Universite de Nice Sophia-Antipolis
Erasmus Mundus programm ”Mathmods”
The report of internship

Supervisor : Mr.Nicolas Champagnat

Stochastic models in population dynamics

Bolor Jargalsaikhan

INRIA, Sophia Antipolis, July 24, 2009

Outline

1. Introduction

Keywords: Markov chain, Quasi-stationary distribution, Birth and Death process, Particle method.

Markov chain, a stochastic process with Markov property, has a wide application in biological modeling, finance, gambling and etc. In particular, birth and death chain where the states represent the current size of a population and where the transitions are limited to birth and death is a special case of homogenous Markov Chain.

When homogenous markov chain is irreducible, positive recurrent and aperiodic, then we have a theory stating that starting from an arbitrary initial distribution, it will converge almost surely to the invariant distribution. Therefore, the asymptotic behavior of a markov chain can be characterized by an invariant distribution. However, in the case where absorbtion is certain, we have a trivial invariant distribution $(1, 0, 0, ..)$ (not necessarily unique). It is interesting that what happens before absorbtion. One way to describe "the asymptotic behavior of the Markov chain before absorbtion" is the Quasi-stationary distribution.

There are many articles where they studied the existence of a quasi stationary distribution of an irreducible and aperiodic markov chain and its asymptotic behavior.

The goal of this document is to describe the existing theory , to extend results to 2D birth and death chain which is reducible and to simulate the conditioned processes.

In the first part of the report, the quasi-stationary distribution of an absorbing Markov chain is introduced. From the papers of J.N.Darroch and E. Seneta, we see that the problem of finding a quasi-stationary distribution, which is an invariant distribution conditioned on non absorbtion, of irreducible, aperiodic, finite Markov chain is equivalent to finding the left eigenvector corresponding to the maximal eigenvalue of a substochastic, primitive matrix. For infinite countable state space, there are some more restriction to ensure the above statement.

Next, we study quasi-stationary distribution of birth and death process in 2D. This problem is the stochastic model of a competition between two species before extinction. Therefore, it gives an answer to questions "Which species will survive?" or "Can they coexist?".

In reducible case, we see that quasi stationary distribution exists but not unique and depends on initial state. Also, we consider a special case where we can find the form of eigenvectors or eigenvalues of substochastic matrix which is related to finding quasi stationary distribution.

Instead of checking maximal eigenvalues and corresponding eigenvectors, the simulation of markov chain conditioned on non-absorbtion is an another approach to answer above questions. This can be done by particle method algorithm. Using the particle method, we study particular cases of birth and death processes and we analyze the behavior of the the process depending on its parameters.

2. Preliminary

2.1 Discrete-time Markov chains

Sequences of independent and identically distributed random variables are stochastic processes. Discrete time homogenous Markov chains are a very special class of stochastic processes which is allowed for some dependence on the past. However, the probabilistic dependence on the past is only through the previous state. Also, this process can be represented by deterministic recurrence equations of random variables.

Definition $(X_n)_{n\geqslant 0}$ is a homogenous Markov Chain with initial distribution λ and transition matrix P if
(i) X_0 has distribution λ;
(ii) For $n \geqslant 0$, conditional on $X_n = i, X_{n+1}$ has distribution $(p_{ij} : j \in I)$ and is independent of $X_0, ..., X_{n-1}$.

P.S. Countable state space and homogenous Markov chains are considered throughout this report . Also transition matrix is not restricted to finite size.

Definition The probability starting from i that $(X_n)_{n\geqslant 0}$ ever hits A is then $h_i^A = P_i(H^A < \infty)$. When A is a closed class, h_i^A is called absorption probability.

Example 2.1.1(Birth and death chain) For $i = 1, 2, ...,$ we have $0 < p_i < 1$ birth and $q_i = 1 - p_i$ death probability. 0 is an absorbing state.
We are going to calculate absorption probability starting from state i.
$h_0 = 1, h_i = p_i h_{i+1} + q_i h_{i-1}$, for $i = 1, 2, ...$ $p_i + q_i = 1$. After solving recurrence relation, we will get $h_i = 1 - A(\gamma_0 + ... + \gamma_{i-1})$ where $\gamma_0 = 1, A = h_0 - h_1, \gamma_i = \frac{q_i q_{i-1} ... q_1}{p_i p_{i-1} ... p_1}$.

In the case $\sum_{i=0}^{\infty} \gamma_i = \infty$, the restriction $0 \leqslant h_i \leqslant 1$ forces $A = 0$ and $h_i = 1$. However, if $\sum_{i=0}^{\infty} \gamma_i < \infty$ then we can take $A > 0$ as long as $1 - A(\gamma_0 + ... + \gamma_{i-1}) \geqslant 0$ for all i. Therefore $A = (\sum_{i=0}^{\infty} \gamma_i)^{-1}$ and $h_i = \sum_{j=i}^{\infty} \gamma_j / \sum_{j=0}^{\infty} \gamma_j$.

Definition We write $i \rightarrow j$ if $P_i(X_n = j$ for some $n \geqslant 0) > 0$.
We say i communicates with j and write $i \leftrightarrow j$ if both $i \rightarrow j$ and $j \rightarrow i$.
The relation "$\leftrightarrow$" is reflexive, symmetric and transitive. Therefore, it is an equivalence relation on I, and thus partitions I into communicating classes.

Definition We say that a class C is closed if $i \in C, i \rightarrow j$ imply $j \in C$. The state i is absorbing if $\{i\}$ is a closed class. A transition matrix P where I is a single class is called irreducible.

Definition Let $(X_n)_{n\geqslant 0}$ be a Markov chain with transition matrix P. We say that a state i is recurrent if $P_i(X_n = i$ for infinitely many $n) = 1$. We say that i is transient if $P_i(X_n = i$ for infinitely many $n) = 0$.

Example2.1.2 1D Random Walk We take state space as $E = \mathbb{Z}$. The nonzero terms of its transition matrix are $p_{i,i+1} = p, p_{i,i-1} = 1 - p$, where $p \in (0, 1)$. The nature of its any state can be verified. For example, let's take 0. $p_{00}(2n + 1) = 0$ and $p_{00}(2n) = \frac{(2n!)}{n!n!} p^n (1 - p)^n$.
By using Stirling's equivalence formula $n! \sim (n/e)^n \sqrt{2\pi n}$, $p_{00}(2n) \sim \frac{[4p(1-p)]^n}{\sqrt{\pi n}}$.
If $p \neq 1/2$, then $4p(1 - p) < 1$ and the series converges. Therefore, in this case it is transient. For $p = 1/2$, the series diverges, so it is recurrent.

Invariant distributions

Definition We say a measure λ is invariant if $\lambda P = \lambda$.

Example 2.1.3(Birth and Death chain) We take state space $E = N$, 0 is an absorbing state. From state i, there is a birth with probability p_i to state $i+1$, death with q_i to state $i - 1$ and with probability r to stay at state i.

To be an invariant measure, it must satisfy
$\pi(i) = p_{i-1}\pi(i - 1) + r_i\pi(i) + q_{i+1}\pi(i + 1)$ for $i \in N$, and $\pi(0) = \pi(1)q_1$.
From this recurrence relation, we will get $\pi(i) = \pi(0)\frac{p_1 p_2 \dots p_{i-1}}{q_1 q_2 \dots q_i}$.

For this solution to be an invariant distribution, it must be a probability.
Therefore, $\pi(0) > 0, \sum_{i=1}^{\infty} \pi(i) = 1$.
$\pi(0)\{1 + \frac{1}{q_1} + \sum_{i=1}^{\infty} \frac{p_1 p_2 \dots p_{i-1}}{q_1 q_2 \dots q_i}\} = 1$.
Thus a stationary distribution exists if and only if $\sum_{i=1}^{\infty} \frac{p_1 p_2 \dots p_{i-1}}{q_1 q_2 \dots q_i} < \infty$.

Theorem 2.1.1(The uniqueness of Invariant measure) The invariant measure of an irreducible recurrent stochastic matrix is unique up to a multiplicative factor.

Theorem 2.1.2 An irreducible recurrent HMC has a unique stationary distribution if its invariant measure x satisfy $\sum_{i \in E} x_i < \infty$.

Example 2.1.4 As we have seen from example 2.1.2, random walk in 1D , when p=1/2, it is recurrent and irreducible. If we take $\pi_i = 1$ for all i. Then $\pi_i = \frac{1}{2}\pi_{i-1} + \frac{1}{2}\pi_{i+1}$, so π is invariant. By theorem 2.1.1, any invariant measure must be scalar multiple of π. Since $\sum_{i \in Z} \pi_i = \infty$, there can be no invariant distribution.

On the other hand, if $p \neq 1/2$ then it is transient. We can construct counter example of uniqueness of invariant distribution. In this case, we have
$\pi_i = A + B(p/q)^i, A, B \in R$ two parameter family of invariant measures.

Convergence to equilibrium

Definition Let us call a state i aperiodic if $p_{ii}^{(n)} > 0$ for all sufficiently large n.

Lemma 2.1.5 Suppose P is irreducible and has an aperiodic state i. Then, for all states j and k, $p_{jk}^{(n)} > 0$ for sufficiently large n. In particular, all states are aperiodic.

Theorem 2.1.6 Let P be irreducible and aperiodic , and suppose that P has an invariant distribution π. Let λ be any distribution. Suppose that $(X_n)_{n \geqslant 0}$ is Markov(λ, P). Then $P(X_n = j) \to \pi_j$ as $n \to \infty$ for all j.

In particular, $p_{i,j}^{(n)} \to \pi_j$ as $n \to \infty$ for all i, j.

2.2 Continuous time Markov processes

Definition A stochastic process $\{X(t), t \geqslant 0\}$, defined on a probability space (Ω, F, P) with values in a countable set S, called the state spaces of the process, is called a continuous-time Markov chain if for any finite set of times $0 \leqslant t_1 < t_2 < ... < t_n < t_{n+1}$ and corresponding set $i_1, i_2, ..., i_{n-1}, i, j$ of states of S such that $P\{X(t_n) = i, X(t_{n-1}) = i_{n-1}, ..., X(t_1) = i_1\} > 0$, we have

$$P\{X(t_{n+1}) = j | X(t_n) = i, X(t_{n-1}) = i_{n-1}, ..., X(t_1) = i_1\} = P\{X(t_{n+1}) = j | X(t_n) = i\}.$$

If for all s, t such that $0 \leqslant s \leqslant t$ and all $i, j \in S$ the conditional probability $P\{X(t) = j | X(s) = i\}$ appearing on the right hand-side of equation depends only on $t - s$, then the process is homogeneous.

Definition $P_{ij}(t) = P\{X(t) = j | X(0) = i\}$, $i, j \in S, t \geqslant 0$ is called the transition function of the process. Let $P_{ij}(t)$ be a transition function. A matrix Q whose (i, j)th component is the number $q_{ij} = P'_{ij}(0)$ is called infinitesimal generator of stochastic process X.

Proposition2.2.1 $P'_{ii}(0) = -\lim_{t \to 0}[1 - P_{ii}(t)]/t = -q_i$ exists, but may be ∞. Moreover, $q_i = 0$ if and only if $P_{ii}(t) = 1$ for all $t \geqslant 0$.
A state i is said to be an absorbing state if $q_i = 0$.

Birth and death process

Let's consider a continuous time Markov chain $\{X(t), t \geqslant 0\}$ with state space $S = \{0, 1, 2...\}$ with stationary transition probabilities $P_{ij}(t)$, i.e., $P_{ij}(t) = P\{X(t + s) = j | X(s) = i\}$. In addition we assume that $P_{ij}(t)$ satisfy the following postulates:
1) $P_{i,i+1}(h) = \lambda_i h + o(h)$, as $h \to 0, i \geqslant 0$.
2) $P_{i,i-1}(h) = \mu_i h + o(h)$, as $h \to 0, i \geqslant 0$.
3) $P_{ii}(h) = 1 - (\lambda_i + \mu_i)h + o(h)$, as $h \to 0, i \geqslant 0$.
4) $P_{ij}(0) = \delta_{ij}$.
5) $\mu_0 = 0, \lambda_0 > 0, \mu_i, \lambda_i > 0, i = 1, 2, ...$
These $o(h)$ in each case may depend on i. The matrix

$$Q = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \\ \mu_1 & -(\mu_1 + \lambda_1) & \lambda_1 & 0 & \ddots \\ 0 & \mu_2 & -(\mu_2 + \lambda_2) & \lambda_2 & \\ & \ddots & & \ddots & \end{bmatrix}$$

is called the infinitesimal generator of the process. The parameters λ_i, μ_i are called, respectively, the infinitesimal birth and death rates. The process $X(t)$ is known as a birth and death process.

In postulate 1 and 2 we are assuming that if the process starts in state i, then in a small interval of time the probabilities of the population increasing or decreasing by 1 are essentially proportional to the length of the interval. Since the $P_{ij}(t)$ are probabilities, we have $P_{ij}(t) \geqslant 0$ and $\sum_{j=1}^{\infty} P_{ij}(t) = 1$.

Perron Frobenius Theorem Let A be a nonnegative irreducible and aperiodic rxr matrix. $\exists$ eigenvalue $\lambda_1 \in R^+$ such that $\lambda_1 > \lambda_i$ for any other eigenvalue λ_i. Moreover, the left eigenvector v and the right eigenvector ω ,such that $v'\omega = 1$, corresponding to λ_1 can be chosen positive. Furthermore,

$$A^n = \lambda_1^n \omega v' + o(n^{m_2-1}|\lambda_2|^n)$$

where λ_2 is the second biggest eigenvalue by modulus, and m_2 is the multiplicity of λ_2.

Corollary Let's suppose all states in A communicates, then there exists eigenvalue ρ_1 of C, with maximal real part. It is real, simple, and less than zero. Also, there are corresponding unique positive left and right eigenvectors v, ω such that $v'\omega = 1$.
$Q(t) = \exp Ct = e^{t\rho_1}\omega v' + o(e^{t\rho'})$ where $v'e = 1$ e is an unit vector and $\rho' < \rho_1$.

3.Quasi-stationary distribution

3.1 Quasi-stationary distribution of finite, discrete time markov chains

We consider a Markov chain, $(X_n, n \geqslant 0)$, in which there is a set T of transient states from which the process is certain to be absorbed into the remaining states. Let's suppose the chain have states $0, 1, 2, ..., s$ with transition matrix

$$P = \begin{bmatrix} 1 & 0' \\ p_0 & Q \end{bmatrix}$$

$p_0 \neq 0$, Q is $s \times s$ irreducible, aperiodic substochastic matrix and $p_0, 0$ are $s \times 1$ vectors.

Definition Let $[q_0(n), q'(n)]$ denote the probability distribution of X_n over all states (s+1) states at time n and denote by $d(n)$ the conditional distribution

$$d(n) = \frac{q(n)}{1 - q_0(n)}.$$

Equivalently, we can say $d_k(n) = P_{d(0)}(X_n = k | X_n \neq 0)$.

Definition If $d(n+1) = d(n) = d$ then, we call d a quasi stationary distribution. Equivalently, we can say that $P_{d(n)}(X_1 = k | X_1 \neq 0) = d_k(n) = d_k$.

Proposition3.1.1 d is a quasi-stationary distribution if and only if d is the left eigenvector of Q with non-negative components.
$\lhd$ $A = [q_0(n), q'(n)]P = [q_0(n+1), q'(n+1)] \Rightarrow A = [q_0(n) + q'(n)p_0, q'(n)Q] = [q_0(n+1), q'(N+1)]$. Therefore d must satisfy $d'Q = \rho d'$. Q is irreducible and

the elements of d are non-negative, so it follows from extended Perron Frobenius theorem that ρ is the real maximal eigenvalue of Q and $d = v$, where v is normalized non-negative left eigenvector of Q corresponding to maximal eigenvalue. $\triangleright$

Limiting conditional distribution We can interpret v as an asymptotic behavior of conditional distribution. Let's denote $P^n = (p_{ij}^{(n)})$.

Proposition3.1.2 Conditional distribution of discrete time finite Markov chain, starting from an arbitrary initial distribution π, converges to the quasi-stationary distribution.

$\triangleleft$ If the process starts in state i with probability π_i, then the probability that it has been absorbed by time n is $\sum\limits_{i\in T} \pi_i p_{i0}^{(n)}$, and given that it is still in T, the conditional probability that it is in state j at time n is

$$P_\pi(X_n = j|X_n \neq 0) = \frac{\sum\limits_{i\in T} \pi_i p_{ij}^{(n)}}{\sum\limits_{i\in T} \pi_i(1 - p_{i0}^{(n)})} = \frac{\pi' Q^n f_j}{\pi' Q^n e}$$

where f_j is $s \times 1$ vector with unity in the j-th row and zeros elsewhere.

Since Q is irreducible and aperiodic, by Perron Frobenius theorem

$Q^n = \rho_1^n \omega v' + o(n^k |\rho_2|^n)$ where ρ_1- maximal eigenvalue of Q, ω, v are corresponding normalized, positive right and left eigenvector, ρ_2- second biggest eigenvalue by module, $k = m_2 - 1$ (m_2 - multiplicity of ρ_2).

$$B = \frac{\sum\limits_{i\in T} \pi_i p_{ij}^{(n)}}{\sum\limits_{i\in T} \pi_i(1 - p_{i0}^{(n)})} = v_j + o(n^k (\frac{|\rho_2|}{\rho_1})^n).$$

When we take a limit as $n \to \infty$, $B \to v_j$. $\triangleright$

3.2. Quasi-stationary distribution of discrete time Markov chain with countable infinite states

When it is infinite state space, there are 2 distinct features from finite state space.

1) Absorbtion is no longer certain.

2) Quasi-stationary distribution depend on initial distribution.

From paper "Ergodic properties of non-negative matrices" by Vere-Jones, D., a finite, irreducible, substochastic matrix Q has the property that its convergence parameter R (the radius of convergence of the functions $P_{ij}(z) = \sum_n p_{ij}^{(n)} z^n$) is strictly greater than unity, and Q is R-positive. However, in infinite case, convergence radius is not necessarily greater than unity unless the matrix is R-positive.

Lemma3.2.1 The following conditions on Q are equivalent, and each implies that the matrix has a convergence parameter R and is R-positive:

1) For some i and j, the sequence $p_{ij}^{(n)} R^{(n)}$ tends to a finite non-zero limit as $n \to \infty$.

2) There exists non-negative, non-zero vectors $\{v_k\}, \{\omega_k\}$ such that $R \sum_{k \in T} v_k p_{kj} = v_j (j \in T)$, $R \sum_{k \in T} p_{ik} \omega_k = \omega_i (i \in T)$ and $\sum_{k \in T} v_k \omega_k < \infty$.

Lemma3.2.2 Suppose that Q is a substochastic matrix with convergence parameter R, and suppose Q is R-positive. Then either $R > 1$, or $R = 1$ and the matrix is stochastic.

We can see the proof of above 2 lemmas from the article "On quasi-stationary distributions in Markov chains with a denumerable infinity of states" by E. Seneta and D. Vere-Jones.

Here, it is the analogue theorem for finite chains described in section 2.

Theorem(Main limit theorem) Suppose that Q is irreducible, aperiodic, and substochastic. If Q is R-positive with $R > 1$, and the left eigenvector $\{v_k\}$ satisfies the condition $\sum v_k < \infty$ then the limit of $P_i(X_n = j | X_n \neq 0) = \frac{p_{ij}^{(n)}}{\sum_{k \in T} p_{ik}^{(n)}}$ exists and defines a probability.

$\lhd$ There is a theorem stating that an individual sequence $\{p_{ij}^{(n)} R^n, n = 1, 2, 3 ...\}$ tend to finite non-zero limits λ_{ij} which can be evaluated in terms of the left and right eigenvectors. $\lambda_{ij} = (v_k \omega_k) / \sum_{k \in T} v_k \omega_k$. ("Geometric ergodicity in denumerable Markov chains" by D.Vere-Jones).

There is also another theorem stating that when Q is R-positive, the necessary and sufficient condition for the sums $\sum_{k \in T} p_{ik}^{(n)} R^n$ to tend to a finite limit as $n \to \infty$ is the convergence of series $\sum v_k < \infty$. ("Ergodic properties of non-negative matrices" by D. Vere-Jones).

Therefore we have

$$\lim_{n \to \infty} P_i(X_n = j | X_n \neq 0) = \frac{\lim_{n \to \infty} p_{ij}^{(n)} R^n}{\lim_{n \to \infty} \sum_{k \in T} p_{ik}^{(n)} R^n} = \frac{v_j}{\sum_{k \in T} v_k} = v_j . \rhd$$

Extensions to an arbitrary initial distribution

Theorem 3.2.1 Suppose that Q is irreducible, aperiodic and substochastic and has a convergence parameter $R > 1$. Then if the quantities $\frac{\sum_{i \in T} \pi_i p_{ij}^{(n)}}{\sum_{i \in T} \sum_{j \in T} \pi_i p_{ij}^{(n)}}$ tend to limits v_j which form a probability distribution, this distribution is a left non-negative eigenvector for some some eigenvalue ρ in the range $1/R \leqslant \rho < 1$. Conversely every left eigenvector satisfying condition $\sum v_k < \infty$ can be reached as a limit of the conditional probabilities for a suitable choice of the initial distribution π.

For the later use, as a test problem for simulation of quasi-stationary distribution we now introduce some cases where we can analytically can find the quasi stationary distribution.

The semi-infinite simple random walk with absorbtion The matrix Q which describes the simple random walk on the non-absorbing states $T = \{1, 2, ...\}$ is irreducible and periodic with period 2.

$$Q = \begin{bmatrix} 0 & b & 0 & 0 & \\ a & 0 & b & 0 & \ddots \\ 0 & a & 0 & b & \\ & \ddots & & \ddots & \end{bmatrix}$$

where $a + b = 1$.

Let's consider the case $a \geqslant b$(absorbtion is certain). Let's denote $v_{ij} = \frac{p_{ij}^{(n)}}{\sum\limits_{k=1}^{\infty} p_{ik}^{(n)}}$

From direct calculation we can check that, we have left and right eigenvectors $\{v_j\}, \{\omega_j\}$ corresponding to the eigenvalue $2\sqrt{ab}$, where

$v_j = v_1 j\left(\sqrt{\frac{b}{a}}\right)^{j-1}$, $j = 1, 2...$ and $\omega_j = \omega_1 j\left(\sqrt{\frac{a}{b}}\right)^{j-1}$, $j = 1, 2....$

In article "On Quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states" by E. Seneta and D. Vere-Jones, they found the quasi stationary distribution explicitly using combinatorics, and approximating probability of first passage to 0 by Stirling's formula.

$$\lim_{n\to\infty} \frac{\sum_{m=1}^{n} v_{ij}^{(m)}}{n} = c_j j\left(\sqrt{\frac{b}{a}}\right)^{j-1} = c_j v_j$$

independently of the initial state, where

$$c_j = \begin{cases} 0 \text{ ,if } a = b = 1/2 \\ \frac{1}{2}[\frac{1-4ab}{a}] \text{ ,if } a > b, j \text{ odd} \\ \frac{1}{2}[\frac{1-4ab}{a}\frac{1}{2\sqrt{ab}}] \text{ , if } a > b, j \text{ even} \end{cases}$$

The extended semi-infinite simple random walk with absorbtion When we to allow to pause in any state, we will obtain more general matrix of the following form: (the matrix consists of non-absorbing states $T = \{1, 2, ...\}$, excluding only absorbing state $\{0\}$)

$$Q1 = \begin{bmatrix} c & b & 0 & 0 & \\ a & c & b & 0 & \ddots \\ 0 & a & c & b & \\ & \ddots & & \ddots & \end{bmatrix}$$

where $a+b+c = 1, a > 0, b > 0, c > 0$ and the substochastic matrix Q_1 is aperiodic. If we want to find the Quasi-stationary distribution following the same technique used in the previous article, the problem becomes rather less manageable (they also mentioned this). However we can prove that this problem's solution is as the same as the previous problem's solution. Therefore,

$$P_i(X_k = j|X_k \neq 0) \rightarrow v_j$$

where $\sum_{j=1}^{\infty} v_j = 1$, if $a > b$.

$\lhd$ First we note that $\{v_j\}$ and $\{\omega_j\}$ of the previous problem are the eigenvectors of the $Q1$ corresponding to eigenvalue $(c + 2\sqrt{ab})$. $Q1$ is irreducible and elements of v is nonnegative, therefore by extended Perron Frobenius theorem v is the Quasi-stationary distribution of Q1. $\rhd$

3.3. Quasi stationary distribution of continuous time, finite Markov Chain

Let's consider $(X_t, t \geqslant 0)$ finite, birth and death process. We take as before "0" as only absorbing state, $T = \{1, 2, ..., n\}$ transient states. Then the matrix R of infinitesimal transition probabilities q_{ij} ($q_{ij} \geqslant 0, i \neq j, \sum_{j=0}^{n} q_{ij} = 0$) have the following form:

$$R = \begin{bmatrix} 0 & 0' \\ q_0 & C \end{bmatrix}$$

where $q_0 \neq 0$ and the matrix C corresponds to the transient set T. $q_0 \neq 0$ condition ensures that absorbtion will eventually occur from any state of T.

As we mentioned in preliminary part $i \neq j$ $P(X_{t+s} = j|X_t = i) = q_{ij}s + o(s)$, we are assuming that if the process starts in state i, then in a small interval of time the probabilities of the population increasing or decreasing by 1 are essentially proportional to the length of the interval.

The matrix of transition function $P(t) = \{P_{ij}(t)\}$, the transition probability from state i to state j at time t, is $P(t) = \exp Rt$ (exponential matrix).

If we denote $Q(t) = \exp Ct$, and

$$P(t) = \begin{bmatrix} 1 & 0' \\ p_0(t) & Q(t) \end{bmatrix}$$

where $p_0(t) \neq 0, t > 0$. Therefore, this matrix form is analogous form to the discrete time case.

Definition Let's suppose, the probability distribution over all (n+1) states at time t is $[\pi_0(t), \pi'(t)]$ then the conditional distribution restricted to the transient set T is

$$d(t) = \frac{\pi(t)}{1 - \pi_0(t)}.$$

Equivalently, we can say $d_k(t) = P_{d(0)}(X_t = k|X_t \neq 0)$.

Definition If $d(t) = d, t \geqslant 0$ where $d'e = 1$ then, we call d a quasi stationary distribution. Equivalently, we can say that $P_d(X_t = k|X_t \neq 0) = d_k$ $(\forall t \geqslant 0, \forall k \geqslant 1)$.

Proposition3.3.1 There is an unique quasi-stationary distribution for continuous time birth and death process if all states of T communicates. Moreover $d = v$, where v is the left eigenvector corresponding to maximum eigenvalue.

$\triangleleft$ Since $\pi'(t_1)Q(t_2) = \pi'(t_1+t_2), t_1, t_2 \geqslant 0$, we have that $d'(t_1)Q(t_2) = \rho(t_1, t_2)d'(t_1+t_2)$, where ρ is a function depending on t_1 and t_2. If d exists it satisfies $d' \exp Ct_2 = \rho(t_1, t_2)d'$. For fixed $t_2 > 0$, $Q(t_2)$ is irreducible aperiodic, substochastic matrix. Therefore, it has the unique maximal eigenvalue $\exp \rho 1 t_2$ and eigenvectors ω, v', where ρ, ω, v' are the corresponding quantities for C. It follows that if d exists (non-negative) then $d = v, \rho(t_1, t_2) = \exp \rho t_2$.
Conversely putting $\pi(0) = v$, $d(t) = d, t \geqslant 0$. Hence we have a unique quasi-stationary distribution $d = v$. $\triangleright$

Proposition3.3.2 When we start from an arbitrary π initial distribution, it approaches to v with an exponential rate as $n \to \infty$.

$\triangleleft$ We consider the probability that the process is in state $j \in T$ at time t, given that it has not yet been absorbed, and started from state $i \in T$ with probability π_i. This is, for $j \in T$,

$$P_\pi(X(t) = j|X(t) \neq 0) = \frac{\sum\limits_{i \in T} \pi_i p_{ij}(t)}{\sum\limits_{i \in T} \pi_i (1 - p_{i0}(t))} = \frac{\pi' Q(t) f_j}{\pi' Q(t) e}$$

where f_j is $s \times 1$ vector with unity in the j-th row and zeros elsewhere and e is an unit vector. By the Perron frobenius theorem, we have

$$P_\pi(X(t) = j|X(t) \neq 0) = \frac{\sum\limits_{i \in T} \pi_i p_{ij}(t)}{\sum\limits_{i \in T} \pi_i (1 - p_{i0}(t))} = v_j + o(e^{t(\rho' - \rho)}),$$

where $\rho' < \rho < 0$. Therefore, $P_\pi(X(t) = j|X(t) \neq 0) \to v_j$ as $t \to \infty$, independently of initial distribution. $\triangleright$

4.Birth and death chain in 2D

4.1 The model of population dynamics in 2D as a discrete time finite Markov chain

We consider a model of population dynamics of 2 species which are in competition. Therefore, we intend to study asymptotic behavior of the population sizes before absorption. (After a long enough time, absorption is certain.) In other words, we would like to see which population will survive or whether they can coexist conditioned on non-extinction.

Let's consider a discrete-time, finite birth and death chain (X_n, Y_n) in $[0, .., n]^2$ with the following transition:
1) $(i, j) \to (i, j)$ with probability r_{ij},
$(i, j) \to (i+1, j)$ w.p. $p_{ij}^{(1)}$, $(i, j) \to (i, j+1)$ w.p. $p_{ij}^{(2)}$
$(i, j) \to (i-1, j)$ w.p. $q_{ij}^{(1)}$, $(i, j) \to (i-1, j)$ w.p. $q_{ij}^{(2)}$.
2) (0,0) is only absorbing state.
3) When either coordinate hits zero there is no birth and death probability for that component and when either coordinate reaches state "n" then there is no birth probability.

Let's denote for $i, j \geqslant 1$, $S_1 = \{(0, i) | n \leqslant i \leqslant 1\}$, $S_2 = \{(i, 0) | n \leqslant i \leqslant 1\}$ and $S_3 = \{(i, j) | n \leqslant i, j \leqslant 1\}$.

When we order our states appropriately as $(0,0), (0, i) \in S_1, (i, 0) \in S_2, (i, j) \in S_3$ $\forall 1 \leqslant i, j \leqslant n$, we will have the following stochastic matrix:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & A_1 & 0 & 0 \\ * & 0 & A_2 & 0 \\ 0 & B & C & A_3 \end{bmatrix}$$

In stochastic non-irreducible matrix P, there are at most 5 nonzero entries in any row. $T = \{(0,1), (1,0), ..\}$ transient set has 3 irreducible classes S_1, S_2, S_3. Moreover, since $r_{ij} > 0$ for any state $s \in N^*$, by lemma2.1.5 $\exists N$ such that $\forall n > N, P_{i_1 j_1 \to i_2 j_2}(n) > 0$ for each stochastic matrixes corresponding to sets S_1, S_2 and S_3. Therefore the matrixes A_1, A_2 and A_3 are aperiodic and irreducible.

Non-irreducible situations were already studied e.g. the paper "On quasi-stationary distributions in absorbing discrete time finite markov chains" by N.Darroch and E.Seneta. This particular case has not been studied in the literatures as far as we know.

4.2 Convergence to Quasi-stationary distributions

We want to find the asymptotic behavior of the conditional distribution:

$$\lim_{n\to\infty} P_{(a,b)}((X_n, Y_n) = (i, j) | (X_n, Y_n) \neq 0).$$

We have several cases:
1) The process starts from S_1 set (S_2 set) and $(i, j) \in S_1$. ($i, j \in S_2$)
2) The process starts from S_3 set and $(i, j) \in S_3$.
3) The process starts from S_3 set and $(i, j) \in S_2$. ($(i, j) \in S_1$)

For this and next section, we denote $\lambda_1, \lambda_2, \lambda_3$ as maximum eigenvalues corresponding to matrixes A_1, A_2, A_3, $\omega^{(1)}, \omega^{(2)}, \omega^{(3)}$ as right positive eigenvectors corresponding to $\lambda_1, \lambda_2, \lambda_3$ eigenvalues, $\nu'^{(i)}\omega^{(i)} = 1$ $\nu^{(1)}, \nu^{(2)}, \nu^{(3)}$ as left positive normalized eigenvectors corresponding to $\lambda_1, \lambda_2, \lambda_3$ eigenvalues.

$\lambda = \max\{\lambda_1, \lambda_2, \lambda_3\}$, $\mu = \max\{|\mu_1|, |\mu_2|, |\mu_3|\}$ where μ_i are second biggest eigenvalues (by modulus) of matrix A_i. $m = \max\{m_1, m_2, m_3\}$ where $m_i = m_i^* - 1, m_i^*$ are multiplicity of eigenvalue μ_i.

$e = (1, 1, 1, ...), f_{(a,b)} = (0, 0.., 0, 1, 0, 0..)$ ((a, b)th component is 1 and all other elements are zero). Also we suppose that dimension of both e and f vectors are flexible for matrix computation.

$$P^n = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & A_1^n & 0 & 0 \\ * & 0 & A_2^n & 0 \\ * & B_n & C_n & A_3^n \end{bmatrix}$$

where $B_n = \sum\limits_{k=0}^{n-1} A_3^{n-k-1} B A_1^k$ and $C_n = \sum\limits_{k=0}^{n-1} A_3^{n-k-1} C A_2^k$.

($C_n = C^n = CA_3^{n-1} + F_{n-1}B = ... = \sum\limits_{k=0}^{n-1} A_3^{n-k-1} C A_2^k$. In a similar way we find B_n)

1) For this case, the corresponding substochastic transition matrix is irreducible and aperiodic, therefore as we have seen in the part 1, asymptotic behavior will converge to the left normalized positive eigenvector of matrix A_1.
$(a, b), (i, j) \in S_1(S_2)$, $\lim\limits_{n\to\infty} P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = v_{ij}^{(1)}$ $(v_{ij}^{(2)})$. where $v_{ij}^{(1)}$ $(v_{ij}^{(2)})$is the eigenvector corresponding to maximal eigenvalue of matrix $A_1(A_2)$.

2) $(a, b), (i, j) \in S_3$

$$P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = \frac{P_{(a,b)(i,j)}(n)}{\sum\limits_{(p,q)\in S_1\cup S_2\cup S_3} P_{(a,b)(p,q)}(n)} =$$

$$= \frac{A_3^n((a, b)(i, j))}{f_{(a,b)}A_3^n e' + \sum\limits_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} C A_2^k e' + \sum\limits_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} B A_1^k e'}$$

By using Perron-Frobenius theorem:

$$P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = \frac{\omega_{(a,b)}^{(3)} v_{(i,j)}^{(3)} + o(n^k(\frac{\mu}{\lambda_3})^n)}{\sum\limits_{(p,q)\in S_3} \omega_{(a,b)}^{(3)} v_{(p,q)}^{(3)} + o(n^k(\frac{\mu}{\lambda_3})^n) + R}$$

where $R = \sum\limits_{(p_4,q_4)} \sum\limits_{(p_5,q_5)} \sum\limits_{(p_1,q_1)} \frac{1}{\lambda_3} \omega_{(a,b)}^{(3)} v_{(p_1,q_1)}^{(3)} B_{(p_1,q_1)(p_2,q_2)} \omega_{(p_2,q_2)}^{(1)} v_{(p_3,q_3)}^{(1)} \sum_{k=0}^{n-1} (\frac{\lambda_1}{\lambda_3})^k +$
$\sum\limits_{(p_3,q_3)} \sum\limits_{(p_2,q_2)} \sum\limits_{(p_1,q_1)} \frac{1}{\lambda_3} \omega_{(a,b)}^{(3)} v_{(p_1,q_1)}^{(3)} C_{(p_1,q_1)(p_2,q_2)} \omega_{(p_2,q_2)}^{(2)} v_{(p_3,q_3)}^{(2)} \sum_{k=0}^{n-1} (\frac{\lambda_2}{\lambda_3})^k$

$(p_3, q_3), (p_2, q_2) \in S_2, (p_1, q_1) \in S_3, (p_4, q_4), (p_5, q_5) \in S_1.$

If $\lambda_3 \neq \lambda$ then numerator is finite, so at least one of the component of R goes to infinity. Therefore,

$$\lim_{n\to\infty} P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = 0$$

If $\lambda_1 = \lambda_3 = \lambda$ or $\lambda_2 = \lambda_3 = \lambda$ then the above statement also holds, because of the term $\sum_{k=0}^{n-1} (\frac{\lambda_i}{\lambda_3})^k \to \infty$.

If $\lambda_3 = \lambda, \lambda_3 > \lambda_1, \lambda_3 > \lambda_2$ then

$$\sum_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} B A_1^k e'/\lambda_3^n = \frac{1}{\lambda}\sum_{k=0}^{n-1} f_{(a,b)}\omega^{(3)} v^{(3)'} B(\frac{A_1}{\lambda_3})^k e' + \frac{1}{\lambda}\sum_{k=0}^{n-1} f_{(a,b)} o((n-k-1)^m (\frac{\mu}{\lambda})^{(n-k-1)}) B(\frac{A_1}{\lambda})^k e'$$

Since $\lambda_1 < \lambda, \mu < \lambda$, $\frac{1}{\lambda}\sum_{k=0}^{n-1} f_{(a,b)} o(n)^k (\frac{\mu}{\lambda})^{n-k-1} B(\frac{\lambda_1}{\lambda})^k \omega_{(a,b)}^{(1)} v^{(1)'} e' \to 0$.

Also the same result holds for $\sum_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} C A_2^k e'/\lambda^n$.

Since, if $\theta > \lambda$ (θ is the maximal eigenvalue of nonnegative aperiodic and irreducible matrix M), $\frac{1}{\lambda}\sum_{k=0}^{n} \frac{M^k}{\lambda^k} \to (\lambda I - M)^{-1}$. Therefore,

$$\lim_{n\to\infty} P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = \frac{v_{(ij)}^{(3)}}{1 + v^{(3)'} B(\lambda_3 I - A_1)^{-1} e' + v^{(3)'} C(\lambda_3 I - A_2)^{-1} e'}.$$

3) $(a, b) \in S_3, (i, j) \in S_2$

$$B = P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = \frac{\sum_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} C A_2^k f_{(i,j)}}{\sum_{(p,q)\in S_1\cup S_2\cup S_3} P_{(a,b)(p,q)}(n)}$$

$$= \frac{\sum_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} C A_2^k f'_{(i,j)}}{f_{(a,b)} A_3^n e' + \sum_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} C A_2^k e' + \sum_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} B A_1^k e'}$$

If $\lambda_3 = \lambda, \lambda_3 > \lambda_1, \lambda_2$ then as we did in previous calculation,

$$B \to \frac{v^{(3)'} C(\lambda_3 I - A_2)^{-1} f'_{(i,j)}}{1 + v^{(3)'} B(\lambda_3 I - A_1)^{-1} e' + v^{(3)'} C(\lambda_3 I - A_2)^{-1} e'}.$$

If $\lambda_2 = \lambda, \lambda_2 > \lambda_1, \lambda_3$ then $\frac{f_{(a,b)} A_3^n e'}{\lambda^n} \to 0$, $\frac{\sum_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} C A_2^k e'}{\lambda_2^n} \to \frac{1}{\lambda_2} f_{(a,b)}(\lambda_2 I - A_3)^{(-1)} C\omega^{(2)} v^{(2)'} e'$, Using the Perron Frobenius theorem, for the following component the leading term converges to zero: $\frac{\sum_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} B A_1^k e'}{\lambda_2^n} \approx \sum_{k=0}^{n-1} \frac{\lambda_3^{n-k-1}\lambda_1^k}{\lambda_2^n} \omega^{(3)}_{(a,b)} v'^{(3)} C\omega^{(1)} v^{(1)'} \to 0$. Therefore in this case,

$$\lim_{n\to\infty} P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = \frac{f_{(a,b)}(\lambda_2 I - A_3)^{(-1)} C\omega^{(2)} v^{(2)}_{(i,j)}}{f_{(a,b)}(\lambda_2 I - A_3)^{(-1)} C\omega^{(2)} v^{(2)'} e} = v^{(2)}_{(i,j)}.$$

If $\lambda_1 = \lambda, \lambda_1 > \lambda_2, \lambda_3$ then the term in numerator goes to zero and denominator is nonzero. Therefore $\lim_{n\to\infty} P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = 0$.

If $\lambda_3 < \lambda_1 = \lambda_2 = \lambda$ then $\frac{f_{(a,b)} A_3^n e'}{\lambda^n} \to 0$, $\frac{\sum_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} C A_2^k e'}{\lambda^n} \to \frac{1}{\lambda} f_{(a,b)}(\lambda I - A_3)^{(-1)} C\omega^{(2)}$, $\frac{\sum_{k=0}^{n-1} f_{(a,b)} A_3^{n-k-1} B A_1^k e'}{\lambda^n} \to \frac{1}{\lambda} f_{(a,b)}(\lambda I - A_3)^{(-1)} B\omega^{(1)}$. Therefore,

$$\lim_{n\to\infty} P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = \frac{f_{(ab)}(\lambda I - A_3)^{-1} C\omega^{(2)} v^{(2)}_{(i,j)}}{f'_{(ab)}(\lambda I - A_3)^{-1} C\omega^{(2)} + f_{(ab)}(\lambda I - A_3)^{-1} B\omega^{(1)}}.$$

If $\lambda_3 = \lambda_1 = \lambda_2 = \lambda$ then when we use Perron-Frobenius theorem, the leading term in numerator and denominator is $o(n\lambda^n)$. Now we prove that other terms are negligible. To have easier notation, we suppose n is odd. (When n is even, it is the same.) $R_1 = \sum_{k=0}^{(n-1)/2} f_{(a,b)}(\lambda_3^{n-k-1} \omega^{(3)} v'^{(3)} + o((n-k-1)^m \mu^{n-k-1})) C A_2^k e' + \sum_{k=(n-1)/2+1}^{n-1} f_{(a,b)}(A_3^{n-k-1} C(\lambda_2^k \omega^{(2)} v'^{(2)} + o(k^m \mu^k)) e'$,

When we change the index: $\sum_{k=(n-1)/2+1}^{n-1} f_{(a,b)}(A_3^{n-k-1} C(\lambda_2^k \omega^{(2)} v'^{(2)} + o(k^m \mu^k)) e' = \sum_{k=0}^{(n-1)/2-1} f_{(a,b)} A_3^k C(\lambda^{n-k+1} \omega^{(2)} v'^{(2)} + o((n-k+1)^m \mu^{n-k+1})) C A_2^k e'$.

Therefore, it is sufficient to see that $o(\mu^n \sum_{k=0}^{(n-1)/2} (\frac{\lambda}{\mu})^k)$ is negligible. $o(\mu^n \sum_{k=0}^{(n-1)/2} (\frac{\lambda}{\mu})^k) \leqslant o(\mu^n (\frac{\lambda}{\mu})^{(n+1/2)}) = \circ(\lambda^n)$.

Therefore, $R_1 \sim \lambda^{n-1} n[f_{(a,b)} \omega^{(3)} v'^{(3)} C\omega^{(2)} v'^{(2)} e']$. After simplifying the fraction we have the following:

$$\lim_{n\to\infty} P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = \frac{v'^{(3)} B\omega^{(1)} v'^{(1)} f_{(i,j)}}{v'^{(3)}(C\omega^{(2)} + B\omega^{(1)})}.$$

In the case when process ends in state from S_1, we have similar results. We can see that not likely to irreducible case, the asymptotic behavior is dependent on initial state. Furthermore, we studied asymptotic behavior of our model through maximal eigenvalues of sub matrixes. If the maximal eigenvalue of A_3 is not the biggest one, when we start from state, $s \in S_3$, then eventually it will converge to state in S_1 or S_2. If it is, then there is a probability to stay in set S_3(co-existence).

4.3 The model of population dynamics in 2D as continuous-time finite Markov chains

We study the analogue model of birth and death chain in 2D, where infinitesimal birth and death rates are given by the following:

1) $(i, j) \to (i, j)$ with rate r_{ij},
$(i, j) \to (i+1, j)$ with rate $p^{(1)}_{ij}$, $(i, j) \to (i, j+1)$ with rate $p^{(2)}_{ij}$
$(i, j) \to (i-1, j)$ with rate $q^{(1)}_{ij}$, $(i, j) \to (i-1, j)$ with rate $q^{(2)}_{ij}$.
2) (0,0) is only absorbing state.
3) When either coordinate hits zero there is no birth and death rate for that component and when either coordinate reaches state "n" then there is no birth rate.

As we mentioned before we are using the notation in 4.2.
It will have the following infinitesimal generator if we order them as we did in 4.1.

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 \\ * & A_1 & 0 & 0 \\ * & 0 & A_2 & 0 \\ 0 & B' & C' & A_3 \end{bmatrix}$$

We have seen from section 3.3 , $P(t) = e^{Qt} = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^k$ (exponential matrix). Let's denote $Q_i(t) = e^{A_i t}, i = 1, 2, 3$, then

$$P(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & Q_1(t) & 0 & 0 \\ * & 0 & Q_2(t) & 0 \\ * & S_1(t) & S_2(t) & Q_3(t) \end{bmatrix}$$

where $S_1(t) = \sum_{m=1}^{\infty} (\sum_{k=0}^{m-1} A_3^{m-k-1} B' A_1^k \frac{t^m}{m!})$, $S_2(t) = \sum_{m=1}^{\infty} (\sum_{k=0}^{m-1} A_3^{m-k-1} C' A_2^k \frac{t^m}{m!})$.

1) $(a, b), (i, j) \in S_1$, (S_2). The states in S_1 communicates with itself, therefore as we did in Proposition 3.3.2:

$$\lim_{n\to\infty} P_{(a,b)}((X_n, Y_n) = (i, j)|(X_n, Y_n) \neq 0) = \frac{f_{(a,b)} Q_1(t) f'_{(i,j)}}{f_{(a,b)} Q_1(t) e'} \to \nu^{(1)}(i, j).$$

Similarly, when $(a, b), (i, j) \in S_2$ then it converges to $(\nu^{(1)}(i, j))$.

Let's suppose process stayed in S_3 for t-x time, next shifted to S_2 or S_1 remaining there x time. $t \in R^+$

2) $(a,b),(i,j) \in S_3$.

$$\lim_{n\to\infty} P_{(a,b)}((X_t,Y_t)=(i,j)|(X_t,Y_t)\neq 0) = \frac{f_{(a,b)}Q_3(t)f'_{(i,j)}}{\sum\limits_{(c,d)\in S_1\cup S_2\cup S_3} P_{(a,b)(i,j)}(t)}$$

Then the denominator of the above fraction is equal to

$R = f_{(a,b)}Q_3(t)e' + f_{(a,b)}(\int_0^t Q_3(t-x)BQ_1(x)dxe' + \int_0^t Q_3(t-x)CQ_2(x)dx)e'$

where B and C matrixes given by following tranisiton probability: (we assume that transition occurs at time x)

$(i,j)\to(i+1,j)$ w.p. $p^{(1)}_{ij}/(1-r_{ij})$, $(i,j)\to(i,j+1)$ w.p. $p^{(2)}_{ij}/(1-r_{ij})$

$(i,j)\to(i-1,j)$ w.p. $q^{(1)}_{ij}/(1-r_{ij})$, $(i,j)\to(i-1,j)$ w.p. $q^{(2)}_{ij}/(1-r_{ij})$. If $\lambda_3 \neq \lambda$ then $\lim\limits_{n\to\infty} P_{(a,b)}((X_t,Y_t)=(i,j)|(X_t,Y_t)\neq 0)=0$. Because when we use corollary of Perron Frobenius theorem for every matrices in fraction, the numerator goes to constant and the dominant term of denominator goes to infinity. $\int_0^t e^{(t-x)\lambda_3}e^{x\lambda_1}dx/e^{t\lambda_3} = \frac{e^{t\lambda_3}-e^{t\lambda_1}}{\lambda_3-\lambda_1}/e^{t\lambda_3} \to \infty$.

Also, if $\lambda_3=\lambda=\lambda_i, i=1,2$ then above statement holds.

If $\lambda_3 > \lambda_2,\lambda_1$ then,

$f_{(a,b)}\int_0^t Q_3(t-x)BQ_1(x)dxe' = f_{(a,b)}\int_0^t (e^{(t-x)\lambda_3}\omega^{(3)}v'^{(3)} + o(e^{(t-x)\mu}))BQ_1(x)dxe'$,

and using $\int_0^t e^{sA}ds = [A^{-1}e^{sA}]_0^t = A^{-1}(e^{tA}-I)$,

we can see that second term is negligible.

$\int_0^t o(e^{(t-x)\mu})e^{A_1x}dx = e^{t\mu}(e^{(A_1-I\mu)t}-I)(A_1-I\mu)^{-1} \Rightarrow e^{t\mu}(e^{(A_1-I\mu)t}-I)(A_1-I\mu)^{-1}/e^{t\lambda_3} \to 0$.

Therefore $f_{(a,b)}\int_0^t Q_3(t-x)BQ_1(x)dxe'/e^{t\lambda_3} \to f_{(a,b)}\omega^{(3)}v'^{(3)}B(A_1-I\lambda_3)^{-1}e'$, and similarly $f_{(a,b)}\int_0^t Q_3(t-x)CQ_2(x)dxe'/e^{t\lambda_3} \to f_{(a,b)}\omega^{(2)}v'^{(2)}C(A_2-I\lambda_3)^{-1}e'$.

$$\lim_{n\to\infty} P = \frac{\omega^{(3)}(a,b)v^{(3)}(i,j)}{\omega^{(3)}(a,b)v'^{(3)}B(A_1-I\lambda_3)^{-1}e' + \omega^{(2)}(a,b)v'^{(2)}C(A_2-I\lambda_3)^{-1}e' + \omega^{(3)}(a,b)}.$$

3) $(a,b)\in S_3,(i,j)\in S_2$.

$$\lim_{n\to\infty} P_{(a,b)}((X_t,Y_t)=(i,j)|(X_t,Y_t)\neq 0) = \frac{f_{(a,b)}\int_0^t Q_3(t-x)CQ_2(x)dxf'_{(i,j)}}{\sum\limits_{(c,d)\in S_1\cup S_2\cup S_3} P_{(a,b)(i,j)}(t)}$$

If $\lambda_3 = \lambda, \lambda_3 > \lambda_2, \lambda_1$ then, as we did in second case, we will have the following result:

$$\lim_{n\to\infty} P = \frac{\omega^{(2)}(a,b)v'^{(2)}C(A_2 - I\lambda_3)^{-1}f'(i,j)}{\omega^{(3)}(a,b)v'^{(3)}B(A_1 - I\lambda_3)^{-1}e' + \omega^{(2)}(a,b)v'^{(2)}C(A_2 - I\lambda_3)^{-1}e' + \omega^{(3)}(a,b)}.$$

If $\lambda_2 = \lambda, \lambda_2 > \lambda_3, \lambda_1$ then,

$$\lim_{n\to\infty} P = \frac{v'^{(3)}C(A_3 - I\lambda_2)^{-1}f'(i,j)}{v'^{(3)}C(A_3 - I\lambda_2)^{-1}e'}.$$

If $\lambda_1 = \lambda, \lambda_1 > \lambda_3, \lambda_2$ then, the numerator will go to zero and denominator is constant. $\lim_{n\to\infty} P = 0$

If $\lambda_1 = \lambda_2 = \lambda_3$, then the second order term are all negligible.

$$\lim_{n\to\infty} P = \frac{v'^{(3)}C\omega^{(2)}v'^{(2)}f'(i,j)}{v'^{(3)}C\omega^{(2)}v'^{(2)}e' + v'^{(3)}C\omega^{(1)}v'^{(1)}f'(i,j)}.$$

After doing similar calculation: If $\lambda = \lambda_2 = \lambda_3 > \lambda_1$ then,

$$\lim_{n\to\infty} P = \frac{v'^{(3)}C\omega^{(2)}v'^{(2)}f'(i,j)}{v'^{(3)}C\omega^{(2)}v'^{(2)}e'}.$$

If $\lambda = \lambda_1 = \lambda_2 > \lambda_3$ then,

$$\lim_{n\to\infty} P = \frac{v'^{(3)}C(A_3 - I\lambda_2)^{-1}f'(i,j)}{v'^{(3)}C(A_3 - I\lambda_2)^{-1}e' + v'^{(3)}B(A_3 - I\lambda_2)^{-1}e'}.$$

5. Particular cases where eigenvalues can be found

We have seen so far that finding eigenvalues and eigenvectors of non-negative matrixes is one way to study the model of population dynamics. In this part we will see particular cases where we can find the form of eigenvalues and eigenvectors.

Neutral birth and death model

This is the case where each individual are exchangeable. It means each individual has the same birth and death probability.

We have the states $(i,j) = \{0 \leqslant i+j \leqslant N\}$ and transition probabilities: $p^1_{ij} = \frac{i}{i+j}\lambda(i+j), p^2_{ij} = \frac{j}{i+j}\lambda(i+j), q^1_{ij} = \frac{i}{i+j}\mu(i+j), q^2_{ij} = \frac{j}{i+j}\mu(i+j)$ and $r_{(i,j)} = 0$

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & A & 0 & 0 \\ * & 0 & A & 0 \\ 0 & B & C & A_1 \end{bmatrix}$$

Proposition5.1. P has the right eigenvectors of the form $v_{i+j}, iv_{i+j}, jv_{i+j}, ijv_{i+j}$.
$\triangleleft$ It follows directly from Proposition 5.2. $\triangleright$

Proposition5.2. If

$$\left\{ \begin{array}{l} iP_d(i+1,j) + jP_d(i,j+1) = (i+j+d)P_d(i,j) \\ iP_d(i-1,j) + jP_d(i,j-1) = (i+j-d)P_d(i,j) \end{array} \right. (*)_d$$

$\deg(P_d) = d$ then P has the eigenvectors of the form $P_d(i,j)v_{i+j}$.
$\triangleleft$

$$A = \begin{bmatrix} 0 & \lambda_{(1,0)} & 0 & 0 \\ \mu_{(2,0)} & 0 & \lambda_{(2,0)} & 0 \\ 0 & \mu_{(3,0)} & 0 & \lambda_{(3,0)} \\ 0 & 0 & \ddots & \ddots \end{bmatrix}$$

We can take the matrix A_3 as the linear operator: $A_1 : \mathbb{R}^d \to \mathbb{R}^d$ where $d = N(N+1)/2$, $u \in \mathbb{R}^d$, $u = (u_{ij}, (ij) \in S)$, for $(i,j) \neq 0$.

$(Pu)_{ij} = \frac{i}{i+j}\lambda_{i+j}u_{i+1,j} + \frac{j}{i+j}\lambda_{i+j}u_{i,j+1} + \frac{i}{i+j}\mu_{i+j}u_{i-1,j} + \frac{j}{i+j}\mu_{i+j}u_{i,j-1}$, Let's write $i+j=k$ then $i\lambda_k u_{i+1,j} + j\lambda_k u_{i,j+1} + i\mu_k u_{i-1,j} + j\mu_k u_{i,j-1} = \theta k u_{(i,j)}$
Therefore it reduces to recurrence relation: $\lambda_k v_{k+1} + \mu_k v_{k-1} = \theta k v_k$ for $k \geqslant 1$ where θ is the maximal eigenvalue of A_3. When $(i,j) = (0,0)$ then $u_{(0,0)} = \theta u_{(0,0)}$.

Now we show that $P_d(i,j)v_{(i+j)}$ satisfies the recurrence relation(the eigenvector). $P(P_d v_{i+j}) = \frac{i}{i+j}\lambda_{i+j}P_d(i+1,j)v_{i+j+1} + \frac{j}{i+j}\lambda_{i+j}P_d(i,j+1)v_{i+j+1} + \frac{i}{i+j}\mu_{i+j}P_d(i-1,j)v_{i+j-1} + \frac{j}{i+j}\mu_{i+j}P_d(i,j-1)v_{i+j-1}$. Here we substitute $(*)_d$ equations and we will get $P(P_d v_{i+j}) = (\frac{i+j+d}{i+j}\lambda_{i+j}v_{i+j+1} + \frac{i+j-d}{i+j}\mu_{i+j}v_{i+j-1})P_d(i,j) = \theta P_d(i,j)v_{i+j}$.

Actually, d can be any number to be an eigenvector. However to have solution in $(*)_d$ we need d to be a degree of polynomial P_d. To see this, let's take the coefficient of maximal degree term of polynomial. $iP_d(i+1,j) + jP_d(i,j+1) = (i+j+d)P_d(i,j) \to \triangleright$.

We can find other eigenvectors using Proposition 5.2. e.g. $ij(i-j), ij(j-i)$. Furthermore $\deg P(i,j) = 4$, $P_4 = i^3 j + ij^3 - 3i^2 j^2 + ij$. The set consisting of these polynomials are vector space.

Conjecture $\forall d \geqslant 2, \exists$ a 1-dimensional vector space of $(*)_d$

Theorem 5.1 If conjecture is true, then all the eigenvectors of P are of the form $P_d(i,j)v_{i+j}$, where v_k solves $\frac{i}{i+j}\lambda_{i+j}P_d(i+1,j)v_{i+j+1} + \frac{j}{i+j}\lambda_{i+j}P_d(i,j+1)v_{i+j+1} + \frac{i}{i+j}\mu_{i+j}P_d(i-1,j)v_{i+j-1} + \frac{j}{i+j}\mu_{i+j}P_d(i,j-1)v_{i+j-1} = \theta P_d(i,j)v_k$.

Another example. From the book "Birth and death processes models with applications" by P.R.Parthasarathy, R.B. Lenin, page 190, If $\eta_k = \xi + 2(k-1)^2 c, \beta_k(N-k)(b-(k-1)c)$ and $v_k = (N+k-1)(b+kc)$ then,

$$\begin{bmatrix} \eta_1 & \beta_1 & 0 & 0 \\ \nu_1 & \eta_2 & \beta_2 & 0 \\ 0 & \ddots & .. & .. \\ .. & .. & \nu_n & \eta_n \end{bmatrix} = \prod_{j=1}^{N} \{X + 2(j-1)(Y + [N-(j-2)]Z)\}$$

where $X = \xi - 2(N-1)b$, $Y = 2b - 3c$ and $Z = 4c$.

We take $\eta_k = 1 + 2/N^2 - 2/N$, $\beta_k = (N-k)/N^2, \nu_k = (N+k-1)/N^2$, $b = 1/N^2, c = 0$. $\forall k \geqslant 2$, $\nu_{k-1} + \eta + \beta_k = 1$. The matrix is substochastic.

6. Simulation of Quasi-stationary distribution

Another approach to study the model is to do simulation by particle method.

6.1. Simulation of Markov chain

By the law of large number, $P_i(X_n = k) \approx \frac{1}{N} \sum_{j=1}^{N} 1_{X_n^j = k} = \overline{X_N}$. Moreover, the speed , $\frac{1}{\sqrt{N}}$, of convergence is given by the central limit theorem.

With probability more than 95%, the true value of $P_i(X_n = k)$ belongs to the (random interval) $[\overline{X_N} - 1,96\frac{\sigma}{\sqrt{N}}, \overline{X_N} + 1,96\frac{\sigma}{\sqrt{N}}] \approx [\overline{X_N} - 1,96\frac{\sqrt{\overline{X_N}(1-\overline{X_N})}}{\sqrt{N}}, \overline{X_N} + 1,96\frac{\sqrt{\overline{X_N}(1-\overline{X_N})}}{\sqrt{N}}]$ since $\sigma^2 = P(X_n = k)(1 - P(x_n = k))$.

By using uniform random variable, we choose next state and in this way we simulate Markov chain. Let's suppose an unit length which is divided into transition probabilities. In unit length, throw a point randomly and we choose next step where it hit.

Example 6.1 We have the following transition probabilities:

$i \to i+1$ with probability $\frac{b}{b+a+c(i-1)}$

$i \to i$ with probability $\frac{a}{b+a+c(i-1)}$

$i \to i-1$ with probability $\frac{c(i-1)}{b+a+c(i-1)}$.

Then the invariant distribution must satisfy

$\pi_k = \pi_{k-1}\frac{b}{b+a+c(k-2)} + \pi_k\frac{a}{b+a+c(k-1)} + \pi_{k+1}\frac{kc}{b+a+ck}$.

We can see that the invariant distribution has the following form $\pi_k = D(\frac{b}{c})^k \frac{a+b+c(k-1)}{(k-1)!}$.

Since the invariant distribution is unique for our case, after normalizing we find $\pi_k = \frac{c}{b(2b+a)} e^{-\frac{b}{c}} (\frac{b}{c})^k \frac{a+b+c(k-1)}{(k-1)!}$.

After implementing it, we had the following result:(for a particular state)

m=5000 ; //number of particles

n=300 ; //steps

pi2 = 0.033 numerical solution

pi1 = 0.0306566 analytical solution

err = - 0.0023434

m=20000 ; //number of particles

pi2 = 0.0309
pi1 = 0.0306566
err = - 0.0002434

6.2. Particle method Let's denote the law of the process $q_n = \mathbb{L}(X_n | X_n \neq 0)$. Suppose $q_{n+1} = K_{q_n} q_n$. From paper "Particle Motions in Absorbing Medium with hard and soft obstacles" by P.Del Moral and A. Doucet, the idea is to approximate $q_n(k)$ by $\frac{1}{N} \sum_{i=1} N \delta_{X_n^i}$ with a dynamics for $(X_n^1, ...X_n^N)$ given by $K_{\frac{1}{N} \sum_{i=1}^N \delta_{X_n^i}}$.
Algorithm comes from rewriting of $K_{\frac{1}{N} \sum_{i=1}^N \delta_{X_n^i}}$ as a 2 step Markov chain,
step1 : mutation
step2 : selection.

This gives in our case the following algorithm:
1) At the same time we run n particles, when it becomes extinct we randomly replace it with another existing particle.
2) At the same time we run n particles and each time every particle is chosen randomly from the previous step extant particles.

The simulation of quasi stationary distribution for the semi-infinite simple random walk with absorbtion:(10000 particles)
pi2 = 0.0134
pi1 = 0.0141707
err = 0.0007707

References

[1] J.N.Norris, Markov chains.

[2] E.Seneta, Non negative matrices and Markov chains.

[3] Pierre Bremaud, Markov chains Gibbs fields, monte carlo simulation and queues, Springer.

[4] P.R.Parthasarathy and R.B.Lenin, Birth and Death processes models with applications, American Sciences press, INC.

[5] J.N. Darroch and E.Seneta, "On quasi stationary distributions in absorbing discrete time finite markov chains", J.Appl.Prob.2 (1965).

[6] J.N. Darroch and E.Seneta, "On quasi stationary distributions in absorbing continuous time finite markov chains", J.Appl.Prob.4(1967).

[7] E.Seneta and D.Vere-Jones, "On quasi stationary distributions in discrete time markov chains with a denumerable infinity of states", J.Appl.Prob.3(1966).

[8] P.Del Moral and A.Doucet, "Particle motions in Absorbing Medium with Hard and Soft obstacles".

[9] M.Kijima and E.Seneta, "Some results for quasi stationary distributions of birth and death processes", J.appl.Prob.28(1991).

[10] D.C. Flashpohler, "Quasi stationary distributions for absorbing continuous time denumerable markov chains", (1973)

[11] N. Champagnat and A.Lambert, "Evolution of discrete populations and the canonical diffusion of adaptive dynamics", Ann.Appl.Prob.17, (2007)