

The Moran Process on a Lattice with Spatial Dependence

Isaac Ravi Brenner

May 2020 - PHYS 580 Biological Physics Final Project

1 Introduction

The Moran Process is a model of finite population biology that can be used to model natural selection, genetic drift, and spontaneous mutation. In each time step, one random individual is chosen for reproduction, and another is chosen for death or replacement, causing the total population size to remain constant. In this paper, I will describe the basics of the typical Moran Process, which uses a well mixed population. Then, I will explore the application of this same process to a 1D and then a 2D lattice, where there is spatial dependence on reproduction and replacement. I start with simpler models and build up to more complicated setups, using python to create simulations of some different possible outcomes. Most of the work of this project was in actually encoding the Moran process on a lattice. Use [this](#) link for a github of the code.

2 An intro to the Moran Process

2.1 Neutral Drift

The Moran process with neutral drift starts with N individuals, $N_A = i$ of type A and $N_b = N - i$ of type B. In neutral drift, these can be thought to represent 2 neutral alleles of the same gene in a reproducing population. In this paper, **A's are represented with the color red, and B's are represented with the color blue.** A random individual is chosen from the population of N individuals, and that individual reproduces and replaces another randomly chosen individual in the population. It could replace another individual or itself. Therefore, transitions can take place from i to $i + 1$ and from i to $i - 1$, and it's also possible that no transition takes place if an individual replaces another of its same type. The transition probabilities are:

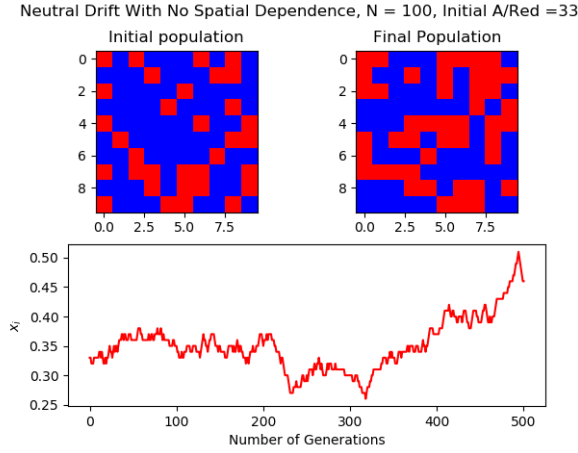
$$P_{i,i+1} = \frac{i}{N} \frac{N-i}{N}$$

$$P_{i,i-1} = \frac{N-i}{N} \frac{i}{N}$$

$$P_{i,i} = 1 - P_{i,i+1} - P_{i,i-1}$$

Another way to think about these probabilities is as the probability that a given type of individual is chosen to reproduce multiplied by the probability that a given type of individual is chosen for replacement. We'll return to this idea later on.

From this, we can calculate the probability of fixation for A, or the probability that it completely takes over the population, which we'll call x_i (The fixation probability for B is always just $1 - x_i$). In the case of neutral drift, $x_i = i/N$. Another way of thinking about this is that when every individual is equally likely to be chosen to reproduce, each individual has a probability of $x_1 = 1/N$ to become the common ancestor of the whole population. Therefore, counting up the number of individuals of each type reveals each one's probability of fixation.



(A note about the figures and graphics in this project: the Moran process is stochastic, and so all of these figures represent random trajectories of these processes. When the process is run multiple times, there will be different results. The figures included are to provide a sense of what might happen given particular parameters. I just chose some possible values and ran the code to come up with the figures. Accompanying gif animations can be found using this [google drive link](#))

2.2 Selection

Now, we will consider a case where A has a selective advantage over B. We'll assign to A fitness f_a , and B fitness $f_b = 1$ for simplicity. This gives some new probabilities of being chosen:

$$P_{i,i+1} = \frac{f_a^i}{f_a^i + f_b(N-i)} \frac{i}{N}$$

$$P_{i,i-1} = \frac{f_b(N-i)}{f_a i + f_b(N-i)} \frac{N-i}{N}$$

$$P_{i,i} = 1 - P_{i,i+1} - P_{i,i-1}$$

Put simply, each type has a probability of being chosen weighted by its relative fitness. In a population of $N = 2$, $i = 1$, and $f_a = 2$, the probability that A is chosen to reproduce is $2/3$, even though it only makes up half of the population.

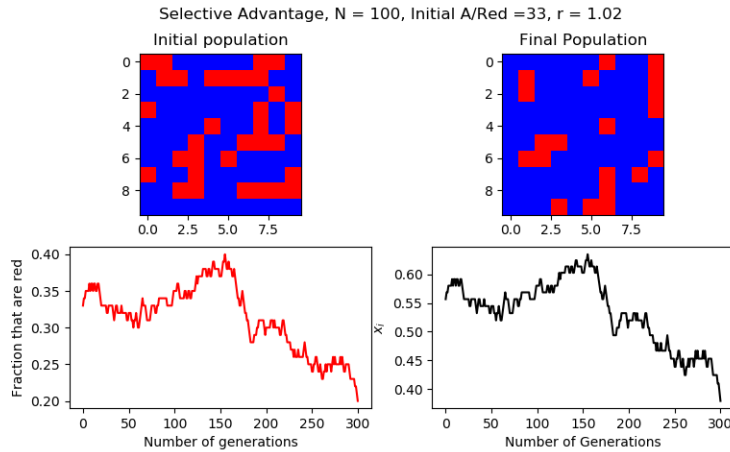
With these probabilities, we can calculate a ratio, $r = P_{i,i+1}/P_{i,i-1}$. In this case, where there is no spatial dependence, $r = \frac{f_a}{f_b}$. This will be useful later on.

When there is a selective advantage, the fixation probability can be written as:

$$x_i = \frac{1-r^{-i}}{1-r^{-N}}$$

Let's test out some limits. $x_0 = 0$ and $x_N = 1$ both check out. $\lim_{r \rightarrow \infty} x_i = 1$ makes sense with our intuition: if r is huge and A has a big selective advantage, then it will almost certainly fix. And $\lim_{r \rightarrow 1} x_i = i/N$ holds as well. This also means that, for a given r , a mutation that is only slightly deleterious but is very common could still have a good chance of fixing in a small population. In a larger population, the individual with a fitness advantage will have a higher chance of fixing for the same r value.

Realistically, most selective advantages are of the form $r = 1 + s$ where $s = .01, .02$ or something similarly small, so that the overall $r = 1.02$, which is only slightly more beneficial than a neutral mutation. This all sets us up to consider spatial dependence.



3 Spatial Dependence - Neutral Drift

3.1 1D Neutral Drift with Spatial Dependence

Eventually, we'll consider reproduction of a population of individuals on an MXM lattice, but first we'll imagine a 1D population, in this case of size $N = 10$ individuals. The probability of being chosen to reproduce is the same as before, i/N for A and $(N - i)/N$ for B. What differs now is that once an individual is chosen for reproduction, it only has 3 options for who to replace— itself, or the individuals on either side of it. Correspondingly, the individuals on the ends have only 2 options, themselves or the individual next to them. Thus, we can calculate the transition probabilities by summing up the probabilities of all the potential transitions that could occur.

We'll call this population $pop(x)$, where $x = 1, 2, \dots, M$. It is quite tedious to write out all of the probabilities, but so we'll just consider that an A is chosen to reproduce (this probability is still i/N). Now the probabilities are as follows:

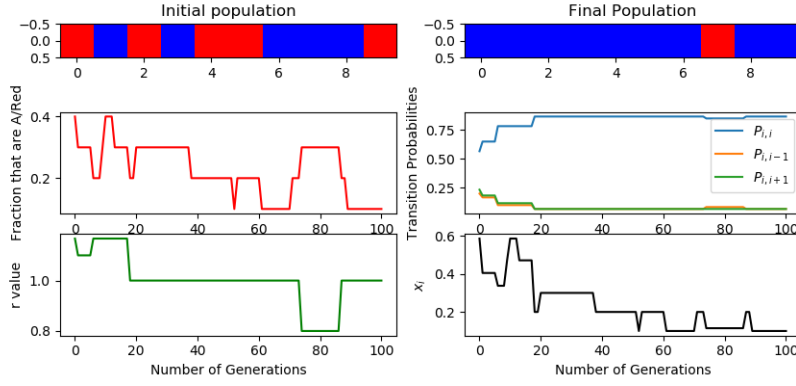
$$\begin{aligned} P_{i,i+1} &= \sum_{k=-1}^1 \frac{pop(x+k)=B}{3} & P_{i,i} &= \sum_{k=1}^1 \frac{pop(x+k)=A}{3} & \text{for } x \neq 1, M \\ P_{i,i+1} &= \sum_{k=0}^1 \frac{pop(x+k)=B}{2} & P_{i,i} &= \sum_{k=0}^1 \frac{pop(x+k)=A}{2} & \text{for } x = 1 \\ P_{i,i+1} &= \sum_{k=-1}^0 \frac{pop(x+k)=B}{2} & P_{i,i} &= \sum_{k=-1}^0 \frac{pop(x+k)=A}{2} & \text{for } x = M \end{aligned}$$

where the numerator in each sum returns a 1 if the condition is satisfied and a 0 if it isn't satisfied. Thus, every individual contributes some fraction to the overall $P_{i,j}$'s. It turns out that under this regime, $P_{i,i}$ is almost always greater than the other transition probabilities, since as a species spreads and begins to dominate it has fewer options of places where it can increase its dominance further.

To calculate the fixation probability now, recall that $r = P_{i,i+1}/P_{i,i-1}$. We can then use this to calculate the fixation probability for the case of neutral drift with 1D spatial dependence. However, due to the nature of this setup, $P_{i,i+1}$ and $P_{i,i-1}$ quickly converge in value, so the overall r ends up being closer to 1, and when $r = 1$, $x_i = i/N$.

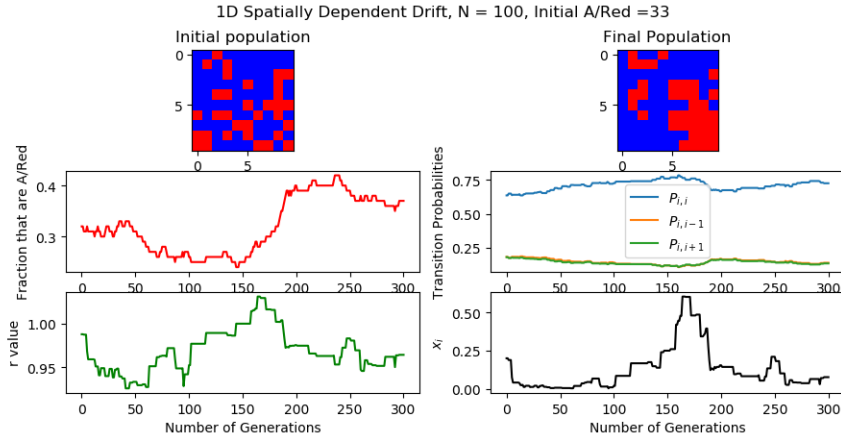
Notice in the graphic that when $r = 1$, x_i = the fraction that are type A, but when $r \neq 1$, $x_i \neq$ the fraction that are type A.

1D Spatially Dependent Drift, $N = 10$, Initial A/Red = 5



3.2 2D Neutral Drift with Spatial Dependence

The 2D case is exactly analogous to the 1D case, but with the probabilities adjusted to account for the additional options for replacement (diagonals are now allowed). Again, in the case of neutral drift, the probability of being chosen to reproduce remains the same, but the transition probabilities are affected by spatial dependence in this way.



4 Spatial Dependence Selective Advantage

4.1 1D Selection with Spatial Dependence

As we consider a selective advantage with spatial dependence, eventually we'll imagine the individuals in our population to be randomly distributed on a square

$M \times M$ lattice. First, we'll consider the 1D case again to establish some basic ideas for dealing with spatial dependence and selection.

In the 1D case, we can imagine 10 individuals randomly in a line. The reproduction radius $r_{reproduction}$ tells us which individuals will be replaced when one reproduces. For simplicity, we'll set $r_{reproduction} = 1$, so if an individual at location $x = 4$ is chosen to reproduce, it can only replace either $x = 3, x = 5$, or itself.

Again, it is tedious to write out all of the probabilities, but some discussion is worthwhile. Now, the probability of an individual being chosen is dependent on the relative fitness of A and B, exactly as we saw in section 2.2. Once an individual is chosen to reproduce, the transition probabilities can be calculated exactly as in section 3.1. Then, the overall transition probabilities for the entire population can be calculated by summing each individual's probability of being chosen multiplied by its transition probabilities.

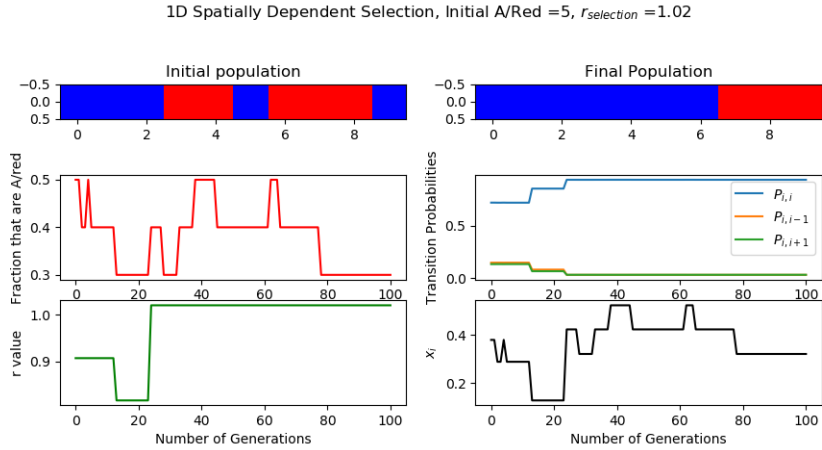
Overall, the transition probabilities for the whole population is the sum of all of the individual probabilities written out as:

$$P_{i,i+1} = P(\text{A is chosen to reproduce}) \times P(\text{A's neighbors that are of type B})$$

$$P_{i,i-1} = P(\text{B is chosen to reproduce}) \times P(\text{B's neighbors that are of type A})$$

$$P_{i,i} = P(\text{A/B chosen to reproduce}) \times P(\text{Choosing self or a neighbor of the same type})$$

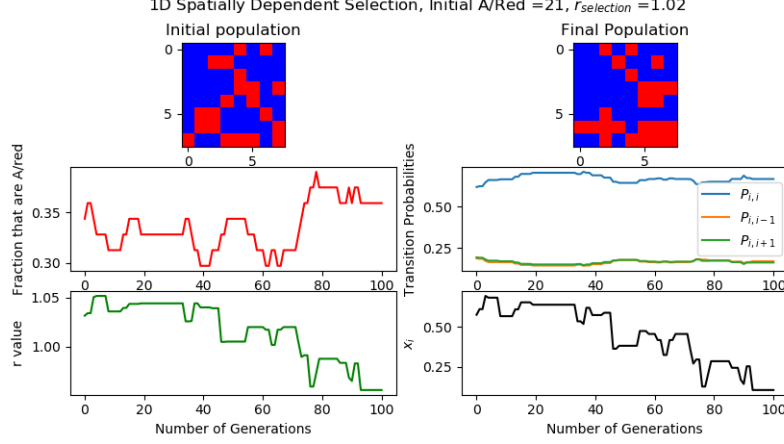
Then, as in the neutral drift case, $r = P_{i,i+1}/P_{i,i-1}$ and $x_i = \frac{1-r^{-i}}{1-r^{-N}}$



4.2 2D Selection with Spatial dependence

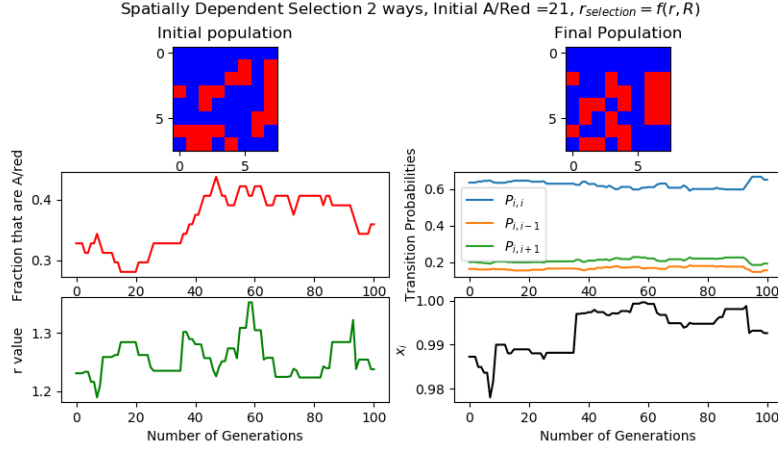
Again here, we'll assume that $r_{reproduction} = 1$, so that each individual can only replace either itself or one of the individuals immediately surrounding it. Individuals in the center have 9 options (8 + themselves), individuals on the edge have 6 options (5 + themselves), and individuals in the corners have 4

options (3 + themselves). Again there is an overall selective advantage to be chosen based on f_a and f_b , but r and x_i are calculated based on the $P_{i,j}$'s, which are dependent on the fitness and the particular spatial configuration of individuals.



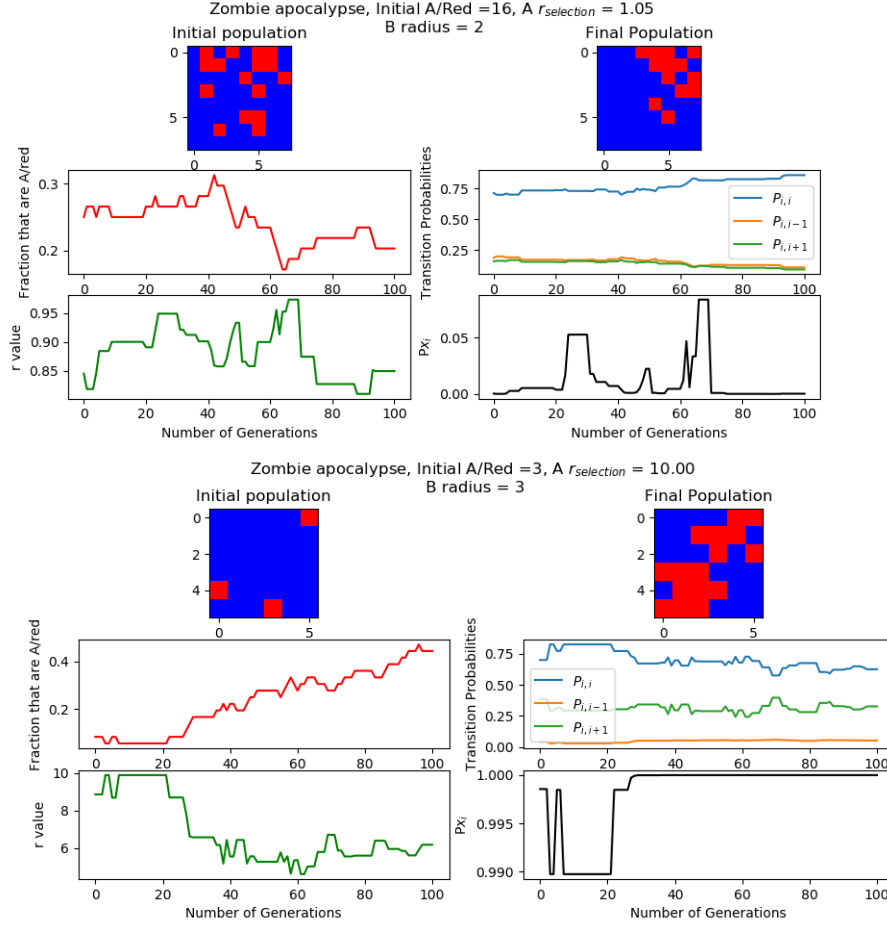
4.3 A Model of Spatially Dependent Selection with a Spatially Dependent Selective Advantage

Now, we'll consider a population of bacteria or another similar organism. The distribution of individuals in the population starts out random, but then a food source is introduced in one corner (the upper left in these figures) that gives A a significant selective advantage. However, the concentration of the food source falls off like $f(r, R) = r/R^2$, so that far away from the food source f_a is close to 1. And at the same time, the bacteria can only replace a neighboring individual as in the previous example. In this case, there is spatial dependence on both who is chosen to reproduce, and on who is replaced given whoever is chosen to reproduce. In this case then, the overall r and x_i very much depend on the particular arrangement of individuals on the lattice



4.4 The Zombie Apocalypse! (Selective Advantage with Spatial Dependence and an increased Radius of Reproduction)

A zombie apocalypse has occurred on an MxM lattice! The zombies, represented in red, are more aggressive than the humans and therefore have a fitness advantage in being chosen to reproduce of $f_z > f_h$ over the humans. The humans however, possess an antidote that turns any zombie back into a human, and the antidote can be shot a distance or $r_{\text{antidote}} = 2$. The antidote turns zombies back into people, but does nothing to humans who receive the antidote. I put 2 different examples here of what could happen with some different parameters. As you can imagine, there is a battle between which has more influence - the advantage in being selected, or the increased radius of reproduction. You can see in the 2nd example that a relatively large radius of reproduction for the blue humans loses out to the aggressiveness of the zombies (red) reproductive advantage.



5 Discussion

The Moran Process is a useful physical model of a well mixed population. For a real finite population, however, it makes sense to factor in spatial dependence when considering neutral drift and selection. Factoring in spatial dependence to the fixation probability requires updating r and x_i after every time step, which is much more complicated than the typical Moran process. This is necessary because, when individuals can only replace their neighbors, the transition probabilities are directly affected by the particular configuration of individuals on the lattice. As I've shown though, building up from simpler 1D examples can help to make this task much less daunting. This spatially dependent model can also form the basis for simple models of disease spread, with some adjustments made to account for the different parameters.