A Brief Exploration of Evolutionary Dynamics

Youze Zheng
hargen.zheng@gmail.com

Euler Circle

July 5, 2022

Table of Contents

What is Evolutionary Dynamics?

Introduction

Meaning of Evolutionary Dynamics

A gradual change in the characteristics of a population over successive generations.

But the topic could be broad and sometimes it involves the discussion of evolutionary game theory.

Deterministic view of evolution

Introduction

Definition

The evolutionary change occurs orderly, instead of due to random fluctuation.

Replication without limit: exponential model

Let $x(t)$ denote the size of population at time t. We can use the differential equation below to represent the rate of reproduction for the whole population:

$$\frac{dx}{dt} = rx.$$

Replication without limit: exponential model

Let $x(t)$ denote the size of population at time t. We can use the differential equation below to represent the rate of reproduction for the whole population:

$$\frac{dx}{dt} = rx.$$

The solution to the differential equation is

$$x(t) = x_0 e^{rt}$$

given that the size of population is x_0 at $t = 0$.

Replication: exponential model

Suppose the natural death rate for a particular population is d, we would have the following differential equation to represent the rate of change of population size:

$$\frac{dx}{dt} = (r - d)x.$$

Replication: exponential model

Suppose the natural death rate for a particular population is d, we would have the following differential equation to represent the rate of change of population size:

$$\frac{dx}{dt} = (r - d)x.$$

The solution the the differential equation is

$$x(t) = x_0 e^{(r-d)t}$$

given that the size of population is x_0 at $t = 0$.

Replication: exponential model

Suppose the natural death rate for a particular population is d, we would have the following differential equation to represent the rate of change of population size:

$$\frac{dx}{dt} = (r - d)x.$$

The solution the the differential equation is

$$x(t) = x_0 e^{(r-d)t}$$

given that the size of population is x_0 at $t = 0$.
If $d > r$, then the size of population will converge to zero. Conversely, if $d < r$, the size of population will grow to infinity.

Bounded replication: logistic model

Suppose the carrying capacity for one specific population is K, with the initial population size x_0, the number of individuals for that population at time t could be represented by the logistic equation as following:

$$\frac{dx}{dt} = rx(\frac{1-x}{K})$$

where r represents the rate of reproduction and population size is denoted by x.

Bounded replication: logistic model

Suppose the carrying capacity for one specific population is K, with the initial population size x_0, the number of individuals for that population at time t could be represented by the logistic equation as following:

$$\frac{dx}{dt} = rx(\frac{1-x}{K})$$

where r represents the rate of reproduction and population size is denoted by x. Since $x(0) = x_0$ and $x(t) = K$ as $t \to \infty$, the solution is

$$x(t) = \frac{Kx_0e^{rt}}{K + x_0(e^{rt} - 1)}.$$

Selection with limited reproduction

The whole environment will have a maximum carrying capacity.

- $x(t)$ denotes the proportion of population A in the environment.
- $y(t)$ denotes the proportion of population B in the environment.

By our definition to $x(t)$ and $y(t)$, we can conclude that $x + y = 1$ and $\phi = ax + by$ because they represent the proportions that range from 0 to 1.

$$\frac{dx}{dt} = x(a - \phi)$$
$$\frac{dy}{dt} = x(b - \phi).$$

Selection with limited reproduction

The whole environment will have a maximum carrying capacity.

- $x(t)$ denotes the proportion of population A in the environment.
- $y(t)$ denotes the proportion of population B in the environment.

By our definition to $x(t)$ and $y(t)$, we can conclude that $x + y = 1$ and $\phi = ax + by$ because they represent the proportions that range from 0 to 1.

$$\frac{dx}{dt} = x(a - \phi)$$
$$\frac{dy}{dt} = x(b - \phi).$$

By replacing y with $1 - x$, we have

$$\frac{dx}{dt} = x(1 - x)(a - b).$$

Visualize selection using simplex

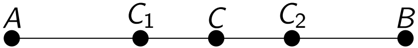

Figure: 1-simplex line segment, $\sigma =< A, B >$

- If population A reproduces at a faster rate throughout the time, meaning $dx/dt > dy/dt$ then at some time, the population distribution would be represented by C_1 and eventually by the end point A.
- Conversely, if $dx/dt < dy/dt$, then the population distribution would at some point be represented by point C_2, and ultimately end point B.
- If $dx/dt = dy/dt$, then the population distribution would remain at point C on the 1-simplex.

Visualize selection using simplex

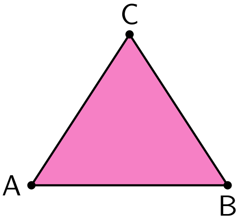

Figure: 2-simplex filled triangle, $\sigma =< A, B, C >$

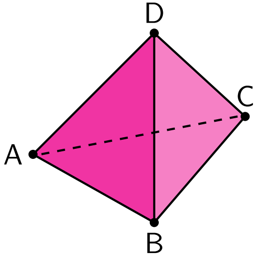

Figure: 3-simplex tetrahedron, $\sigma =< A, B, C, D >$

More complicated case with simplicial complex

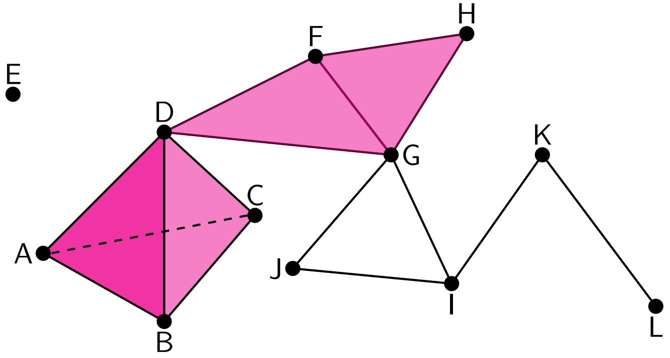

Figure: Simplicial Complex $\mathcal{K}$

Stochastic view of evolution

Introduction

Definition

Evolutionary changes by random, chance, or probability.

Moran Process

Named after the Australian statistician (geneticist) Patrick Moran in 1958, the Moran Process tells that at each time step, a random individual is chosen for reproduction and a random individual is chosen for elimination. The eliminated individual is replaced by the individual that is reproduced, so the total population remains unchanged. Since the size of population is unchanged, the process would end with one type of individual dominates the whole population.

Moran Process

Suppose there are only two types of individuals, type A and type B with total size of population N. If there are i individuals for type A population, then there are $N - i$ type B individuals for all $i = 0, 1, ..., N$.

- The probability of choosing one type A individual is i/N.

Moran Process

Suppose there are only two types of individuals, type A and type B with total size of population N. If there are i individuals for type A population, then there are $N - i$ type B individuals for all $i = 0, 1, ..., N$.

- The probability of choosing one type A individual is i/N.
- The probability of choosing one type B individual is $(N - i)/N$.

Birth-death Model I

Suppose we have a transition matrix $P = [p_{ij}]$, which represent the probability for an individual to move from position i to j, we would have the following expressions to represent every entries:

$$
\begin{aligned}
p_{i,i-1} &= \frac{i(N-i)}{N^2} \\
p_{i,i+1} &= \frac{i(N-i)}{N^2} \\
p_{i,i} &= 1 - p_{i,i-1} - p_{i,i+1} = \frac{(N-i)^2}{N^2} \\
p_{0,0} &= p_{N,N} = 1 \\
p_{0,i} &= p_{N,i} = 0
\end{aligned}
$$

Birth-death Model I

Let x_i denotes the probability for the number of population A to change from i to N. Then the probability for population B to dominate is $1 - x_i$. We would have the following general equations

$$\begin{aligned} x_0 &= 0 \\ x_N &= 1 \\ x_i &= p_{i,i-1}x_{i-1} + p_{i,i}x_i + p_{i,i+1}x_{i+1} \end{aligned}$$

for all $i = 1, 2, ..., (N-1)$.

Birth-dealth Model I

$$\begin{aligned} x_i &= p_{i,i-1}x_{i-1} + p_{i,i}x_x + p_{i,i+1}x_{i+1} \\ x_i &= p_{i,i-1}x_{i-1} + (1 - 2p_{i,i+1}) + p_{i,i+1}x_{i+1} \\ 2p_{i,i+1}x_i &= p_{i,i-1}x_{i-1} + p_{i,i+1}x_{i+1} \end{aligned}$$

for all $i = 0, 1, ..., N$.
Dividing both sides by $p_{i,i+1}$, we would get

$$2x_i = x_{i+1} + x_{i-1}$$

Birth-dealth Model I

$$\begin{aligned} x_i &= p_{i,i-1}x_{i-1} + p_{i,i}x_x + p_{i,i+1}x_{i+1} \\ x_i &= p_{i,i-1}x_{i-1} + (1 - 2p_{i,i+1}) + p_{i,i+1}x_{i+1} \\ 2p_{i,i+1}x_i &= p_{i,i-1}x_{i-1} + p_{i,i+1}x_{i+1} \end{aligned}$$

for all $i = 0, 1, ..., N$.
Dividing both sides by $p_{i,i+1}$, we would get

$$2x_i = x_{i+1} + x_{i-1}$$

Arithmatic sequence with common difference of $d = (x_N - x_0)/N - 0 = 1/N$.

Birth-dealth Model I

$$
\begin{aligned}
x_i &= p_{i,i-1}x_{i-1} + p_{i,i}x_x + p_{i,i+1}x_{i+1} \\
x_i &= p_{i,i-1}x_{i-1} + (1 - 2p_{i,i+1}) + p_{i,i+1}x_{i+1} \\
2p_{i,i+1}x_i &= p_{i,i-1}x_{i-1} + p_{i,i+1}x_{i+1}
\end{aligned}
$$

for all $i = 0, 1, ..., N$.
Dividing both sides by $p_{i,i+1}$, we would get

$$2x_i = x_{i+1} + x_{i-1}$$

Arithmatic sequence with common difference of
$d = (x_N - x_0)/N - 0 = 1/N$.

$$x_i = \frac{i}{N}$$

for all $i = 0, 1, ..., N$.

Birth-death Model II

$$x_i = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} \gamma_k}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k}$$

where $\gamma_i = p_{i,i-1} / p_{i,i+1}$.
If we take fitness into consideration, we would get

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

for all $i = 1, 2, ..., N$, where r is the fitness for population A (Geometric sequence).

Compute Fitness by linear regression

Gradient descent algorithm

$$f_{\vec{w},b}(\vec{x}) = \vec{w} * \vec{x} + b$$

where $\vec{w} = \begin{bmatrix} w_1 & w_2 & w_3 & \dots & w_n \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix}$, b is a constant, and n be the number of features.

Gradient descent algorithm

$$f_{\vec{w},b}(\vec{x}) = \vec{w} * \vec{x} + b$$

where $\vec{w} = \begin{bmatrix} w_1 & w_2 & w_3 & \dots & w_n \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix}$, b is a constant, and n be the number of features.

$$J_{(w,b)} = \frac{1}{2m} \sum_{i=1}^{m} (\hat{y^{(i)}} - y^{(i)})^2$$

where w and b are parameters and m is the total number of the traning data $(x^{(i)}, y^{(i)})$.

Method I: Gradient descent algorithm

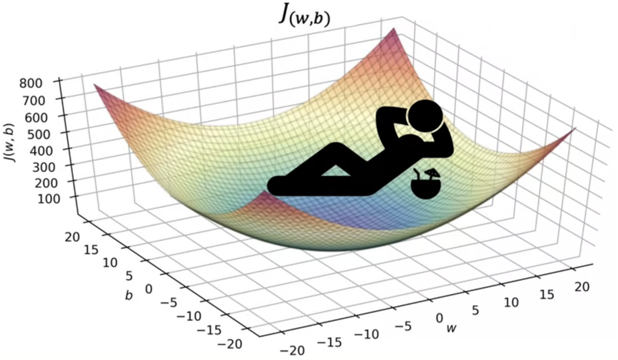

Figure: $3D$ plot for $J_{(w,b)}$, w, and b

Method I: Gradient descent algorithm

To do that, we will use the following two equations, which are main parts of the gradient descent algorithm:

$$w_1 = w_1 - \alpha \, \frac{1}{m} \sum_{i=1}^{m} (f_{\vec{w},b}(\vec{x}^{(i)}) - y^{(i)}) x_1^{(i)}$$

$$\vdots$$

$$w_n = w_1 - \alpha \, \frac{1}{m} \sum_{i=1}^{m} (f_{\vec{w},b}(\vec{x}^{(i)}) - y^{(i)}) x_n^{(i)}$$

$$b = b - \alpha \, \frac{1}{m} \sum_{i=1}^{m} (f_{w,b}(x^{(i)}) - y^{(i)})$$

where α is called the learning rate, and w, b are parameters for the training model and the cost function.

Method II: Ordinary Least Squares (OSL)

We want to model a linear equation as following:

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}$$

where $\boldsymbol{X}$ is a $1 \times n$ matrix representing n input features, $\boldsymbol{y}$ is a 1×1 square matrix representing n estimated outcomes, and $\hat{\beta}$ is a $n \times 1$ column matrix representing the weight of each input feature.
The OSL technique can help us compute the vector $\hat{\beta}$ that represents the weight of each input feature $\boldsymbol{X}$ by the following equation:

$$\hat{\beta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}.$$

Proof on my expository paper!

Remarks

- Gradient descent is a choice; it approximates the result (learning rate). However, ordinary least squares technique is better option for linear regression, as it gives the exact closed form solution.
- The relationship between input features and output fitness may not be linear.
- Evolutionary dynamics is a broad topic. Still a lot of things to explore.

Thank you! Questions on Euler Circle discord server!