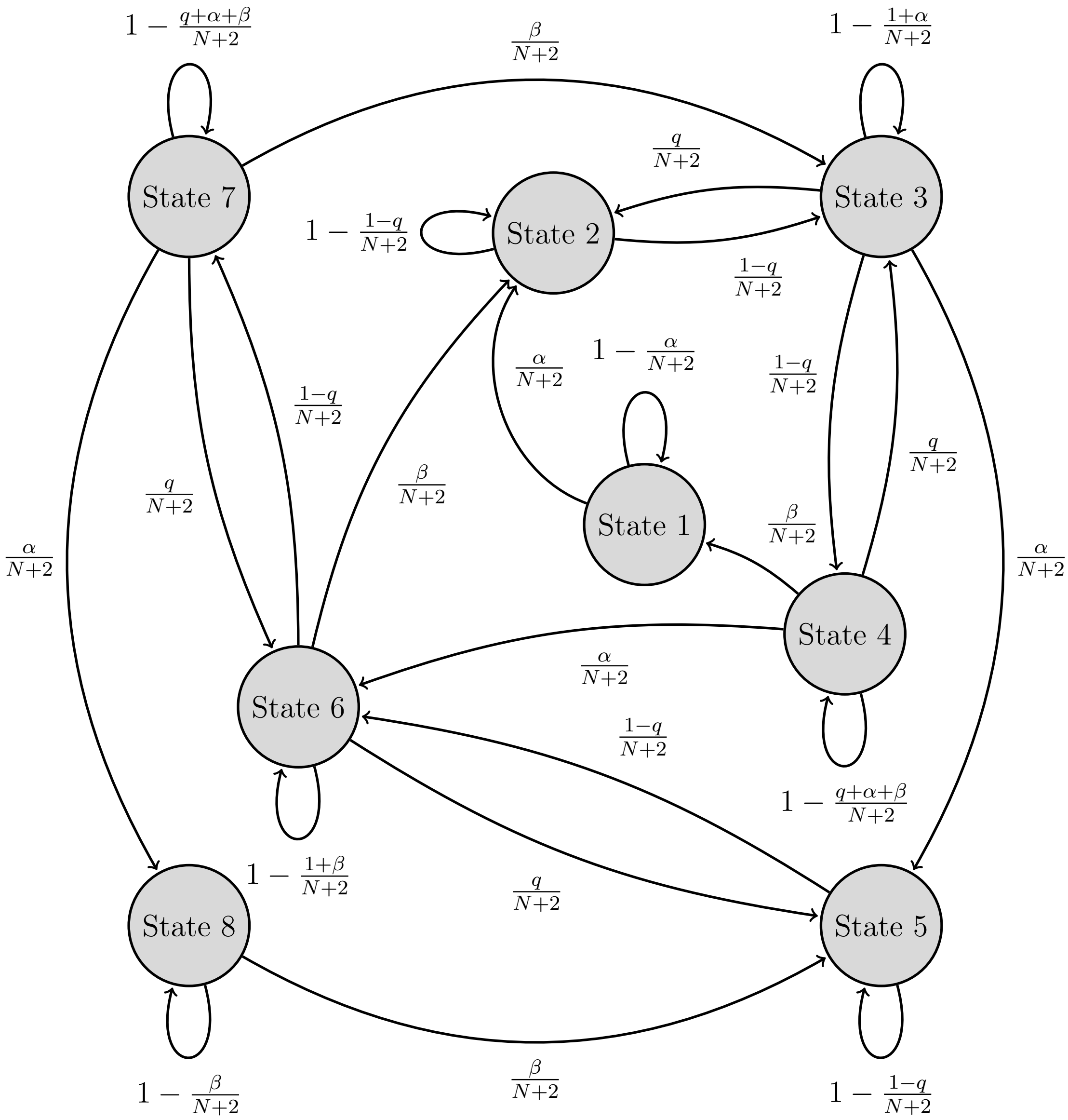


FIGURE 5. Open Markov Chain for $N = 3$.

Theorem 3.6. *There exists a stationary distribution π for the Open Markov chain with $N = 3$. The stationary distribution is a function of q, α, and β, and is explicitly derived for $N = 3$ as*

$$\pi = \frac{1}{D}\begin{bmatrix} \beta^3(q^2-2q+1) \\ \alpha\beta^2(q^2+q\alpha+\beta) \\ -\alpha\beta^2(q^2+q\alpha+q\beta-1-\alpha-\beta) \\ \alpha\beta^2(q^2-2q+1) \\ \alpha^2\beta(q^2+2q\beta+\alpha\beta+\alpha+\beta^2) \\ -\alpha^2\beta(q^2+q\alpha+q\beta-q-\alpha-\beta) \\ \alpha^2\beta(q^2-2q+1) \\ \alpha^3(q^2-2q+1) \end{bmatrix}$$

Where

$$\begin{aligned} D &= q^2\alpha^3 + q^2\alpha^2\beta + q^2\alpha\beta^2 + q^2\beta^3 - q\alpha^3\beta \\ &- 2q\alpha^3 + q\alpha^2\beta^2 - q\alpha^2\beta - q\alpha\beta^3 - q\alpha\beta^2 \\ &- 2q\beta^3 + \alpha^3\beta^2 + 2\alpha^3\beta + \alpha^3 + \alpha^2\beta^3 \\ &+ 2\alpha^2\beta^2 + \alpha^2\beta + 2\alpha\beta^3 + \alpha\beta^2 + \beta^3 \end{aligned}$$

Further, there exists is a stationary distribution π $\forall N$.

Proof. Per Definition 3.5 the stationary distribution is similar to an eigenvector of the transition matrix. Specifically, it is the eigenvector corresponding to the $\lambda = 1$ eigenvalue of

$$\pi = \pi\mathbf{P}$$

Rearranging,

$$(\mathbf{I} - \mathbf{P})\pi = \vec{0}$$

By the law of total probability we know that the entries of π must sum to 1. To account for this on the LHS we append a row vector of $1's$ to $(\mathbf{I} - \mathbf{P})$, call this new matrix A. On the RHS we append a 1 to the end of the zero vector, call this new vector b.

$$\mathbf{A}\pi = b$$

Solving for π, we obtain the stationary distribution. □

Theorem 3.7. *There exists a unique stationary distribution π for the Open Markov chain if $\alpha, \beta > 0$ and $q < 1$, $\forall N$.*

Proof. For $q = 1$, particles will always move to their leftmost possible position in the Markov chain. Once this configuration is reached there will be no particle movement, thus states in the chain may fail the condition for positive recurrence and the chain as a whole is now reducible.

For $\alpha, \beta = 0$, there is a similar argument. If $\alpha = 0$ then no new particles will enter the chain and eventually all will exit. If $\beta = 0$ then eventually the chain will be completely populated with particles. In either case, the chain fails irreducibility and positive recurrence. □

Additionally, we can examine how varying q over $[0, 1)$ affects the unique stationary distribution.

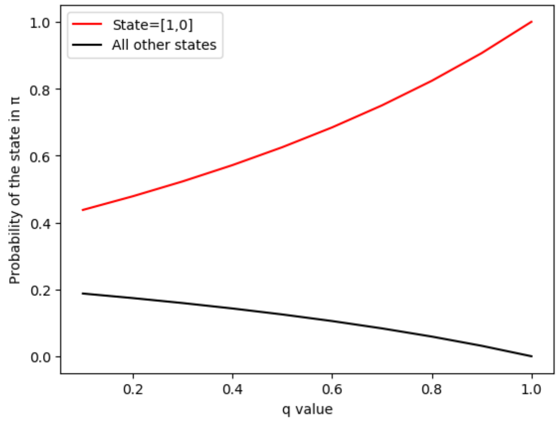
Example 3.8. Consider the cases

$$N = 2 \qquad\qquad N = 3$$

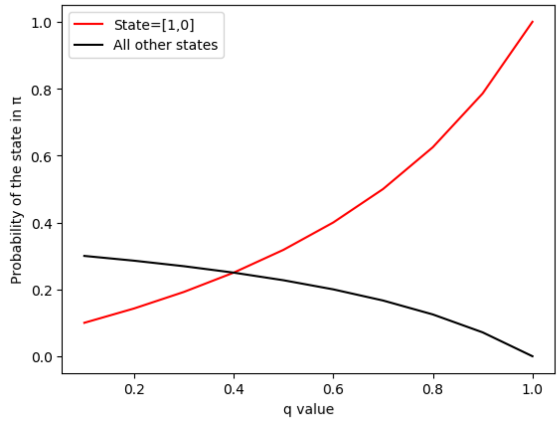
Varying q over $[0, 1)$, we obtain the graphs in Figures 6 & 8 for different relationships between α and β.

From these experimental results, it is clear to see that when $\alpha = \beta$, there is significant overlap between the distribution curves.

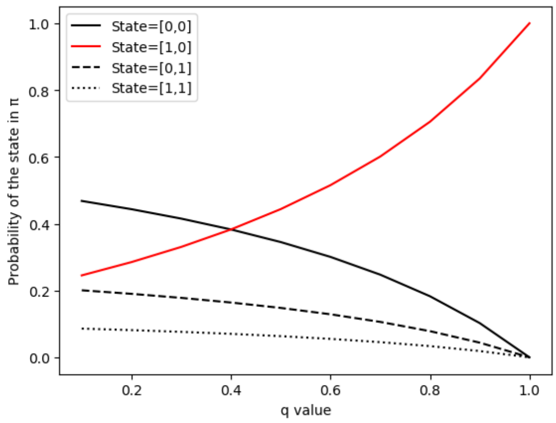
Notice that in the graphs in Example 3.8 the number of probability clusters corresponds to the value of N, leading to Conjecture 3.10.



(A) $\alpha, \beta = 1$

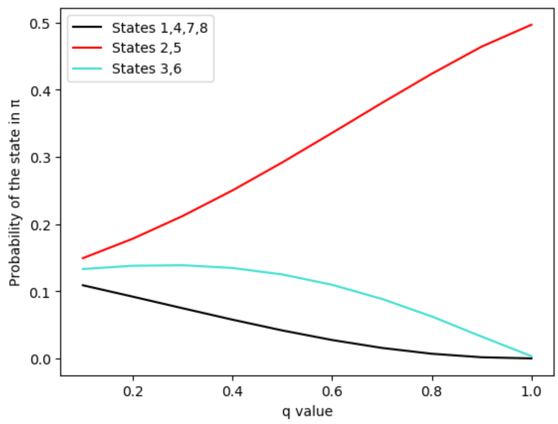


(B) $\alpha, \beta = 0.1$

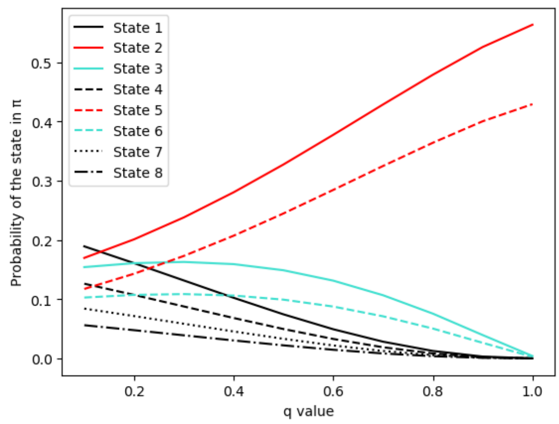


(C) $\alpha = 0.3,\ \beta = 0.7$

FIGURE 6. Stationary distributions for N=2 as a function of q.



(A) $\alpha, \beta = 0.5$



(B) $\alpha = 0.3,\ \beta = 0.7$

FIGURE 7. Stationary distributions for N=3 as a function of q (see Section 3.1 for state visualizations).

Definition 3.9. A *cluster* Γ is a grouping of states that all have equal probability of occurring. For example, consider some random events A, B, and C.

$$\mathbb{P}[A] = \mathbb{P}[B] = \frac{1}{4}$$
$$\mathbb{P}[C] = \frac{1}{2}$$

In this case A and B are in a cluster, and C is in another.

$$\Gamma_1 = \{A, B\}$$
$$\Gamma_2 = \{C\}$$

Additionally, let the total number of clusters in a distribution equal γ.

Conjecture 3.10. For the Open Markov chain with $\alpha = \beta$ the number of clusters γ in the stationary distribution π is equal to N, regardless of the value of q.

$$\gamma_{\alpha=\beta} = N$$

If follows that that γ is bounded below by N and above by 2^N.

$$N \leq \gamma \leq 2^N$$

3.2. **Mixing Time on Open Loop.** TODO

Massachusetts Institute of Technology, Department of Mathematics, Cambridge MA 02139