

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Diogo Taylor Ciciliotti

PROCESSO DE MORAN EM GRAFOS EVOLUTIVOS E O
TEOREMA DE PERRON-FROBENIUS

Vitória
2022

Agradecimentos

Em primeiro lugar, gostaria de agradecer à minha família pela força, paciência e pelo amor incondicional. Agradeço à minha mãe por ser um exemplo constante de resiliência, me mostrando como acordar todos os dias sem abaixar a cabeça para os problemas da vida. Agradeço ao meu pai por ser essa pessoa extraordinária, que defende a sua família e que ama seus filhos. Agradeço à minha irmã, que me ensina a ser uma pessoa melhor sempre que consigo vê-la.

Agradeço aos meus amigos, os presenciais e aos virtuais. Aos colegas da graduação e do mestrado. As pessoas que me apoiam, que compartilharam lágrimas e risadas nessa longa caminhada. Sem vocês, a vida seria um bem mais cinza e sem-graça.

Gostaria de agradecer à todos os professores que de alguma maneira contribuíram para minha formação acadêmica e pessoal. Em especial, agradeço ao meu professor orientador Fabiano Petronetto do Carmo, meu pai acadêmico. Agradeço pelas palavras de ajuda, pela inspiração profissional, pela paciência com as dificuldades que surgiram e pelo engrandecimento pessoal. Obrigado por estar sempre ao meu lado nessa jornada.

Agradeço ao Fundação de Amparo à Pesquisa do Estado do Espírito Santo (FAPES) pela concessão de bolsa de estudo durante o curso de Mestrado.

Resumo

Nesse trabalho, iremos apresentar alguns dos resultados mais importantes relacionados ao Processo de Moran simples - uma cadeia de Markov em tempo discreto - e as probabilidades de fixação de um gene mutante como resultado do Teorema de Perron-Frobenius. Posteriormente, iremos generalizar a estrutura populacional do modelo, identificando os indivíduos em grafos orientados e destacando estruturas populacionais que atuam como supressores ou amplificadores de seleção natural.

Palavras-chave: Processo de Moran; cadeias de Markov; *fitness*; probabilidade de fixação; dinâmica evolutiva em grafos; teorema isotérmico; teorema de Perron-Frobenius

Abstract

In this work we shall present some of the most important results related to the (simple) Moran Process - a discrete-time Markov chain - and the fixation probability of a mutant gene as a result of the Perron-Frobenius Theorem. In addition, we will generalize the model population structure, labeling the individuals in directed graphs and highlighting population structures that act as natural selection suppressors or natural selection amplifiers.

Palavras-chave: Moran process; Markov chain; *fitness*; fixation probability; evolutionary dynamics on graphs; isothermal theorem; Perron-Frobenius theorem

Conteúdo

1	Introdução	7
2	Preliminares	9
2.1	Vetores e Matrizes	9
2.2	Cadeias de Markov	16
2.3	Grafos	20
3	Processo de Moran Simples	22
3.1	Modelo do Processo de Moran	22
3.2	Probabilidades de seleção de tipo e transição de estado	23
3.3	<i>Fitness</i> e neutralidade	25
3.4	Matriz de Markov para o Processo de Moran simples	30
3.5	Vetor de probabilidades de fixação	31
4	Grafos Evolutivos	38
4.1	Estrutura Populacional	38
4.2	Probabilidades de fixação	39
4.2.1	O Ciclo direcionado	39
4.2.2	O Ciclo Bidirecionado	42
4.2.3	A Linha e o <i>Burst</i>	43
4.2.4	<i>Drift</i> Genético e Seleção	45
4.3	O Teorema Isotérmico	45

4.4	Supressores de Seleção	48
4.5	Amplificadores de seleção	49
5	Conclusão	51
	Apêndices	52
A	Raio Espectral e Norma Matricial	52
B	Teorema de Perron-Frobenius	59
B.1	Teorema dos Discos de Gershgorin	59
B.2	Fórmula de Collatz-Wielandt	60
B.3	Teorema de Perron-Frobenius	63
	Referências Bibliográficas	66

1 Introdução

O Processo de Moran é um processo estocástico criado para modelar a difusão de mutações genéticas dos organismos de uma população. Esse processo tem sua origem no trabalho *Random processes in genetics* (Moran, 1958), no qual ele adaptou o modelo criado pelo geneticista Sewell Wright baseado nos princípios de genética mendeliana e no surgimento espontâneo de gametas mutantes ([11]). Embora seu surgimento esteja relacionado com a ecologia, o Processo de Moran e os processos similares podem ser adaptados para modelar a propagação de doenças, comportamento de votantes, difusão de ideologias em redes sociais e evolução de sistemas sociais e demográficos ([7], [8]).

O Processo de Moran é uma cadeia de Markov em tempo discreto, com população total de N indivíduos. O modelo também considera que a população é bem-distribuída, e portanto a configuração dos indivíduos no espaço é desprezada. Em cada passo de tempo da cadeia de Markov, um dos indivíduos é escolhido para reproduzir e outro é escolhido para morrer, mantendo a população constante ao final de cada geração. O indivíduo que reproduziu cria uma cópia exata de si mesmo, e seus genes são transmitidos para a geração futura. Para este trabalho, iremos considerar que os indivíduos possuem duas opções de genes, denotados \mathbb{A} e \mathbb{B} , e que o modelo não possui mutações (um indivíduo de gene \mathbb{A} não pode gerar um descendente de gene \mathbb{B} , e vice-versa).

Deste modo, quando só há indivíduos de um tipo no Processo (por exemplo, o modelo tem apenas indivíduos do tipo \mathbb{A}), então a cadeia de Markov atinge um estado de absorção e o Processo de Moran fica completamente determinado. Como neste trabalho consideraremos apenas dois tipos de genes populacionais, então os estados de absorção são 0 (apenas indivíduos do tipo \mathbb{B} existem no modelo) e N (apenas indivíduos do tipo \mathbb{A} existem no modelo).

A seguir, apresentaremos uma descrição informal dos conceitos básicos desenvolvidos nos capítulos.

No Capítulo 2 - Preliminares, iremos descrever os principais resultados das áreas de conhecimento relacionadas ao trabalho: matrizes não-negativas e irredutíveis; processos estocásticos e cadeia de Markov e Teoria de Grafos. Além dos resultados, o capítulo introduz o leitor as notações que serão utilizadas no corpo do texto.

No Capítulo 3 - Processo de Moran Simples, iremos definir precisamente o modelo

estocástico e utilizar as teorias desenvolvidas para descrever o Processo de Moran e algumas das principais propriedades. Introduziremos a noção de *fitness* associadas aos tipos genéticos e observaremos que, nesse processo, o tipo genético com maior *fitness* tem maior probabilidade de se fixar na população inteira.

Por fim, no Capítulo 4 - Grafos Evolutivos, iremos generalizar a estrutura populacional identificando os indivíduos em vértices de um grafo, e as arestas indicam as possíveis interações de indivíduos. O objetivo nesse capítulo é determinar a fixação da população a partir do surgimento de um indivíduo mutante em um vértice aleatório. Iremos também caracterizar os grafos cujo comportamento de fixação do indivíduo mutante é numericamente igual à fixação do indivíduo mutante no Processo de Moran Simples.

Nos Apêndices, apresentaremos os principais teoremas que correlacionam o raio espectral de uma matriz com sua norma. Utilizaremos esses resultados para demonstrar o Teorema de Perron-Frobenius, um resultado clássico (aqui demonstrado na sua forma de matrizes não-negativas e irredutíveis) e que permite análise assintótica do comportamento do Processo de Moran simples para obtenção das probabilidades de fixação dos indivíduos.

2 Preliminares

Nessa seção, iremos introduzir alguns conceitos básicos que são necessários no decorrer do trabalho. Em particular, iremos destacar noções relacionadas à processos estocásticos e cadeias de Markov. As notações e definições introduzidas aqui são baseadas nos trabalhos de Karlin & Taylor [9], Grinstead & Snell [4], e as definições relativas à vetores e matrizes são baseadas em [10].

2.1 Vetores e Matrizes

Definição 2.1 (Espaço vetorial). Um espaço vetorial sobre o corpo K é a terna $(V, +, \cdot)$, em que V é um conjunto não-vazio e $+$ e \cdot são respectivamente as operações de soma em V e produto por escalar satisfazendo as propriedades:

- Associatividade da soma: $u, v, w \in V \Rightarrow (u + v) + w = u + (v + w)$
- Comutatividade da soma: $u, v \in V \Rightarrow u + v = v + u$
- Vetor zero: Existe um vetor $0 \in V$ tal que para todo $u \in V$ temos $u + 0 = u$
- Inverso aditivo: Para todo vetor $u \in V$, existe vetor $-u \in V$ tal que $u + (-u) = 0$
- Distributividade: Sejam $\alpha, \beta \in K$ e $u, v \in V$. Então $(\alpha + \beta)u = \alpha u + \beta u$ e $\alpha(u + v) = \alpha u + \alpha v$.
- Multiplicação por 1: Seja $1 \in K$ o elemento identidade da multiplicação do corpo K . Para todo $u \in V$, temos $1 \cdot u = u$

Os elementos do espaço vetorial são ditos vetores. Neste trabalho, consideraremos os espaços vetoriais sobre o corpo dos números reais \mathbb{R} . O conjunto $\mathbb{R}^{n \times m}$ representa o espaço vetorial das matrizes de ordem $n \times m$ (com as operações de soma e produto por escalar usuais). A notação de matriz em parênteses da forma $A = (a_{i,j})$ indica a matriz em que o elemento da i -ésima linha e da j -ésima coluna é $a_{i,j}$ e a notação $(x)_i$

referente à n -upla¹ $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ indica a i -ésima coordenada do vetor x , a saber, x_i . Por fim, escrevemos $0 = 0_n$ o vetor nulo (com todas as coordenadas nulas) de \mathbb{R}^n e $0 = 0_{n \times n}$ e $I = I_{n \times n}$ respectivamente as matrizes nulas (com todas os elementos nulos) e matriz identidade (com os elementos da diagonal principal iguais a 1 e demais elementos nulos) de $\mathbb{R}^{n \times n}$.

Definição 2.2 (Não-negativa). Uma matriz $A \in \mathbb{R}^{n \times m}$ é não-negativa se todos os seus elementos são não-negativos. Um vetor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ é não-negativo se o respectivo vetor-linha é não-negativo.

Sejam $A = (a_{i,j}), B = (b_{i,j}) \in \mathbb{R}^{n \times m}$. Escrevemos $A \geq B$ se $a_{i,j} \geq b_{i,j}$ para todo $i, j = 1, \dots, n$ e $A > B$ se $A \geq B$ e $A \neq B$. Também escrevemos $A \gg B$ se $a_{i,j} > b_{i,j}$ para todo $i, j = 1, \dots, n$.

Lema 2.1. Sejam $A = (a_{i,j}), B = (b_{i,j}) \in \mathbb{R}^{n \times n}$ matrizes quadradas não-negativas e $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ vetor não-negativo. Se $A \geq B$, então $Ax \geq Bx$.

Demonstração. Como $A \geq B \geq 0$ e $x \geq 0$, temos que $a_{i,j} \geq b_{i,j}$ e $a_{i,j}, b_{i,j}, x_i \geq 0$ para todos $i, j = 1, \dots, n$. Daí, da definição de produto entre matriz e vetor, vemos que, como todo índice $i, j = 1, \dots, n$, temos que:

$$(Ax)_i = \sum_{j=1}^n a_{i,j}x_j \geq \sum_{j=1}^n b_{i,j}x_j = (Bx)_i$$

Pela desigualdade acima, segue que $Ax \geq Bx$. □

Lema 2.2. Sejam $A = (a_{i,j}), B = (b_{i,j}) \in \mathbb{R}^{n \times n}$ matrizes quadradas. Então se $0 \leq A \leq B$, então $0 \leq A^m \leq B^m$ para todo $m \in \mathbb{N}$.

Demonstração. Escreva $A^m = (a_{i,j}^{(m)})$ e $B^m = (b_{i,j}^{(m)})$. Provaremos este lema por indução sobre m .

Como $B \geq A \geq 0$, então temos que $b_{i,j} \geq a_{i,j} \geq 0$ para todos índices $i, j = 1, \dots, n$. Pela definição de produto de matrizes, temos que:

¹Neste trabalho, chamaremos as n -uplas simplesmente de vetores. Assumiremos também os isomorfismos entre o espaço vetorial de n -uplas $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ e o espaço vetorial dos vetores-coluna $x = (x_{i,1}) \in \mathbb{R}^{n \times 1}$ no produto Ax e entre n -uplas $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ e o espaço vetorial dos vetores-linha $x = (x_{1,j}) \in \mathbb{R}^{1 \times n}$ no produto xA .

$$b^{(2)}_{i,j} = \sum_{k=1}^n b_{i,k} b_{k,j} \geq \sum_{k=1}^n a_{i,k} a_{k,j} = a^{(2)}_{i,j} \geq 0$$

Como os índices são quaisquer, segue que $B \geq A \geq 0$.

Agora, suponha que o lema é válido para $1, 2, \dots, m$. Por hipótese de indução, temos que $b^{(m)}_{i,j} \geq a^{(m)}_{i,j} \geq 0$ para todos índices $i, j = 1, \dots, n$. Como $B^{m+1} = B^m B$ e $A^{m+1} = A^m A$, pela definição de produto entre matrizes,

$$b^{(m+1)}_{i,j} = \sum_{k=1}^n b^{(m)}_{i,k} b_{k,j} \leq \sum_{k=1}^n a^{(m)}_{i,k} a_{k,j} = a^{(m+1)}_{i,j} \geq 0$$

E portanto temos que $B^{m+1} \geq A^{m+1} \geq 0$. Por indução, segue o enunciado. \square

Definição 2.3 (Positiva). Uma matriz $A = (a_{i,j}) \in \mathbb{R}^{n \times m}$ é positiva se $a_{i,j} > 0$ para todos os índices $i \in \{1, \dots, n\}, j \in \{1, \dots, m\}$. Um vetor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ é positivo se $x \gg 0$.

Corolário 2.2.1. Sejam $A = (a_{i,j}), B = (b_{i,j}) \in \mathbb{R}^{n \times n}$ matrizes quadradas não-negativas e $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ vetor positivo. Se $A > B$, então $Ax > Bx$.

Demonstração. Pelo Lema 2.1, temos que $Ax \geq Bx$. Portanto, resta mostrar que $Ax \neq Bx$. De fato, de $A > B$, segue que existe índices $i, j \in \{1, \dots, n\}$ tais que $a_{i,j} > b_{i,j}$. Como $x \gg 0$, então $x_j > 0$ e portanto $a_{i,j}x_j > b_{i,j}x_j$. Da definição de matriz e produto e como os termos do somatório são não-negativos, temos que:

$$(Ax)_i = \sum_{k=1}^n a_{i,k}x_k > \sum_{k=1}^n b_{i,k}x_k = (Bx)_i$$

Daí, $Ax \neq Bx$ e portanto $Ax > Bx$. \square

Lema 2.3. Seja $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$ matriz quadrada. Então $A \gg 0$ se e somente se, $Ax \gg 0$ para todo vetor $x \in \mathbb{R}^n, x > 0$.

Demonstração. Suponha que $A \gg 0$ e seja $x = (x_1, \dots, x_n) > 0$. vetor qualquer. Como $x \geq 0, x \neq 0$, então existe índice $j \in \{1, \dots, n\}$ tal que $(x)_j = x_j > 0$. Portanto, para todo $i \in \{1, \dots, n\}$, temos que:

$$(Ax)_i = \sum_{k=1}^n a_{i,k}x_k \geq a_{i,j}x_j > 0$$

Como i é índice arbitrário, então $Ax \gg 0$.

Por outro lado, suponha que $Ax \gg 0$ para todo vetor $x \in \mathbb{R}^n, x > 0$. Suponha por absurdo que $a_{i,j} = 0$ para algum par de índices $i, j \in \{1, \dots, n\}$ e defina o vetor (canônico) e_j como $(e_j)_k = 0$ se $k \neq j$ e $(e_j)_j = 1$. Como e_j é não-negativo e não-nulo, então $e_j > 0$. Então:

$$(Ae_j)_i = \sum_{k=1}^n a_{i,k}(e_j)_k = a_{i,j}(e_j)_j + \sum_{\substack{k=1, \dots, n \\ k \neq j}} a_{i,k}(e_j)_k = 0$$

pois $a_{i,j} = 0$ e $k = 1, \dots, n, k \neq j \Rightarrow (e_j)_k = 0$. Absurdo pois por hipótese $e_j > 0 \Rightarrow Ae_j \gg 0$. \square

Definição 2.4 (Estocástica). Uma matriz estocástica (ou linha-estocástica) é uma matriz quadrada não-negativa $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$ satisfazendo:

$$\sum_{j=1}^n a_{i,j} = 1, \quad \text{para todo } i = 1, \dots, n$$

Definição 2.5 (Primitiva). Uma matriz quadrada não-negativa $A \in \mathbb{R}^{n \times n}$ é primitiva se existe $N \in \mathbb{N}$ tal que A^N é positiva.

Definição 2.6 (Irreduzível). Uma matriz quadrada não-negativa $A \in \mathbb{R}^{n \times n}$ é irreduzível se, para todo $i, j \in \{1, \dots, n\}$, existe $p = p(i, j) \in \mathbb{N}$ tal que $a_{i,j}^{(p)} > 0$, onde $A^p = (a_{i,j}^{(p)})$.

Uma matriz quadrada não-negativa que não é irreduzível é dita reduzível. Observamos que, em matrizes irreduzíveis, o expoente p pode variar de acordo com os índices i e j . Isto não ocorre nas matrizes primitivas, nas quais existe um expoente p tal que $a_{i,j}^{(p)} > 0$ para todos índices i, j . Esta observação prova o seguinte lema.

Lema 2.4. Toda matriz quadrada não-negativa primitiva é irreduzível.

Lema 2.5. Sejam $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$ matriz quadrada não-negativa e irredutível e $x = (x_1, \dots, x_n) \geq 0$. Se $Ax = 0$, então $x = 0$.

Demonstração. Suponha por absurdo que exista índice $k \in \{1, \dots, n\}$ tal que $x_k > 0$. Então, para todo $i = 1, \dots, n$, temos que:

$$0 \leq a_{i,k}x_k \leq \sum_{j=1}^n a_{i,j}x_j = (Ax)_i = 0 \Rightarrow a_{i,k}x_k = 0$$

Como $x_k > 0$, segue que $a_{i,k} = 0$ para todo $i = 1, \dots, n$.

Agora, vamos mostrar que se existe expoente $m \in \mathbb{N}$ tal que $a_{i,k}^{(m)} = 0$ para todo $i = 1, \dots, n$, então $a_{i,k}^{(m+1)} = 0$ para todo $i = 1, \dots, n$. De fato, se $i = 1, \dots, n$, então da igualdade $A^{m+1} = A^m A$, temos que:

$$a_{i,k}^{(m+1)} = \sum_{j=1}^n a_{i,k}^{(m)} a_{i,k} = 0$$

Assim, pelo princípio indutivo, não existe expoente m tal que $a_{i,k}^{(m)} > 0$ para $i = 1, \dots, n$. Portanto A é redutível, um absurdo. \square

Corolário 2.5.1. Seja $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$ matriz quadrada não-negativa. Se A é matriz irredutível, então os vetores-coluna e os vetores-linhas da matriz A são não-nulos.

Demonstração. Na demonstração do Teorema 2.5, provamos que se existe expoente $m \in \mathbb{N}$ tal que o k -ésimo vetor-coluna de A^m é nulo para algum $k \in \{1, \dots, n\}$, então o k -ésimo vetor-coluna da matriz A^{m+1} também será nulo. Deste modo, se existir algum vetor-coluna de A nulo, então a matriz A é redutível, e portanto provamos o enunciado por contraposição.

Por outro lado, se A possui algum vetor-linha nulo, então A^T possui (pelo menos) um vetor-coluna nulo. Resta mostrar que, se A é irredutível, então A^T é também irredutível e obtemos uma contradição. De fato, dado um par de índices arbitrários $i, j \in \{1, \dots, n\}$, então por hipótese existe índice $p \in \mathbb{N}$ tal que $a_{j,i}^{(p)} > 0$. \square

Corolário 2.5.2. Sejam $A = (a_{i,j}), B = (b_{i,j}) \in \mathbb{R}^{n \times n}$ matrizes quadradas não-negativas irredutíveis. Então se A é positiva, então AB é positiva.

Demonstração. Sejam $i, j \in \{1, \dots, n\}$ par de índices arbitrários fixados. Por definição:

$$ab_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}$$

onde $AB = (ab_{i,j})$. Como A é positiva, então $a_{i,k} > 0$ para todo $k = 1, \dots, n$. Como B é irredutível, pelo Corolário 2.5.1, existe algum $k \in \{1, \dots, n\}$ tal que $b_{k,j} > 0$ e portanto $ab_{i,j} > 0$. Como i, j são índices arbitrários, então $AB \gg 0$. \square

Lema 2.6. Seja $A \in \mathbb{R}^{n \times n}$ matriz quadrada não-negativa tal que existe expoente $p \in \mathbb{N}$ tal que A^p é irredutível. Então A é irredutível.

Demonstração. Sejam $i, j \in \{1, \dots, n\}$. Como A^p é irredutível, então existe expoente $q \in \mathbb{N}$ tal que o elemento da i -ésima linha e da j -ésima coluna da matriz $(A^p)^q = A^{pq}$ é positivo. Deste modo, $a_{i,j}^{(pq)} > 0$ e portanto A é irredutível. \square

Teorema 2.7. Seja $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$ matriz quadrada não-negativa e irredutível. Então $(A + I)^{n-1} \gg 0$.

Demonstração. Pelo Lema 2.3, basta mostrar que $(A + I)x^{n-1} \gg 0$ para todo vetor $x \in \mathbb{R}^n, x > 0$. De fato, seja $x > 0$ arbitrário.

Caso 1: Suponha que exista índice $K \in \{1, \dots, n\}$ tal que $(x)_K = 0$ e defina o vetor $y = y(x) = (A + I)x = Ax + x$. Vemos que, como A é não-negativa,

$$(x)_i > 0 \Rightarrow (y)_i = (Ax)_i + (x)_i > 0, \quad i \in \{1, \dots, n\}$$

e portanto $y_i = 0 \Rightarrow x_i = 0$. Logo a quantidade de entradas nulas em y deve ser menor ou igual a quantidade de entradas nulas em x .

Suponha por absurdo que exista algum vetor x tal que a quantidade de entradas nulas em y é a mesma que do vetor x e escreva $S \subset \{1, \dots, n\}$ conjunto de índices tais que $i \in S \Rightarrow (x)_i = (y)_i = 0$ e $i \notin S \Rightarrow (x)_i > 0$ e $(y)_i > 0$. Vemos que $S \neq \{1, \dots, n\}$ pois $x > 0$ e $K \in S$ pois $(x)_K = 0$. Portanto:

$$(y)_i = \left(\sum_{k=1}^n a_{i,k}(x)_k \right) + (x)_i = \left(\sum_{\substack{k=1, \dots, n \\ k \notin S}} a_{i,k}(x)_k \right) + (x)_i$$

Assim, para todo $i \in S$, temos que $(x)_i = (y)_i = 0$ e

$$0 = (y)_i = \left(\sum_{\substack{k=1, \dots, n \\ k \notin S}} a_{i,k}(x)_k \right) + (x)_i = \sum_{\substack{k=1, \dots, n \\ k \notin S}} a_{i,k}(x)_k$$

Como $k \notin S \Rightarrow (x)_k > 0$, então temos que $i \in S \Rightarrow a_{i,k} = 0$ para todo $k \notin S$.

Vamos mostrar agora que, se existe expoente $m \in \mathbb{N}$ tal que para todos índices i, j com $i \in S$ e $j \notin S$ temos $a_{i,j}^{(m)} = 0$, então $a_{i,j}^{(m+1)} = 0$ para todos índices i, j com $i \in S$ e $j \notin S$. De fato, sejam $i, j \in \{1, \dots, n\}$ tais que $i \in S$ e $j \notin S$. Da igualdade $A^{m+1} = A^m A$, temos que:

$$\begin{aligned} a_{i,j}^{(m+1)} &= \sum_{k=1}^n a_{i,k}^{(m)} a_{k,j} = \left(\sum_{\substack{k=1, \dots, n \\ k \in S}} a_{i,k}^{(m)} a_{k,j} \right) + \left(\sum_{\substack{k=1, \dots, n \\ k \notin S}} a_{i,k}^{(m)} a_{k,j} \right) = \\ &= \left(\sum_{\substack{k=1, \dots, n \\ k \in S}} a_{i,k}^{(m)} \cdot 0 \right) + \left(\sum_{\substack{k=1, \dots, n \\ k \notin S}} 0 \cdot a_{k,j} \right) = 0 \end{aligned}$$

Pelo princípio indutivo, segue que não existe expoente m tal que $a_{i,j}^{(m)} > 0$ para índices da forma $i \in S, j \notin S$, um absurdo, pois A é irredutível. Assim, o vetor $y = (A + I)x$ tem uma quantidade de entradas nulas estritamente menor do que o vetor x . Como x é não-nulo, então a quantidade de entradas nulas é no máximo $n - 1$ e portanto o vetor $(A + I)x$ tem no máximo $n - 2$ entradas nulas. Aplicando o argumento recursivamente para o conjunto de vetores $\{x, (A + I)x, (A + I)^2x, \dots, (A + I)^{n-2}x, (A + I)^{n-1}x\}$, vemos que $(A + I)^{n-1}x$ não possui entradas nulas e portanto $(A + I)^{n-1}x \gg 0$

Caso 2: Suponha que $x \gg 0$. Como A é matriz não-negativa, pela expansão de $(I + A)^{n+1}$, vemos que:

$$(A + I)^{n-1} = I + A + \binom{n-1}{2} A^2 + A^{n-1} \geq I$$

Portanto, como x é positivo,

$$(A + I)^{n-1} x \geq Ix \gg 0$$

e segue o enunciado. □

2.2 Cadeias de Markov

Definição 2.7 (Espaço de Probabilidade). Um espaço de probabilidade é uma terna $(\Omega, \mathcal{F}, \mathbf{P})$, em que Ω é o espaço amostral (espaço de todos os possíveis resultados), \mathcal{F} é uma σ -álgebra² de Ω (os conjuntos $E \in \mathcal{F}$ são chamados de eventos) e $\mathbf{P} : \mathcal{F} \rightarrow [0, 1]$ é uma medida de probabilidade σ -aditiva³

Definição 2.8 (Variável Aleatória). Dado um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbf{P})$, uma variável aleatória $X : \Omega \rightarrow \mathbb{R}$ é uma função tal que o conjunto $\{w \in \Omega : X(w) \leq r\} \in \mathcal{F}$ para todo $r \in \mathbb{R}$. O valor $X(w)$, $w \in \Omega$ nos dá uma ideia de medição do resultado w . Definimos a probabilidade $\mathbf{P}(X \leq r) := \mathbf{P}(\{w : X(w) \leq r\})$. Se Ω é finito ou infinito enumerável, então a variável aleatória X é denominada discreta.

Definição 2.9 (Esperança). Dado um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbf{P})$ e uma variável aleatória discreta $X : \Omega \rightarrow \mathbb{R}$, definimos a esperança de X como $\mathbb{E}[X] = \sum_{i=1}^{\infty} x_i \mathbf{P}(X = x_i)$, em que $\Omega = \{w_1, \dots, w_n, \dots\}$ e $X(w_i) = x_i$, $i \in \mathbb{N}$.

Definição 2.10 (Processos Estocásticos). Um processo estocástico é uma família $\{X_t, t \in T\}$, em que X_t é uma variável aleatória para cada índice $t \in T$. Intuitivamente, o conjunto de índices T representa uma ideia de unidades de tempo e X_t representa os possíveis resultados de uma repetição de experimento.

Definição 2.11 (Estado de espaço). O estado de espaço S é o conjunto que contém todos os possíveis valores que as variáveis X_t podem assumir para todos $t \in T$. Se S

²Uma σ -álgebra \mathcal{F} satisfaz: $\emptyset \in \mathcal{F}$; \mathcal{F} é fechado por uniões enumeráveis; e $E \in \mathcal{F} \Rightarrow E^c \in \mathcal{F}$

³ $\mathbf{P}(\emptyset) = 0$; $\mathbf{P}(\Omega) = 1$ e $\mathbf{P}(\bigcup_{i \in I} E_i) = \sum_{i \in I} \mathbf{P}(E_i)$ se $\{E_i, i \in I\}$ for coleção enumerável de eventos dois a dois disjuntos.

é finito ou S é infinito enumerável, dizemos que o processo estocástico é discreto. Se $S \subset \mathbb{R}$, então dizemos que o processo estocástico assume valores reais.

Definição 2.12 (Parâmetro de índices). O parâmetro de índices T é uma coleção de índices. Se T é finito ou infinito enumerável, dizemos que o processo estocástico é em tempo discreto. Se $T = [0, +\infty)$, dizemos que o processo estocástico é em tempo contínuo.

Definição 2.13 (Martingal). Seja $\{X_t, t \in T\}$ processo estocástico que assume valores reais em tempo discreto ou contínuo. Dizemos que $\{X_t\}$ é um martingal se $\mathbb{E}[|X_t|] < \infty$ para todo $t \in T$ e, se para quaisquer $t_1 < t_2 < \dots < t_{n+1}$, então $\mathbb{E}[X_{t_{n+1}} | X_{t_1} = a_1, \dots, X_{t_n} = a_n] = a_n$ para todos $t_1, \dots, t_{n+1} \in T$ e $a_1, \dots, a_n \in \mathbb{R}$.

Definição 2.14 (Processo de Markov). Um processo de Markov é um processo estocástico com a propriedade de, dado um valor X_t , então as probabilidades dos valores X_s , $s > t$ são independentes dos valores X_r , $r < t$. Em termos formais, um processo é dito de Markov se satisfaz $\mathbf{P}(a < X_t \leq b | X_{t_1} = a_1, X_{t_2} = a_2, \dots, X_{t_n} = a_n) = \mathbf{P}(a < X_t \leq b | X_{t_n} = a_n)$ para todos $t_1 < t_2 < \dots < t_n < t \in T$ e $a, b, a_1, \dots, a_n \in S^4$. Se S é discreto, dizemos que o processo de Markov é uma cadeia de Markov.

Uma cadeia de Markov $\{X_t\}$ é um processo estocástico de Markov em tempo discreto e estado de espaço discreto. De maneira geral, adotamos $T = \{0, 1, \dots\}$ e, para cada $t \in T$, a variável X_t representa o resultado do t -ésimo experimento. Dizemos que X_t está no estado i se é dado que $X_t = i$, $i \in S$.

A probabilidade de X_{t+1} ficar no estado j , dado que X_t está no estado i , denotado por $\mathbf{P}_{i,j}^{t,t+1}$ é:

$$P_{i,j}^{t,t+1} = \mathbf{P}(X_{t+1} = j | X_t = i)$$

As probabilidades $P_{i,j}^{t,t+1}$ são conhecidas como probabilidades de transição em um passo de tempo. A notação ainda sugere que essas probabilidades de transição, além de dependerem dos estados i e j (inicial e final), são também dependentes do passo de tempo no qual ocorrem (t para $t + 1$). Quando as probabilidades de transição não dependem do tempo ($P_{i,j}^{t_1,t_1+1} = P_{i,j}^{t_2,t_2+1}$ para todos $t_1, t_2 \in T$), dizemos que a cadeia de Markov tem as probabilidades de transições estacionárias. Supondo que a cadeia

⁴Dizemos que $\{X_t\}$ satisfazem a propriedade de Markov

de Markov é discreta e tem probabilidades de transições estacionárias, é possível então definir a matriz de Markov.

Definição 2.15 (Matriz de Markov). Sejam $\{X_t\}$ uma cadeia de Markov discreta com probabilidades de transições estacionárias ($P_{i,j}^{t,t+1} = P_{i,j}$ para todo $t \in T$) e $S = \{0, \dots, N\}$ estado de espaço. Definimos a matriz de Markov $M \in \mathbb{R}^{(N+1) \times (N+1)}$ (ou matriz de transições de probabilidade) associada a cadeia de Markov como:

$$M = \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \dots & P_{0,N} \\ P_{1,0} & P_{1,1} & P_{1,2} & \dots & P_{1,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{N,0} & P_{N,1} & P_{N,2} & \dots & P_{N,N} \end{bmatrix} \quad (1)$$

Como $P_{i,j} \geq 0$ e $\sum_{j \in S} P_{i,j} = 1$, então a matriz de Markov M é não-negativa e linha-estocástica.

Teorema 2.8. Seja M matriz de Markov como em (1). A probabilidade de, ao começar no estado i , depois de n passos de tempo, a cadeia estar no estado j , denotado por $P_{i,j}^{t,t+n}$ é o elemento $P_{i,j}^{(n)}$ da matriz $M^n = (P_{i,j}^{(n)})_{ij}$. Isto é, os elementos da n -ésima potência da matriz de Markov M representam as probabilidades de transições de estado em n passos de tempo.

Demonstração. Inicialmente, vemos que pela Lei da Probabilidade Total (aplicada a probabilidades condicionais), temos que:

$$\mathbf{P}(X_{t+2} = j | X_t = i) = \sum_{k \in S} \mathbf{P}(X_{t+2} = j | X_{t+1} = k, X_t = i) \cdot \mathbf{P}(X_{t+1} = k | X_t = i)$$

Como $\{X_t\}$ tem a propriedade de Markov, segue que:

$$\mathbf{P}(X_{t+2} = j | X_t = i) = \sum_{k \in S} \mathbf{P}(X_{t+2} = j | X_{t+1} = k) \cdot \mathbf{P}(X_{t+1} = k | X_t = i)$$

e portanto

$$P_{i,j}^{t,t+2} = \sum_{k \in S} P_{i,k} \cdot P_{k,j}$$

Observe que o produto $P_{i,k} \cdot P_{k,j}$, $k \in S$ é o resultado da multiplicação (produto interno) da i -ésima linha da matriz M com a j -ésima coluna na matriz M , e portanto é o elemento $P_{i,j}^{(2)}$ da matriz M^2 . O resultado geral segue de indução sobre n e demonstração análoga a que foi desenvolvida. \square

Agora, iremos buscar entender o comportamento em longo prazo da cadeia de Markov quando ela começa em uma distribuição de probabilidade dos estados de espaço.

Definição 2.16 (Vetor de probabilidades). Um vetor-linha de probabilidades $u = (u_0, \dots, u_N) \in \mathbb{R}^{N+1}$ é um vetor que satisfaz $u_i \geq 0$ para todo $i = 0, \dots, N$ e $\sum_{i=0}^N u_i = 1$.

Vamos considerar um vetor de probabilidades u para representar o estado inicial da cadeia de Markov. A ideia é que a entrada u_i do vetor u indica a probabilidade da cadeia de Markov iniciar no estado i .

Teorema 2.9. Sejam M matriz de Markov e u vetor de probabilidades que representam a distribuição inicial. Então a probabilidade da cadeia estar no estado i depois de n passos de tempo é indicado pela i -ésima coordenada do vetor

$$u^{(n)} = uM^n$$

Demonstração. Pela Lei da Probabilidade Total,

$$\mathbf{P}(X_n = i) = \sum_{k \in S} \mathbf{P}(X_n = i | X_0 = k) \cdot \mathbf{P}(X_0 = k)$$

Pelo Teorema (2.8) e pela definição do vetor u , temos que:

$$\mathbf{P}(X_n = i) = \sum_{k \in S} P_{k,i}^{(n)} u_k$$

E portanto $u_i^{(n)} = \mathbf{P}(X_n = i)$ pode ser obtido pelo produto interno do vetor u com a i -ésima coluna da matriz M^n . Deste modo, como i é arbitrário, então $u^{(n)} = uM^n$. \square

2.3 Grafos

Definição 2.17 (Grafos). Um grafo G consiste em um conjunto finito de vértices $V(G) = \{1, \dots, n\}$ e um conjunto finito de arestas $E(G) = \{e_1, \dots, e_m\}$ consistindo de pares distintos e não-ordenados de vértices em $V(G)$. A notação $e_k = (i, j)$ indica que a aresta conecta os vértices $i, j \in V(G)$. Neste caso, dizemos que os vértices i e j são adjacentes.

Definição 2.18 (Grafos orientados). Seja G um grafo com $V(G) = \{1, \dots, n\}$ e $E(G) = \{e_1, \dots, e_m\}$. Indicamos que um grafo é orientado se assinalamos que cada aresta de G possui uma orientação fixada, isto é, a aresta $e_k = (i, j)$ é distinta da aresta $e_l = (j, i)$ para todo par de vértices de G . Uma orientação é uma indicação de direção de cada aresta, tornando o grafo inicial em um grafo orientado.

Definição 2.19 (Matriz associada). A matriz associada ao grafo direcionado G , em que $V(G) = \{1, \dots, n\}$, é uma matriz $M = M(G) \in \mathbb{R}^{n \times n}$ onde um vértice i é adjacente a um vértice j se e somente se $a_{i,j} \neq 0$.

Definição 2.20 (Matriz de adjacência). Uma matriz de adjacência relativa a um grafo é uma matriz quadrada cujas entradas da matriz indicam se os pares de vértices são (ou não são) adjacentes no grafo. Dado um grafo G com conjunto de vértices $V(G) = \{1, \dots, n\}$, a matriz de adjacência relativa a G é $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$ satisfazendo $a_{i,j} = 1$ se e somente se existe uma aresta no grafo G saindo do vértice i e chegando no vértice j .

Definição 2.21 (Caminhos). Um caminho do grafo é um subgrafo no qual dois vértices são adjacentes, tem dois vértices terminais (ou vértices de grau 1) e tem outros vértices (se existirem) com grau 2. Equivalentemente, é possível listar os vértices em ordem v_1, \dots, v_m tais que os vértices satisfazem (v_i, v_{i+1}) com $i = 1, \dots, m - 1$. Um grafo é conectado se para cada dois pares de vértices i e j , existe um caminho direcionado de i para j . Um grafo é dito fortemente conectado se, para cada par de vértices i e j , existem os caminhos orientados de i para j e também de j para i .

Teorema 2.10. Seja $M \in \mathbb{R}^{n \times n}$ matriz não-negativa. Então M é irredutível se e somente se o grafo associado é fortemente conectado.

Demonstração. Sejam G grafo e $M = M(G)$ matriz não-negativa associada ao grafo. Vemos que, se M é redutível, então existe matriz de permutação $P \in \mathbb{R}^{n \times n}$ tal que

$$P^T M P = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

Onde A, B e C são matrizes bloco quadradas e 0 é matriz nula. Nesse caso, os vértices $\{1, \dots, p\}$ em $P^T M P$ correspondem as primeiras p linhas de $P^T M P$ e os $n - p$ vértices restantes correspondem as linhas restantes. Da estrutura de $P^T M P$, vemos que não existem arestas indo de $\{p+1, \dots, n\}$ para $\{1, \dots, p\}$. Portanto o grafo não é fortemente conectado, já que não existem arestas de $\{p+1, \dots, n\}$ para $\{1, \dots, p\}$.

Portanto, resta mostrar que uma matriz M é redutível se, e somente se, existe permutação de linhas e colunas tais que:

$$P^T M P = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

De fato, se M é redutível, então M tem entradas nulas. Então existe matriz de permutação de linhas e colunas P tal que:

$$P^T M P = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

E podemos obter uma decomposição de índices do grafo $V(G) = V_1 \cup V_2$, com $V_1 \cap V_2 = \emptyset$. Supondo que o conjunto de vértices $V(G)$ é fortemente conectado, temos um absurdo por não existe caminho de $i \in V_1$ para $j \in V_2$. \square

3 Processo de Moran Simples

Nessa seção, iremos aplicar os conceitos de processos estocásticos e cadeias de Markov que foram desenvolvidos para definir um modelo simples de Processo de Moran⁵. Nesse processo, dois tipos de indivíduos (duas variações genéticas de um mesmo ser vivo, por exemplo) disputam pela dominância na população.

O autor destaca que grande parte do conteúdo desse capítulo é baseado no trabalho de Chalub & Souza [2].

3.1 Modelo do Processo de Moran

Considere uma população de N indivíduos. Os indivíduos se dividem em dois tipos: \mathbb{A} e \mathbb{B} (também chamados de alelos⁶). Em cada geração (passo de tempo), um indivíduo é escolhido aleatoriamente para reproduzir (fazendo uma cópia de si mesmo) e um indivíduo é escolhido aleatoriamente para ser eliminado (substituído pelo descendente gerado), garantindo que o tamanho da população permaneça constante. É possível que o mesmo indivíduo seja escolhido para reproduzir e morrer no mesmo passo de tempo. Neste caso, o indivíduo é substituído pelo seu próprio descendente.

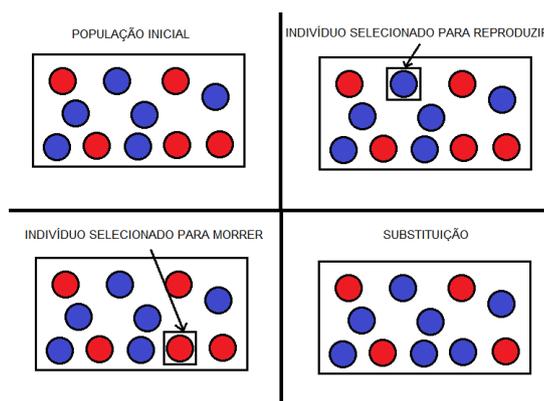


Figura 1: Exemplo de uma geração em um Processo de Moran com população de 11 indivíduos.

⁵Em referência ao geneticista australiano P.A.P. Moran que inventou o modelo no artigo *Random processes in genetics* (1958).

⁶Alelos são mutações de um mesmo gene, responsáveis pelas variações genéticas dos indivíduos.

É possível que no modelo um alelo tenha vantagem ou desvantagem reprodutiva em relação ao outro. Para isso, definiremos um parâmetro positivo, denominada *fitness*, associada a cada um dos alelos. Os indivíduos do alelo com a maior *fitness* possuem uma tendência maior a serem escolhidos para reproduzir. Além disso, iremos considerar que a reprodução ocorre sem mutações, e portanto um indivíduo com alelo \mathbb{A} só pode gerar descendentes com alelo \mathbb{A} e um indivíduo com alelo \mathbb{B} só pode gerar descendentes com alelo \mathbb{B} . Este modelo é conhecido como **Processo de Moran (simples)**.

O Processo de Moran é uma cadeia de Markov discreta em tempo discreto e pode ser definido pelos estados de espaço $i = 0, 1, \dots, N$ que representam o número de indivíduos com alelo \mathbb{A} . Para isso, definimos as variáveis aleatórias $X(t) := X_t$ como a quantidade de indivíduos com alelo \mathbb{A} na t -ésima geração (e portanto existem $N - i$ indivíduos com alelo \mathbb{B}). Na geração seguinte, a transição do estado de espaço só pode ocorrer para os estados $i - 1, i$ ou $i + 1$.

No estado 0, a população consiste apenas de indivíduos \mathbb{B} ⁷. Deste modo, no próximo passo de tempo, o indivíduo a ser escolhido para reproduzir e o indivíduo escolhido para morrer possuem o mesmo alelo e portanto o estado de espaço permanece em 0 em todas as gerações seguintes. O mesmo ocorre de maneira análoga no estado N , no qual só existem indivíduos \mathbb{A} . Os estados 0 e N são então classificados como estados de absorção do processo, e são os únicos estados com tal propriedade. Os estados $i = 1, \dots, N - 1$ são então classificados como estados de transição.

3.2 Probabilidades de seleção de tipo e transição de estado

Defina $\mathbf{p} = (p_0, \dots, p_N) \in [0, 1]^{N+1}$ vetor tal que p_i indica a probabilidade de um indivíduo \mathbb{A} ser escolhido para reproduzir dado que o Processo de Moran está no estado i . Consideraremos também que \mathbf{p} depende do estado de espaço atual (número de indivíduos \mathbb{A}), mas não depende do passo de tempo em que a cadeia se encontra⁸. O vetor \mathbf{p} é chamado de vetor de probabilidade de seleção de tipo (PST). Como 0 e N são estados de absorção, temos que $p_0 = p_N = 0$ e, como $1, \dots, N - 1$ são estados de

⁷Um indivíduo \mathbb{B} é um indivíduo que possui o alelo \mathbb{B} . Análogo para indivíduo \mathbb{A} .

⁸ \mathbf{p} tem a propriedade de Markov ou propriedade de perda de memória.

transição, temos que $0 < p_i < 1, i = 1, \dots, N - 1$ para todo PST. A probabilidade de um indivíduo \mathbb{B} ser escolhido para reproduzir é $1 - p_i$.

Para o Processo de Moran simples, assumiremos também que, no estado i , os indivíduos \mathbb{A} tem probabilidade $\frac{i}{N}$ de serem escolhidos para morrerem e os indivíduos \mathbb{B} tem probabilidade $\frac{N-i}{N}$. Isso significa que cada indivíduo, independente do alelo que possui, tem a mesma probabilidade $\frac{1}{N}$ de ser escolhido para morrer.

Seja $P_{i,j}$ a probabilidade de transição do estado i para o estado j em um passo de tempo. O número esperado de indivíduos de indivíduos na próxima geração, dado que o estado de espaço é i , é $\mathbb{E}[X_{t+1}|X_t = i] = \sum_{j=0}^N i P_{i,j}$.

Como 0 e N são estados de absorção, então:

$$P_{0,0} = 1; P_{0,j} = 0, j = 1, \dots, N$$

$$P_{N,N} = 1; P_{N,j} = 0, j = 0, \dots, N - 1$$

Seja i um estado de transição. Logo existem pelo menos um indivíduo \mathbb{A} e um indivíduo \mathbb{B} , e portanto existem quatro possibilidades que podem ocorrer na próxima geração:

- (a) Um indivíduo \mathbb{A} é escolhido para reproduzir e morrer (não necessariamente o mesmo). O estado de espaço continua em i na geração seguinte. Este evento ocorre com probabilidade $p_i \frac{i}{N}$.
- (b) Um indivíduo \mathbb{B} é escolhido para reproduzir e morrer (não necessariamente o mesmo). O estado de espaço continua em i na geração seguinte. Este evento ocorre com probabilidade $(1 - p_i) \left(\frac{N-i}{N}\right)$.
- (c) Um indivíduo \mathbb{A} é escolhido para reproduzir e um indivíduo \mathbb{B} é escolhido para morrer. O estado de espaço é $i + 1$ na geração seguinte. Este evento ocorre com probabilidade $p_i \left(\frac{N-i}{N}\right)$.
- (d) Um indivíduo \mathbb{B} é escolhido para reproduzir e um indivíduo \mathbb{A} é escolhido para morrer. O estado de espaço é $i - 1$ na geração seguinte. Este evento ocorre com probabilidade $(1 - p_i) \frac{i}{N}$.

Portanto, temos que:

$$P_{i,j} = \begin{cases} (1 - p_i) \frac{i}{N}, & j = i - 1 \\ p_i \frac{i}{N} + (1 - p_i) \left(\frac{N-i}{N}\right), & j = i \\ p_i \left(\frac{N-i}{N}\right), & j = i + 1 \\ 0, & |i - j| > 1 \end{cases} \quad (2)$$

3.3 *Fitness e neutralidade*

Agora, iremos introduzir o conceito de *fitness* no modelo com o objetivo de se determinar as probabilidades p_i . Segundo Orr [13], *fitness* é uma das teorias biológicas mais aceitas e envolve a capacidade de certos organismos de sobreviverem e reproduzirem, e, como consequência, contribuem com seus genes para a próxima geração. Em teoria de jogos evolutivos, como em Imhof & Nowak [6], os indivíduos \mathbb{A} e indivíduos \mathbb{B} são identificados como estratégia \mathbb{A} e estratégia \mathbb{B} e o conceito de *fitness* é entendido como o *payoff* (pagamento, recompensa) do jogo. Definiremos *fitness* (ou *fitness* reprodutiva) como um parâmetro positivo que, associado a cada alelo, determina se os indivíduos que possuem aquele alelo tem vantagem ou desvantagem reprodutiva em relação ao outro alelo e conseqüentemente, são mais prováveis ou não de serem escolhidos para reproduzir.

Sejam $\varphi^{(\mathbb{A})}, \varphi^{(\mathbb{B})} : \{0, 1, \dots, N\} \rightarrow \mathbb{R}_+^*$ funções positivas tais que $\varphi^{(\mathbb{A})}(i)$ e $\varphi^{(\mathbb{B})}(i)$ são, respectivamente, as *fitness* reprodutivas dos indivíduos \mathbb{A} e dos indivíduos \mathbb{B} no estado de espaço i (quando há i indivíduos \mathbb{A}). No Processo de Moran simples, assumiremos que todos os indivíduos do mesmo alelo possuem a mesma *fitness* reprodutiva e que as funções *fitness* reprodutivas têm a propriedade de Markov. A relação entre *fitness* reprodutiva e as probabilidades de seleção de tipo p_i pode ser definida como:

$$p_i = \frac{i\varphi^{(\mathbb{A})}(i)}{i\varphi^{(\mathbb{A})}(i) + (N - i)\varphi^{(\mathbb{B})}(i)} \quad (3)$$

Deste modo, as únicas variáveis que definem completamente os diferentes modelos de Processos de Moran (simples) são as funções $\varphi^{(\mathbb{A})}$ e $\varphi^{(\mathbb{B})}$. É também possível definir um Processo de Moran a partir das probabilidades p_i . Neste caso, veremos que só é possível determinar o valor da razão $\frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)}$.

Definição 3.1 (Neutro). Dizemos que o Processo de Moran é neutro se as variáveis aleatórias X_t satisfazem $\mathbb{E}[X_{t+1}|X_t = i] = i$ para todo $i = 0, \dots, N$. ■

Para os estados de absorção, temos que $P_{0,0} = P_{N,N} = 1$. Logo:

$$\begin{aligned}\mathbb{E}[X_{t+1}|X_t = 0] &= 0 \cdot P_{0,0} + \sum_{i=1}^N iP_{0,i} = 0 \\ \mathbb{E}[X_{t+1}|X_t = N] &= \sum_{i=0}^{N-1} iP_{N,i} + N \cdot P_{N,N} = 1 \cdot N = N\end{aligned}$$

Para os estados de transição, temos que $P_{i,j} = 0$ para $|i - j| > 1$. Logo:

$$\mathbb{E}[X_{t+1}|X_t = i] = (i - 1)P_{i,i-1} + iP_{i,i} + (i + 1)P_{i,i+1}$$

Observando que $P_{i,i-1} + P_{i,i} + P_{i,i+1} = 1$, segue que:

$$\begin{aligned}\mathbb{E}[X_{t+1}|X_t = i] &= (i - 1)P_{i,i-1} + i(1 - P_{i,i-1} - P_{i,i+1}) + (i + 1)P_{i,i+1} \\ &= iP_{i,i-1} - P_{i,i-1} + i - iP_{i,i-1} - iP_{i,i+1} + iP_{i,i+1} + P_{i,i+1} \\ &= -P_{i,i-1} + i + P_{i,i+1} \quad (2) \\ &= (p_i - 1)\frac{i}{N} + i + p_i \left(1 - \frac{i}{N}\right) \\ &= \frac{ip_i}{N} - \frac{i}{N} + i + p_i - \frac{ip_i}{N} \\ &= p_i + i - \frac{i}{N}\end{aligned} \tag{4}$$

Logo temos que a neutralidade do Processo de Moran está relacionado diretamente com PST (especificamente com as probabilidades de seleção dos estados de transição $i = 1, \dots, N - 1$). Para um entendimento mais completo sobre a noção de neutralidade e sua relação com PST, iremos introduzir uma outra definição de *fitness*, denotada *fitness* Darwiniana, que representa uma ideia de sobrevivência dos indivíduos de um alelo na geração seguinte.

Definição 3.2 (*Fitness* Darwiniana). Sejam X_t as variáveis aleatórias do Processo de Moran e i um estado de transição. Então as *fitness* Darwinianas dos alelos \mathbb{A} e \mathbb{B} são, respectivamente:

$$\Psi^{(\mathbb{A})}(i) = \frac{\mathbb{E}[X_{t+1}|X_t = i]}{i}, \quad \Psi^{(\mathbb{B})}(i) = \frac{(N - \mathbb{E}[X_{t+1}|X_t = i])}{N - i} \quad \blacksquare \quad (5)$$

Notamos também que as *fitness* reprodutivas ($\varphi^{(\mathbb{A})}$ e $\varphi^{(\mathbb{B})}$) representam uma noção comparativa entre os alelos, isto é, a razão $\frac{\varphi^{(\mathbb{A})}}{\varphi^{(\mathbb{B})}}$ é o valor que determina as probabilidades p_i . De fato, por (3), temos que:

$$p_i = \frac{i \frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)}}{i \frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)} + N - i} \quad (6)$$

Por outro lado, as *fitness* Darwinianas ($\Psi^{(\mathbb{A})}$ e $\Psi^{(\mathbb{B})}$) são valores com uma noção absoluta, sendo a razão de indivíduos de um alelo esperados na geração seguinte pela quantidade de indivíduos do mesmo alelo da geração atual. As *fitness* reprodutivas e Darwinianas se relacionam pelo seguinte lema:

Lema 3.1. Sejam $\varphi^{(\mathbb{A})}(i)$, $\varphi^{(\mathbb{B})}(i)$, $\Psi^{(\mathbb{A})}(i)$ e $\Psi^{(\mathbb{B})}(i)$ respectivamente as *fitness* reprodutivas e Darwinianas dos indivíduos \mathbb{A} e \mathbb{B} e i um estado de transição. Então:

$$\frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)} = \frac{\Psi^{(\mathbb{A})}(i)}{\Psi^{(\mathbb{B})}(i)} \iff \mathbb{E}[X_{t+1}|X_t = i] = N \cdot p_i$$

Quando ocorrem uma das igualdades (e conseqüentemente a equivalência) para todo estado de transição i , dizemos que o Processo de Moran é consistente com as *fitness* reprodutivas e Darwinianas.

Demonstração. Por (5), temos:

$$\frac{\Psi^{(\mathbb{A})}(i)}{\Psi^{(\mathbb{B})}(i)} = \frac{(N - i)}{i} \cdot \frac{\mathbb{E}[X_{t+1}|X_t = i]}{(N - \mathbb{E}[X_{t+1}|X_t = i])}$$

Por outro lado, de (6) segue que:

$$p_i \left(i \frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)} + N - i \right) = i \frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)}$$

$$\begin{aligned} \Rightarrow \frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)}(ip_i - i) &= -p_i(N - i) \\ \Rightarrow \frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)} &= \frac{(N - i)}{i} \cdot \frac{p_i}{(1 - p_i)} \end{aligned}$$

Por fim,

$$\frac{\frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)}}{\frac{\Psi^{(\mathbb{A})}(i)}{\Psi^{(\mathbb{B})}(i)}} = \frac{p_i}{1 - p_i} \cdot \frac{(N - \mathbb{E}[X_{t+1}|X_t = i])}{\mathbb{E}[X_{t+1}|X_t = i]} = \frac{N \cdot p_i - p_i \mathbb{E}[X_{t+1}|X_t = i]}{\mathbb{E}[X_{t+1}|X_t = i] - p_i \mathbb{E}[X_{t+1}|X_t = i]}$$

e portanto

$$\frac{\frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)}}{\frac{\Psi^{(\mathbb{A})}(i)}{\Psi^{(\mathbb{B})}(i)}} = 1 \iff N \cdot p_i = \mathbb{E}[X_{t+1}|X_t = i]$$

□

Observamos aqui que as razões $\frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)}$ e $\frac{\Psi^{(\mathbb{A})}(i)}{\Psi^{(\mathbb{B})}(i)}$ estão bem-definidas para todo estado de transição i . De fato, $\varphi^{(\mathbb{B})}(i) > 0$ por definição, e, por (4) e (5), $\Psi^{(\mathbb{B})}(i) = 0 \iff p_i = N - i + \frac{i}{N}$. Como $\frac{i}{N} > 0$ e $N - i \geq 1$, então teríamos que $p_i > 1$, uma contradição. Assim $\Psi^{(\mathbb{B})}(i) \neq 0$.

Teorema 3.2 (Neutralidade e *fitness*). As seguintes afirmações são equivalentes:

- (a) O Processo de Moran (simples) é neutro
- (b) $p_i = \frac{i}{N}$ para todo estado de transição i
- (c) O Processo de Moran (simples) é consistente com as *fitness* reprodutivas e Darwinianas

Demonstração. Iremos provar que (a) \Rightarrow (b), (b) \Rightarrow (c) e (c) \Rightarrow (a).

(a) \Rightarrow (b) Suponha que o Processo de Moran é neutro, e portanto, vale $\mathbb{E}[X_{t+1}|X_t = i] = i$ para todo estado de espaço. Pela equação (4), temos:

$$i = p_i + i - \frac{i}{N}$$

$$\therefore p_i = \frac{i}{N}$$

(b) \Rightarrow (c) Suponha que $p_i = \frac{i}{N}$ para todo estado de transição i . Pela equação (4), temos:

$$\mathbb{E}[X_{t+1}|X_t = i] = \frac{i}{N} + i - \frac{i}{N} = i = \frac{N \cdot i}{N} = N \cdot p_i$$

Pelo Lema (3.1), segue que o Processo de Moran é consistente com as *fitness* reprodutivas e Darwinianas

(c) \Rightarrow (a) Suponha que $\mathbb{E}[X_{t+1}|X_t = i] = N \cdot p_i$ para todo espaço de transição i .

Já mostramos que $\mathbb{E}[X_{t+1}|X_t = 0] = 0$ e $\mathbb{E}[X_{t+1}|X_t = N] = N$, e, logo, resta mostrar que a igualdade $\mathbb{E}[X_{t+1}|X_t = i] = i$ é válida para todo estado de transição i . Pela equação (4), temos:

$$N \cdot p_i = p_i + i - \frac{i}{N}$$

$$N^2 \cdot p_i = N \cdot p_i + Ni - i$$

$$p_i(N^2 - N) = Ni - i$$

Podemos assumir que N é um inteiro maior que 1⁹. Neste caso:

$$p_i = \frac{i(N-1)}{N(N-1)} = \frac{i}{N}$$

Pela hipótese inicial,

$$\mathbb{E}[X_{t+1}|X_t = i] = N \cdot p_i = i$$

⁹Se $N = 1$, então o Processo de Moran consiste de dois estados estacionários: 0 e 1. Neste caso, não há estados de transição.

Deste modo, provamos o enunciado. □

O teorema acima nos mostra que a neutralidade do Processo de Moran (simples) só acontece quando temos que $p_i = \frac{i}{N}, i = 1, \dots, N - 1$. Essa probabilidade coincide com a probabilidade de se escolher um indivíduo \mathbb{A} aleatoriamente (dentre N indivíduos totais), o que significa que cada indivíduo (independente se é um indivíduo \mathbb{A} ou \mathbb{B}) tem probabilidade $\frac{1}{N}$ de ser escolhido para reproduzir. Isso nos dá uma ideia intuitiva de que nenhum dos alelos tem vantagem ou desvantagem evolutiva sobre o outro (mesma *fitness* reprodutiva). Essa noção pode ser entendida pelo corolário abaixo.

Corolário 3.2.1. O Processo de Moran é neutro se, e somente se, $\varphi^{(\mathbb{A})}(i) = \varphi^{(\mathbb{B})}(i), i = 1, \dots, N - 1$.

Demonstração. Da demonstração do Lema 3.1, temos:

$$\frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)} = \frac{(N - i)}{i} \cdot \frac{p_i}{(1 - p_i)}$$

Suponha que o Processo de Moran é neutro. Daí, pelo Teorema 3.3:

$$\frac{\varphi^{(\mathbb{A})}(i)}{\varphi^{(\mathbb{B})}(i)} = \frac{(N - i)}{i} \cdot \frac{\frac{i}{N}}{\frac{N-i}{N}} = 1$$

Por outro lado, suponha agora que $\varphi^{(\mathbb{A})}(i) = \varphi^{(\mathbb{B})}(i), i = 1, \dots, N - 1$. Por (6), temos:

$$p_i = \frac{i}{i + N - i} = \frac{i}{N}$$

E o enunciado segue observando novamente o Teorema 3.3. □

3.4 Matriz de Markov para o Processo de Moran simples

Vemos que, por (6), podemos assumir que $\varphi^{(\mathbb{A})}(i) = r_i$ e $\varphi^{(\mathbb{B})}(i) = 1$ para todo $i = 0, \dots, N$. Neste caso, temos que:

$$p_i = \frac{ir_i}{ir_i + N - i} \quad (7)$$

Com isso, temos que, por (2):

$$P_{0,0} = 1$$

$$P_{1,1} = 1$$

$$P_{i,i-1} = \frac{N-i}{ir_i + N - i} \cdot \frac{i}{N} \quad (8)$$

$$P_{i,i} = 1 - P_{i,i-1} - P_{i,i+1} \quad (9)$$

$$P_{i,i+1} = \frac{N-i}{N} \cdot \frac{ir_i}{ir_i + N - i} \quad (10)$$

Notamos agora que as probabilidades dos estados de transição $P_{i,i-1}$, $P_{i,i}$ e $P_{i,i+1}$ dependem do estado i e da *fitness* r_i (que também só varia com i). Deste modo, chamando $P_{i,i-1} = \beta_i$ e $P_{i,i+1} = \alpha_i$, vemos que a matriz de Markov M associada ao Processo de Moran Simples (que é uma cadeia de Markov) pode ser escrita como:

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ \beta_1 & 1 - \beta_1 - \alpha_1 & \alpha_1 & 0 & \dots & 0 \\ 0 & \beta_2 & 1 - \beta_2 - \alpha_2 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix} \quad (11)$$

3.5 Vetor de probabilidades de fixação

Definição 3.3 (Extinção e fixação). Sejam 0 e N os (únicos) estados de absorção de um Processo de Moran. Dizemos que o estado 0 é estado de *extinção* do alelo \mathbb{A} e o estado N é o estado de *fixação* do alelo \mathbb{A} .

O teorema abaixo caracteriza o vetor de probabilidade de fixação de uma maneira algébrica.

Teorema 3.3. Seja M matriz de Markov associada ao processo de Moran simples. Então existe um único vetor $\tilde{\mathbf{F}} = (F_1, \dots, F_{N-1}) \in \mathbb{R}^{N-1}, 0 < F_i < 1$ tais que $\mathbf{F} = (0, \tilde{\mathbf{F}}, 1), M\mathbf{F}^T = \mathbf{F}^T$ e

$$\tilde{\mathbf{F}}^T = (I - M)^{-1}(0, \dots, 0, \alpha_{N-1})^T$$

Antes de iniciarmos a demonstração do teorema acima, vamos apresentar alguns resultados necessários para a mesma. Primeiramente, vemos que a igualdade:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ \beta_1 & 1 - \beta_1 - \alpha_1 & \alpha_1 & 0 & \dots & 0 \\ 0 & \beta_2 & 1 - \beta_2 - \alpha_2 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} 0 \\ F_1 \\ F_2 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ F_1 \\ F_2 \\ \vdots \\ 1 \end{bmatrix}$$

é equivalente à

$$\begin{bmatrix} 1 - \beta_1 - \alpha_1 & \alpha_1 & 0 & \dots & 0 \\ \beta_2 & 1 - \beta_2 - \alpha_2 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \beta_{N-1} & 1 - \beta_{N-1} - \alpha_{N-1} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_{N-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \alpha_{N-1} \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_{N-1} \end{bmatrix}$$

Observamos aqui que a sub-matriz principal \tilde{M} , definida como

$$\tilde{M} = \begin{bmatrix} 1 - \beta_1 - \alpha_1 & \alpha_1 & 0 & \dots & 0 \\ \beta_2 & 1 - \beta_2 - \alpha_2 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \beta_{N-1} & 1 - \beta_{N-1} - \alpha_{N-1} \end{bmatrix} \quad (12)$$

é sub-estocástica, não-negativa e irredutível.

Demonstração do Teorema 2.3. Inicialmente, vimos que a igualdade $M\mathbf{F}^T = \mathbf{F}^T$ é equivalente a $\tilde{M}\tilde{\mathbf{F}}^T + (0, \dots, 0, \alpha_{N-1})^T = \tilde{\mathbf{F}}^T$, em que \tilde{M} é matriz sub-estocástica, não-negativa e irredutível.

Considere a matriz $\tilde{M}' = \tilde{M} + (\beta_1, 0, \dots, \alpha_{N-1})^T$. Como o vetor $(\beta_1, 0, \dots, \alpha_{N-1})$ é não-negativo, então a matriz \tilde{M}' é não-negativa e irredutível. Além disso, \tilde{M}' é estocástica (todas as linhas tem soma 1) e $\tilde{M} < \tilde{M}'$. Pelo Teorema de Perron-Frobenius aplicado a \tilde{M}' , segue que

$$1 = \min_i \sum_j m'_{ij} \leq \rho(\tilde{M}') \leq \max_i \sum_j m'_{ij} = 1 \Rightarrow \rho(\tilde{M}') = 1$$

Além disso, pelo Corolário B.3.1 do Apêndice B, temos que:

$$0 \leq \tilde{M} < \tilde{M}' \Rightarrow \rho(\tilde{M}) < \rho(\tilde{M}') = 1$$

E portanto $\rho(\tilde{M}) < 1$. Pelo Lema de Neumann, segue imediatamente que $\sum_{k=0}^{\infty} \tilde{M}^k$ converge. Afirmamos que a matriz $I - \tilde{M}$ é invertível e o vetor $(I - \tilde{M})^{-1} (0, \dots, 0, \alpha_{N-1})^T$ tem todas as coordenadas positivas. De fato:

$$\lim_N (I - \tilde{M}) \left(\sum_{k=0}^N \tilde{M}^k \right) = \lim_N (I - \tilde{M}^{N+1}) = I$$

Portanto $I - \tilde{M}$ é invertível e

$$(I - \tilde{M})^{-1} = \sum_{k=0}^{\infty} \tilde{M}^k$$

Além disso, como \tilde{M} é irredutível, temos que $\sum_{k=0}^{\infty} \tilde{M}^k$ é matriz positiva e, por fim,

$$(I - \tilde{M})\tilde{\mathbf{F}}^T = (0, \dots, 0, \alpha_{N-1})^T$$

$$\Rightarrow \tilde{\mathbf{F}}^T = \left(\sum_{k=0}^{\infty} M^k \right) \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \alpha_{N-1} \end{bmatrix}$$

E segue que $F_i > 0$.

Por outro lado, temos que:

$$\begin{bmatrix} 1 - \beta_1 - \alpha_1 & \alpha_1 & 0 & \dots & 0 \\ \beta_2 & 1 - \beta_2 - \alpha_2 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \beta_{N-1} & 1 - \beta_{N-1} - \alpha_{N-1} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \begin{bmatrix} \beta_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Escrevendo $\mathbf{1} = (1, 1, \dots, 1)^T$, então vemos que:

$$\tilde{M}(\mathbf{1} - \tilde{\mathbf{F}})^T = (\mathbf{1} - \tilde{\mathbf{F}})^T - (\beta_1, 0, \dots, 0)^T$$

E portanto:

$$(I - \tilde{M})(\mathbf{1} - \tilde{\mathbf{F}})^T = (\beta_1, 0, \dots, 0)^T$$

$$\Rightarrow (\mathbf{1} - \tilde{\mathbf{F}})^T = \left(\sum_{k=0}^{\infty} M^k \right) \begin{bmatrix} \beta_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

E, com o argumento análogo ao anterior, segue que as entradas do vetor $(\mathbf{1} - \tilde{\mathbf{F}})^T$ são positivas. Isso significa que $F_i < 1$ e então concluímos o enunciado. \square

Vamos encontrar explicitamente o vetor de probabilidades de fixação \mathbf{F} . Da igualdade $M\mathbf{F}^T = \mathbf{F}^T$, obtemos a seguinte relação:

$$F_i = \beta_i F_{i-1} + (1 - \alpha_i - \beta_i) F_i + \alpha_i F_{i+1}, 1 \leq i \leq N - 1$$

$$\Rightarrow \beta_i(F_{i-1} - F_i) + \alpha_i(F_{i+1} - F_i) = 0 \quad (13)$$

Denotando $F_0 = 0, F_N = 1$ e introduzindo as notações $x_i = F_i - F_{i-1}, i = 1, \dots, N$ e $\gamma_i = \frac{\beta_i}{\alpha_i}, i = 1, \dots, N - 1$, obtemos a relação recursiva:

$$x_{i+1} = \gamma_i x_i$$

Deste modo, vemos que:

$$x_1 = F_1 - F_0 = F_1$$

$$x_2 = \gamma_1 x_1 = \gamma_1 F_1$$

E indutivamente vemos que:

$$x_k = F_k - F_{k-1} = F_1 \left(\prod_{i=1}^{k-1} \gamma_i \right), k = 2, 3, \dots, N - 1$$

E

$$x_N = F_N - F_{N-1} = 1 - F_1 \left(\prod_{i=1}^{N-1} \gamma_i \right)$$

Por outro lado, pela soma telescópica,

$$\sum_{k=1}^N x_k = \sum_{k=1}^N (F_k - F_{k-1}) = F_N - F_0 = 1$$

E, combinando as equações, segue que:

$$1 = \sum_{k=1}^N x_k = \sum_{k=1}^N \left(F_1 \prod_{i=1}^{k-1} \gamma_i \right) = F_1 \left[1 + \left(\sum_{k=1}^{N-1} \prod_{i=1}^{k-1} \gamma_i \right) \right]$$

Resolvendo para F_1 , obtemos:

$$F_1 = \frac{1}{1 + \sum_{k=1}^{N-1} \prod_{i=1}^{k-1} \gamma_i}$$

Que representa a probabilidade de fixação a partir de um único indivíduo do tipo \mathbb{A} . Para um estado de transição j qualquer, vemos que:

$$\begin{aligned} F_j &= \sum_{k=1}^j x_k = \sum_{k=1}^j \left(F_1 \prod_{i=1}^{k-1} \gamma_i \right) = F_1 \prod_{i=1}^0 \gamma_i + \sum_{k=2}^j \left(F_1 \prod_{i=1}^{k-1} \gamma_i \right) = F_1 + \sum_{k=2}^j \left(F_1 \prod_{i=1}^{k-1} \gamma_i \right) \\ &= F_1 \left[1 + \sum_{k=2}^j \left(\prod_{i=1}^{k-1} \gamma_i \right) \right] = F_1 \left(1 + \sum_{k=1}^{j-1} \prod_{i=1}^k \gamma_i \right) = \frac{1 + \sum_{k=1}^{j-1} \prod_{i=1}^k \gamma_i}{1 + \sum_{k=1}^{N-1} \prod_{i=1}^{k-1} \gamma_i} \end{aligned}$$

Assumindo que os indivíduos do tipo \mathbb{A} tem fitness reprodutiva constante r e os indivíduos do tipo \mathbb{B} tem fitness reprodutiva constante 1, então segue que $\gamma_i = \frac{1}{r}$ é um fator constante no produtório e temos que:

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

Então, a probabilidade de fixação quando a população inicial é composta por apenas um único indivíduo do tipo \mathbb{A} é:

$$\rho := F_1 = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^N}}$$

4 Grafos Evolutivos

Até aqui, desenvolvemos a teoria do Processo de Markov simples em populações homogêneas, em que todos os indivíduos da população estão em posições equivalentes. Nessa seção, iremos analisar alguns efeitos e consequências de estruturas populacionais nessa dinâmica evolutiva. Faremos isso através de uma associação entre os indivíduos e os vértices de um grafo. As arestas do grafos, neste caso, representariam interações de competição, de modo que, se existe uma aresta do vértice i para o vértice j , existe uma probabilidade (positiva) do descendente do indivíduo do vértice i substituir o indivíduo do vértice j .

O autor destaca que o conteúdo desse capítulo é inspirado no trabalho de Nowak ([12]).

4.1 Estrutura Populacional

Inicialmente, identificamos os indivíduos da população com $i = 1, 2, \dots, N$. Em cada passo de tempo, um indivíduo é escolhido aleatoriamente para se reproduzir. A probabilidade do descendente do indivíduo i substituir o indivíduo j é denotado por $w_{i,j}$, e portanto, o processo pode ser determinado através de uma matriz $N \times N$, $W = [w_{i,j}]$. Como as entradas de W são probabilidades, então W é matriz linha-estocástica.

Podemos ainda identificar os indivíduos da população com vértices de um grafo. Se $w_{i,j} > 0$, então existe uma aresta orientada do vértice i para o vértice j . Se $w_{i,j} = 0$, então não existe aresta que liguem estes dois vértices. Portanto, a matriz W define um grafo orientado.

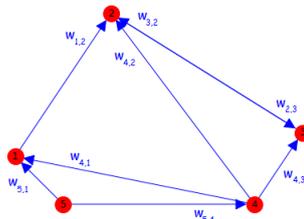


Figura 2: Exemplo de estrutura populacional G com população $N = 5$

$$W = \begin{bmatrix} 0 & w_{1,2} & 0 & 0 & 0 \\ 0 & 0 & w_{2,3} & 0 & 0 \\ 0 & w_{3,2} & 0 & 0 & 0 \\ w_{4,1} & w_{4,2} & w_{4,3} & 0 & 0 \\ w_{5,1} & 0 & 0 & w_{5,4} & 0 \end{bmatrix}; W \text{ é a matriz de adjacência relativa ao grafo } G$$

A estrutura acima é uma generalização do processo de Moran simples, em que o caso homogêneo é recuperado em um grafo completo (em que, dado dois vértices i e j quaisquer, existe as arestas que ligam i para j e que ligam j para i) com mesmo peso $w_{i,j}$. Se a população tem N indivíduos, então $w_{i,j} = \frac{1}{N}$ para todos i, j .

4.2 Probabilidades de fixação

Nessa sessão, iremos analisar diferentes estruturas de grafos e o comportamento das probabilidades de fixação do processo. Observa-se que a estrutura do grafo está relacionada diretamente com a matriz W das probabilidades de substituição.

O objetivo é determinar a probabilidade de fixação ao surgir apenas um indivíduo de gene mutante (1 indivíduo do tipo \mathbb{A} e $N - 1$ indivíduos do tipo \mathbb{B} e verificar quais das estruturas possuem comportamento de fixação idêntico ao Processo de Moran descrito no capítulo anterior.

4.2.1 O Ciclo direcionado

Considere um grafo simples assimétrico de N vértices, denotado de ciclo direcionado. Para cada vértice i , a única aresta conectada neste vértice liga até $i + 1$. O vértice N liga-se ao vértice 1 e então obtém-se estrutura semelhante a um ciclo.

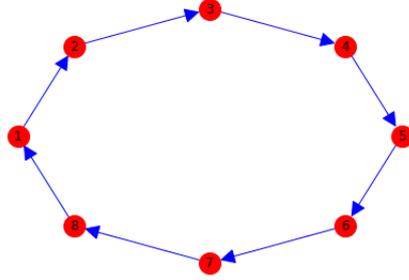


Figura 3: Exemplo de um ciclo direcionado com população $N = 8$

$$W = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}; W \text{ é forma geral da matriz de adjacência do ciclo direcionado}$$

Vamos supor inicialmente que toda a população consista de indivíduos do tipo \mathbb{B} e que, em algum momento, ocorre uma mutação e é gerado um indivíduo do tipo \mathbb{A} com *fitness* relativa r .

Definiremos um **caminho de indivíduos do tipo** como um caminho no grafo nos quais os indivíduos dos vértices do caminho são do mesmo tipo, de modo que os vértices iniciais e finais estão conectados com indivíduos de um tipo diferente. Nessa definição, entendemos que os caminhos de indivíduos de tipo são os "maiores" possíveis (de maiores comprimentos) e não conseguem ser estendidos (ou continuados). Diremos que dois caminhos de indivíduos de tipo são distintos se eles possuem pelo menos um vértice diferente - e conseqüentemente todos os vértices dos caminhos são distintos. A ideia geral por trás dessa definição é construir um análogo a definição de componente conexa no grafo, mas atuando nas colorações \mathbb{A} e \mathbb{B} .

Começando com um único indivíduo do tipo \mathbb{A} , somente um caminho de indivíduos do tipo \mathbb{A} poderá surgir (a saber, a partir da orientação do grafo). Por exemplo, na Figura 3, se o indivíduo mutante surge no vértice 4, então os possíveis caminhos de indivíduos do tipo \mathbb{A} são $\{4\}$ na primeira geração e $\{4, 5\}, \{4, 5, 6\}, \dots$ nas gerações

seguintes. Mais ainda, é possível que em uma certa geração o caminho de indivíduos de tipo seja $\{6, 7, 8, 1\}$, por exemplo. Para isso, basta os indivíduos 3 e 4 sejam escolhidos para reproduzir, nesta ordem (mas não necessariamente em gerações consecutivas). Nessa configuração, não é possível que se formem dois ou mais caminhos distintos de indivíduos do tipo \mathbb{A} .

Seja i o espaço de estado do processo estocástico correspondente. Vemos que, para o estado de espaço no passo de tempo seguinte seja $i - 1$, é necessário de um indivíduo do tipo \mathbb{B} seja escolhido para reproduzir, especificamente o indivíduo imediatamente anterior ao caminho de indivíduos do tipo \mathbb{A} . Como existem $N - i$ indivíduos do tipo \mathbb{B} e que esses indivíduos tem a mesma probabilidade, escrevendo $P_{i,i-1}$ como a probabilidade de transição do estado i para o estado $i - 1$, temos que:

$$P_{i,i-1} = \frac{N - i}{i \cdot r + (N - i) \cdot 1} \cdot \frac{1}{N - i} = \frac{1}{ir + N - i}$$

Por outro lado, para o estado de espaço na geração seguinte ser de $i + 1$, é necessário que um indivíduo do tipo \mathbb{A} seja escolhido para reproduzir, especificamente o indivíduo que está no final do caminho. Admitindo que os indivíduos do tipo \mathbb{A} tem mesma probabilidade de serem escolhidos para reproduzir, temos que, escrevendo $P_{i,i+1}$ como a probabilidade de transição do estado de espaço i para $i + 1$, temos que:

$$P_{i,i+1} = \frac{i \cdot r}{i \cdot r + (N - i) \cdot 1} \cdot \frac{1}{i} = \frac{r}{ir + N - i}$$

A razão entre essas duas probabilidades é dada por:

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

Vemos que essa quantidade não depende do estado de espaço e é idêntica aquela obtida no processo de Moran com seleção constante. De modo análogo, pelas equações (13), obtemos que:

$$\rho = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^N}}$$

E portanto a probabilidade de fixação no grafo simples assimétrico coincide com a probabilidade de fixação no Processo de Moran simples.

4.2.2 O Ciclo Bidirecionado

Considere um grafo simples simétrico de N vértices, denotado de ciclo. Neste grafo, para cada dois vértices i e $i + 1$, existem duas arestas orientadas entre elas, a saber, a aresta $(i, i + 1)$ e a aresta $(i + 1, i)$. Também conecta-se os vértices 1 e N de maneira simétrica. Para esse estudo, vamos supor que todas as arestas tenham o mesmo peso. O grafo obtido dessa construção assemelhasse ao ciclo direcionado, porém nas duas direções.

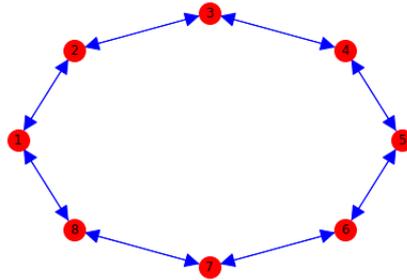


Figura 4: Exemplo de um ciclo bidirecionado com população $N = 8$

$$W = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \dots & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \dots & \frac{1}{2} & 0 \end{bmatrix} \quad W \text{ é forma geral da matriz de adjacência do ciclo bidirecionado}$$

Como anteriormente, começando com apenas um único indivíduo mutante do tipo \mathbb{A} , só poderá haver um único caminho de indivíduos do tipo \mathbb{A} . Com os mesmos argumentos do caso anterior, vemos que:

$$P_{i,i-1} = \frac{1}{ir + N - i} \quad \text{e} \quad P_{i,i+1} = \frac{r}{ir + N - i}$$

E obtemos a mesma probabilidade de fixação do caso anterior.

4.2.3 A Linha e o *Burst*

Um grafo orientado é denominado linha se, do vértice i , a única possibilidade de colocar o descendente é no vértice $i + 1$ e, no vértice N , o descendente substitui a si mesmo. Não há arestas que cheguem no vértice 1.



Figura 5: Exemplo de uma linha com população $N = 5$

$$W = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

A probabilidade de fixação, portanto, não depende da *fitness* dos indivíduos do tipo \mathbb{A} , e sim da posição em que está localizado o único indivíduo do tipo \mathbb{A} . Deste fato,

nessa configuração de grafo, o tipo do indivíduo que está no vértice 1 é o que será fixado. Assumindo que o indivíduo do tipo \mathbb{A} surja em um vértice aleatoriamente, a probabilidade de fixação portanto é:

$$\rho = \frac{1}{N}$$

E, com probabilidade $\frac{(N-1)}{N}$, o indivíduo do tipo \mathbb{A} surge nos vértices $2, 3, \dots, N$ e eventualmente é extinto. Assim, as probabilidades de fixação no grafo linha diferem-se das probabilidades de fixação do Processo de Moran.

Agora considere o grafo definido pela matriz W abaixo, no qual um vértice central se conecta unidirecionalmente com os outros $N - 1$ vértices periféricos.

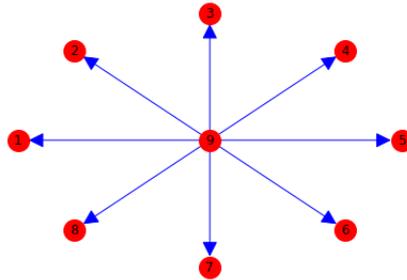


Figura 6: Exemplo de um *burst* com população $N = 9$

$$W = \begin{bmatrix} 0 & \frac{1}{N-1} & \frac{1}{N-1} & \cdots & \frac{1}{N-1} & \frac{1}{N-1} \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \quad W \text{ é forma geral da matriz de adjacência do } burst$$

O grafo induzido por W é denotado *burst*. A fixação de um único indivíduo do tipo \mathbb{A} ocorre apenas no caso em que ele surge no centro, com probabilidade $\frac{1}{N}$. Novamente, a probabilidade de fixação não depende da *fitness* e coincide com a probabilidade do grafo

linha. Esses dois grafos são exemplos de estruturas que suprimem a seleção natural, dado que as probabilidades de fixação independem da *fitness* do indivíduo mutante e são a mesma probabilidade de fixação do Processo de Moran Neutro.

4.2.4 *Drift* Genético e Seleção

Vimos que a probabilidade de fixação do Processo de Moran é dado por:

$$\rho_M = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^N}}$$

Dizemos que um grafo G é ρ -equivalente ao Processo de Moran se a probabilidade de fixação no grafo G , ρ_G , coincide com ρ_M .

Se, para um indivíduo mutante "vantajoso" (com *fitness* $r > 1$), a probabilidade de fixação em G for maior do que a probabilidade de fixação no Processo de Moran ($\rho_G > \rho_M$), então o grafo G favorece mais a seleção do que o *drift*, pois aumenta a probabilidade de fixação de um mutante "vantajoso". Deste modo, dizemos que G é um amplificador de seleção. Caso $r > 1$ e $\rho_G < \rho_M$, então o grafo G favorece mais o *drift* do que a seleção, pois desfavorece um indivíduo mutante "vantajoso". Neste caso, dizemos que G é um supressor de seleção.

Similarmente, para os casos nos quais o indivíduo mutante é "desvantajoso" ($r < 1$), se $\rho_G > \rho_M$, dizemos que G é um supressor de seleção. Caso $\rho_G < \rho_M$, então G é um amplificador de seleção.

Se $\rho_G = \frac{1}{N}$ para qualquer r , o grafo G é o maior supressor possível de seleção, pois elimina completamente os efeitos causados pela diferença de *fitness* dos tipos de indivíduos. Nos exemplos acima, temos que o ciclo direcionado e o ciclo são ρ -equivalentes ao Processo de Moran e os grafos linha e *burst* eliminam completamente a seleção.

4.3 O Teorema Isotérmico

Definimos a temperatura de um vértice como a soma de todos os pesos que levam aquele vértice. Deste modo, a temperatura de um vértice j , denotada por T_j , é dada

pela expressão:

$$T_j = \sum_{i=1}^N w_{i,j}$$

que corresponde a soma da j -ésima coluna da matriz W .

Um vértice com alta temperatura tem uma tendência maior a mudar seu estado do que um vértice com temperatura menor. Se todos os vértices do grafo tem a mesma temperatura, dizemos que o grafo é isotérmico. Provaremos a seguir o Teorema Isotérmico: um grafo é ρ -equivalente ao Processo de Moran se, e somente se, o grafo é isotérmico.

Para um grafo isotérmico, segue que a soma das colunas é uma constante. Como W é linha-estocástico, segue que, para j vértice qualquer:

$$N \cdot T_j = \sum_{j=1}^N T_j = \sum_{j=1}^N \left(\sum_{i=1}^N w_{i,j} \right) = \sum_{i=1}^N \left(\sum_{j=1}^N w_{i,j} \right) = \sum_{i=1}^N 1 = N$$

$$\therefore T_j = \frac{N}{N} = 1$$

E portanto W é coluna-estocástica.

Teorema 4.1. Um grafo induzido pela matriz W é ρ -equivalente ao Processo de Moran se, e somente se, a matriz W é linha-estocástica e coluna-estocástica.

Demonstração. Inicialmente, defina o vetor $v = (v_1, \dots, v_N)$, em que $v_i = 1$ se o vértice i é um indivíduo do tipo \mathbb{A} , e $v_i = 0$ se o vértice i é um indivíduo do tipo \mathbb{B} . O vetor v define uma 2-coloração no grafo. Seja m o estado de espaço do processo estocástico. Segue que $m = \sum_i v_i$. A probabilidade do estado de espaço aumentar em uma geração é dada por:

$$P_{m,m+1} = \frac{r \sum_i \sum_j w_{i,j} v_i (1 - v_j)}{r \cdot m + 1(N - m)}$$

De fato, o termo geral da expressão acima é positivo apenas quando $v_i = 1, v_j = 0$ e $w_{i,j} > 0$. Isso significa que o vértice i é um indivíduo do tipo \mathbb{A} , o vértice j é um indivíduo do tipo \mathbb{B} e existe uma aresta orientada saindo de i e chegando em j . De modo análogo, temos que a probabilidade do estado de espaço diminuir na geração seguinte é dada por:

$$P_{m,m-1} = \frac{\sum_i \sum_j w_{i,j}(1-v_i)v_j}{r \cdot m + 1(N-m)}$$

Deste modo, a probabilidade de fixação é a mesma do Processo de Moran se, para qualquer coloração v , temos:

$$\frac{P_{m,m-1}}{P_{m,m+1}} = \frac{1}{r}$$

Temos a igualdade acima se, e somente se,

$$\sum_i \sum_j w_{i,j}v_i(1-v_j) = \sum_i \sum_j w_{i,j}(1-v_i)v_j$$

A igualdade acima deve ser verdade para todo vetor v . Em particular, deve ser verdade para todos os vetores da forma $v_k = 1$ e $v_i = 0, \forall i \neq k$. Neste caso, a equação acima é reduzida para:

$$\sum_j w_{k,j} = \sum_j w_{j,k}, \forall k$$

Como W é linha-estocástica, segue que:

$$\sum_j w_{j,k} = \sum_j w_{k,j} = 1$$

E portanto temos que a matriz W é coluna-estocástica e o grafo correspondente é isotérmico. \square

Nos exemplos acima, temos que o grafo ciclo direcionado e o grafo ciclo são ambos isotérmicos. Todos os grafos simétricos (que satisfazem $w_{i,j} = w_{j,i}$) são grafos isotérmicos.

4.4 Supressores de Seleção

Uma raiz é um tipo de vértice no qual não existem arestas direcionadas que chegam nele, e portanto tem temperatura zero. Se o grafo tem apenas uma raiz, então a probabilidade de fixação é $\rho = \frac{1}{N}$, pois o indivíduo mutante deve surgir na raiz, caso contrário não é capaz de fixar a população. Um mutante surge aleatoriamente na raiz com probabilidade $\frac{1}{N}$ e portanto todos os grafos com apenas uma raiz são supressores totais de seleção.

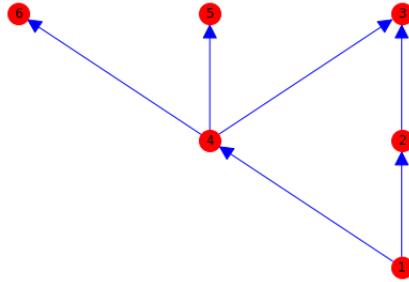


Figura 7: Exemplo de grafo com população $N = 6$. O vértice 1 é a única raiz.

Por outro lado, se um grafo tem múltiplas raízes, então um indivíduo mutante não é capaz de fixar-se na população. Se o indivíduo mutante surge em uma destas raízes, então ele também não será capaz de extinguir-se. Deste modo, é possível que em um grafo com múltiplas raízes, não aconteçam nem a fixação e nem a extinção.

Lema 4.2. Existe um grafo supressor de seleção com probabilidade de fixação de um mutante vantajoso ρ_G , em que $\frac{1}{N} < \rho_G < \rho_M$.

Demonstração. Seja V_1, V_2 uma decomposição da população em dois grupos tal que $V(G) = V_1 \cup V_2$ e $V_1 \cap V_2 = \emptyset$. Vamos construir um grafo da seguinte maneira: para os vértices de V_1 , faça um grafo completo (com todas as possíveis arestas). Depois, construa arestas saindo dos vértices de V_1 e chegando em V_2 , mas não o recíproco. O

grafo dos vértices V_2 tem estrutura qualquer, mas cada $i \in V_2$, deve existir $j \in V_1$ tal que exista i e j estão conectados. Nesta construção, definimos que os vértices de V_1 formam uma fonte, e os vértices de V_2 são receptores. A ideia é que, para o indivíduo mutante se fixar em $V(G)$, basta que ele se fixe em V_1 . Como V_2 não tem acesso aos vértices de V_1 , então é impossível que o indivíduo mutante se fixe em $V(G)$ se ele surgir aleatoriamente em V_2 . Portanto, temos que a probabilidade de fixação no grafo construído é:

$$\rho_G = \frac{1 - \frac{1}{r}}{1 - \frac{1}{rN_1}}$$

Se o mutante é vantajoso, então $r > 1$ e portanto $\frac{1}{N} < \rho_G < \rho_M$. □

Em geral, grafos cujo tamanho da fonte é pequeno em comparação ao tamanho dos receptores tendem a ser supressores de evolução. De fato, assumindo que a probabilidade do indivíduo mutante surgir aleatoriamente na fonte é pequena em relação a surgir nos vértices receptores, segue que, mesmo com maior *fitness* relativa, o mutante não consegue se fixar em vértices receptores.

4.5 Amplificadores de seleção

A estrutura do grafo pode ser balanceada de modo à favorecer a seleção ao invés do *drift*. Um exemplo é o grafo estrela de N vértices, construído a partir de uma simetria de arestas do grafo *burst*.

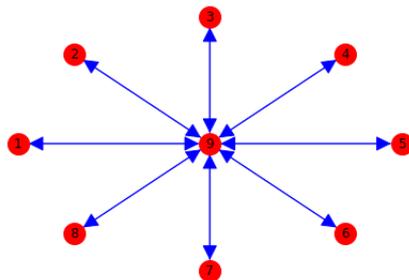


Figura 8: Exemplo de uma estrela com população $N = 9$

Para N suficientemente grande, a probabilidade de fixação de um mutante inserido aleatoriamente se aproxima de

$$\rho_M = \frac{1 - \frac{1}{r^2}}{1 - \frac{1}{r^{2N}}}$$

Como mostrado em ([14]) utilizando aproximações para as probabilidades de fixação como soluções de um conjunto de *EDO's*. Isso significa que uma fitness relativa r em um grafo estrela é equivalente a uma fitness r^2 em um processo de Moran, e, portanto, a estrela é um amplificador de seleção. De fato, um mutante vantajoso com fitness relativa $r > 1$ se comporta como um mutante com fitness relativa $r^2 > r$ em um processo de Moran simples. Um mutante desvantajoso, com fitness relativa $r < 1$, tem uma probabilidade de fixação equivalente a um mutante ainda mais desvantajoso $r^2 < r$ em um processo de Moran.

5 Conclusão

Ao fim deste trabalho, conseguimos alcançar os objetivos inicialmente propostos, tais quais, apresentar os conceitos fundamentais da teoria matricial, de cadeias de Markov e grafos e utilizá-los para construção do modelo estocástico do Processo de Markov, além de caracterizar a fixação de um gene neste modelo através do desenvolvimento do Teorema de Perron-Frobenius. Também identificamos a existência de estruturas populacionais que podem agir como supressores de seleção ou amplificadores de seleção.

Também encontramos uma condição necessária e suficiente para que a estrutura populacional mantenha as probabilidades de fixação encontradas no modelo simples através da demonstração do Teorema Isotérmico. Como esse resultado permite uma análise da estrutura através da soma dos elementos da coluna, o resultado tem aplicação computacional simples em relação às estimativas numéricas de simulações de modelos.

Destacamos que é possível generalizar as hipóteses do Teorema 3.3 para uma classe mais genérica de matrizes pois na demonstração utilizamos apenas o fato da submatriz principal M ser sub-estocástica, não-negativa e irreduzível, e não hipóteses da estrutura da matriz de Moran. Essa classe de matrizes, denotada de Classe de Kimura, é formalmente definida em ([2]), bem como algumas de suas principais propriedades topológicas.

A Raio Espectral e Norma Matricial

O raio espectral de uma matriz quadrada M , denotado por $\rho(M)$, é o maior valor dos módulos dos autovalores de M . Isto é, sejam $\lambda_1, \dots, \lambda_n$ os autovalores (não necessariamente distintos) da matriz $M \in \mathbb{C}^{n \times n}$. Definimos o raio espectral de M como

$$\rho(M) = \max\{|\lambda_1|, \dots, |\lambda_n|\}$$

Nesse apêndice, iremos destacar alguns resultados que relacionam raio espectral e as normas matriciais, baseado nos trabalhos de [1], [3] e [5]

Lema A.1. Seja $\|\cdot\|$ uma norma matricial sub-multiplicativa em $M_n(\mathbb{C})$. Então para todo $M \in M_n(\mathbb{C})$,

$$\rho(M) \leq \|M\|$$

Demonstração. Seja $\lambda \in \mathbb{C}$ autovalor de M e $v \neq 0$ um autovetor correspondente. De $Mv = \lambda v$, obtemos que:

$$MV = \lambda V, \quad \text{em que } V = [v | v | \dots | v] \in M_n(\mathbb{C})$$

Portanto,

$$|\lambda| \|V\| = \|\lambda V\| = \|MV\| \leq \|M\| \|V\|$$

Como $v \neq 0$, segue que $V \neq 0$ e logo $\|V\| > 0$. Simplificando ambos os lados da equação por $\|V\|$, obtemos que $|\lambda| \leq \|M\|$. Tomando o máximo dos módulos dos autovalores de M (M tem finitos autovalores), temos que $\rho(M) \leq \|M\|$. \square

Lema A.2 (Lema de Neumann). Seja $M \in \mathcal{M}_n(\mathbb{C})$. Então $\sum_{k=0}^{\infty} M^k$ converge se e somente se $\rho(M) < 1$.

Demonstração. Suponha que $\sum_{k=0}^{\infty} M^k$ converge e portanto existe $\lim_N S_N$, em que $S_N := \sum_{k=0}^N M^k$. Portanto,

$$\lim_n M^N = \lim_N (S_N - S_{N-1}) = \left(\lim_N S_N \right) - \left(\lim_N S_{N-1} \right) = 0$$

Seja (λ, v) um par de autovalor $\lambda \in \mathbb{C}$ e v um autovetor qualquer associado ao autovalor λ da matriz M . Segue que:

$$(M^N)v = \lambda^N v$$

e portanto

$$0 = \lim_N (M^N)v = \lim_N \lambda^N v$$

Como v é autovetor, então $v \neq 0$ e então $\lim_N \lambda^N = 0$. Assim, $|\lambda| < 1$, e, como há uma quantidade finita de autovalores de M , então $\rho(M) < 1$.

Por outro lado, suponha que $\rho(M) < 1$. Pelo Apêndice A, existe norma submultiplicativa tal que $\|M\| < 1$.

$$\|S_M - S_N\| = \left\| \sum_{k=N+1}^M M^k \right\| \leq \sum_{k=N+1}^M \|M^k\| \leq \sum_{k=N+1}^M \|M\|^k$$

Como $\|M\| < 1$, então $\lim_k \|M\|^k = 0$ e portanto é possível tomar M, N suficientemente grandes tais que $\|S_M - S_N\| < \varepsilon$ para todo $\varepsilon > 0$ dado. Deste modo, $\{S_N\}_{N \in \mathbb{N}}$ é sequência de Cauchy, e como $\mathcal{M}_n(\mathbb{C})$ é espaço de Banach, temos que $\sum_{k=0}^{\infty} M^k$ converge. \square

Seja $v = (v_1, \dots, v_n) \in \mathbb{R}^n$. Definimos a 1-norma (vetorial) como $\|v\|_1 := \sum_{i=1}^n |v_i|$. A 1-norma matricial é definida como:

$$\|M\|_1 := \max_{\|x\|_1=1} \|Mx\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^n |m_{i,j}|$$

A 1-norma matricial é sub-multiplicativa. De fato:

$$\|AB\|_1 = \max_{\|x\|_1=1} \|ABx\|_1 = \max_{\|x\|_1=1} \left\{ \|A\left(\frac{Bx}{\|Bx\|_1}\right)\|_1 \cdot \|Bx\|_1 \right\} \leq \|A\|_1 \cdot \|B\|_1$$

Lema A.3. Sejam $M \in M_n(\mathbb{C})$ e $\varepsilon > 0$. Então existe uma norma matricial sub-multiplicativa $\|\cdot\|$ tal que

$$\|M\| \leq \rho(M) + \varepsilon$$

Demonstração. A Forma Canônica de Jordan da matriz M é

$$M = P \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{n_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 \dots & 0 & J_{n_k}(\lambda_k) & \dots \end{bmatrix} P^{-1}$$

em que P é matriz não-singular, $\lambda_1, \dots, \lambda_k$ são os autovalores de M , as matrizes-bloco $J_{n_i}(\lambda_i), i = 1, \dots, k$ são matrizes bidiagonais, com entradas na diagonal principal λ_i e na diagonal acima da diagonal principal iguais a 1 (e demais entradas nulas) e $n_1 + \dots + n_k = n$. Defina para $\delta > 0$ arbitrário a matriz:

$$D(\delta) = \begin{bmatrix} D_{n_1}(\delta) & 0 & \dots & 0 \\ 0 & D_{n_2}(\delta) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & D_{n_k}(\delta) \end{bmatrix}, \quad \text{em que } D_{n_l}(\delta) = \text{diag}(\delta, \delta^2, \dots, \delta^{n_l})$$

Agora, como as matrizes $D_{n_i}(\varepsilon)$ são diagonais, então o produto $M D_{n_i}(\varepsilon)$ resulta na matriz cuja j -ésima coluna é a j -ésima coluna da matriz M multiplicada por ε^j . Também vemos que o produto $D_{n_i}(\frac{1}{\varepsilon}) M$ resulta na matriz cuja i -ésima linha é a i -ésima linha da matriz M multiplicada por $\frac{1}{\varepsilon^i}$. Deste modo, os elementos que tem coordenadas da forma $m_{i,i+1}$ estão na diagonal acima da diagonal principal e portanto

$$D_{n_i} \left(\frac{1}{\varepsilon} \right) J_{n_i} D_{n_i}(\varepsilon) = \begin{bmatrix} \lambda_l & \varepsilon & 0 & \dots & 0 \\ 0 & \lambda_l & \varepsilon & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda_l & \varepsilon \\ 0 & \dots & 0 & 0 & \lambda_l \end{bmatrix} =: B_{n_i}(\lambda_l, \varepsilon)$$

Portanto,

$$D \left(\frac{1}{\varepsilon} \right) P M P^{-1} D(\varepsilon) = \begin{bmatrix} B_{n_1}(\lambda_1, \varepsilon) & 0 & \dots & 0 \\ 0 & B_{n_2}(\lambda_2, \varepsilon) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & B_{n_k}(\lambda_k, \varepsilon) \end{bmatrix}$$

Seja $A \in M_n(\mathbb{C})$ matriz qualquer e defina a norma

$$\|A\| := \left\| D \left(\frac{1}{\varepsilon} \right) P A P^{-1} D(\varepsilon) \right\|_1$$

Assim, $\|M\| = \max_{j=1, \dots, n} |\lambda_j| + \varepsilon = \rho(M) + \varepsilon$ e portanto segue o enunciado. \square

Um dos resultados mais importantes que correlaciona a norma matricial e o raio espectral é a Fórmula de Gelfand.

Teorema A.4 (Fórmula de Gelfand). Seja $\|\cdot\|$ uma norma matricial e $M \in \mathbb{R}^{n \times n}$. Então:

$$\rho(M) = \lim_{k \rightarrow \infty} \|M^k\|^{\frac{1}{k}}$$

Demonstração. Se M é a matriz nula, então vemos que $\rho(M) = 0$ e $\|M\|^k = 0$ para todo k . Assim, o enunciado segue trivialmente. Suponha então M matriz não nula e tome $\varepsilon > 0$ qualquer. Defina a matriz

$$M_+ = \frac{1}{\rho(M) + \varepsilon} \cdot M$$

Deste modo, se λ é autovalor de M , então $\frac{\lambda}{\rho(M) + \varepsilon}$ é autovalor de M_+ . Mais ainda, como $\frac{1}{\rho(M) + \varepsilon} > 0$, então:

$$\left| \frac{\lambda}{\rho(M) + \varepsilon} \right| = \frac{1}{\rho(M) + \varepsilon} |\lambda|$$

E portanto $\rho(M_+) = \frac{1}{\rho(M) + \varepsilon} \rho(M) = \frac{\rho(M)}{\rho(M) + \varepsilon} < 1$.

E como $\rho(M_+) < 1$, pelo Lema de Neumann, a série $\sum_{k=0}^{\infty} M_+^k$ converge, então $\lim_k M_+^k = 0$. Deste modo, $\lim_k \|M_+^k\| = 0$ (pela continuidade da norma). Assim, existe índice $N_1 \in \mathbb{N}$ tal que:

$$k > N_1 \Rightarrow \|M_+^k\| < 1$$

Pela definição de M_+ , segue que

$$\|M\| = \|M_+\|(\rho(M) + \varepsilon)$$

E, para $k > N_1$, vemos que:

$$\|M^k\| = \|M_+^k\|(\rho(M) + \varepsilon)^k < (\rho(M) + \varepsilon)^k$$

$$\therefore \|M^k\|^{\frac{1}{k}} < \rho(M) + \varepsilon$$

Agora, por outro lado, defina a matriz M_- como

$$M_- = \frac{1}{\rho(M) - \varepsilon} \cdot M$$

Como M é matriz não-nula¹⁰, temos que $\rho(M) > 0$ e podemos tomar ε tal que $\varepsilon < \rho(M)$. Vemos que $\frac{1}{\rho(M) - \varepsilon} > 0$ e analogamente ao caso anterior, obtemos que

$$\rho(M_-) = \frac{\rho(M)}{\rho(M) - \varepsilon} > 1$$

E, como $\rho(M_-) > 1$, pelo Lema de Neumann, a série $\sum_{k=1}^{\infty} M_-^k$ não converge. Pela desigualdade triangular,

$$\left\| \sum_{k=1}^{\infty} M_-^k \right\| \leq \sum_{k=1}^{\infty} \|M_-^k\|$$

e portanto a série $\sum_{k=1}^{\infty} \|M_-^k\|$ não converge. Como os termos da série são não-negativos, então a série é não-limitada. Deste modo, existe índice $N_2 \in \mathbb{N}$ tal que

$$k > N_2 \Rightarrow \|M_-^k\| > 1$$

E, portanto, como

$$\|M^k\| = \|M_-^k\|(\rho(M) - \varepsilon)^k$$

Então

$$k > N_2 \Rightarrow \left\| \|M^k\|^{1/k} \right\| > \rho(M) - \varepsilon$$

Assim, se $k > \max\{N_1, N_2\}$, segue que

$$\rho(M) - \varepsilon < \|M^k\|^{1/k} < \rho(M) + \varepsilon$$

¹⁰A matriz nula é a única matriz real com todos os autovalores nulos. Logo, se M não é matriz nula, existe algum autovalor $\lambda \in \mathbb{C}$ tal que $|\lambda| > 0$ e portanto $\rho(M) \geq |\lambda| > 0$.

E portanto temos o limite $\rho(M) = \lim_k \|M^k\|^{\frac{1}{k}}$ □

Sejam $M = (m_{i,j}) \in \mathbb{R}^{n \times n}$. Defina a 2-norma matricial como:

$$\|M\|_2 = \left(\sum_{i=1}^n \sum_{j=1}^n (m_{i,j})^2 \right)^{\frac{1}{2}}$$

Lema A.5. Sejam $A = (a_{i,j}), B = (b_{i,j}) \in \mathbb{R}^{n \times n}$ matrizes quadradas. Se $0 \leq A \leq B$, então $\|A\|_2 \leq \|B\|_2$. Mais ainda, se $0 \leq A \leq B$, então $\rho(A) \leq \rho(B)$.

Demonstração. Como $0 \leq A \leq B$, então para todo par de índices $i, j = 1, \dots, n$, vemos que $0 \leq a_{i,j} \leq b_{i,j}$. Portanto, $0 \leq a_{i,j}^2 \leq b_{i,j}^2$ para todos índices e

$$\sum_{i=1}^n \sum_{j=1}^n (a_{i,j})^2 \leq \sum_{i=1}^n \sum_{j=1}^n (b_{i,j})^2$$

Como a função $x \mapsto x^{\frac{1}{2}}$ é crescente em $(0, +\infty)$ e pela definição de 2-norma matricial, temos que

$$\|A\|_2 = \left(\sum_{i=1}^n \sum_{j=1}^n (a_{i,j})^2 \right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^n \sum_{j=1}^n (b_{i,j})^2 \right)^{\frac{1}{2}} = \|B\|_2$$

Agora, suponha que $0 \leq A \leq B$ e tome $k \in \mathbb{N}$ qualquer. Pelo Lema 2.2, temos que $0 \leq A^k \leq B^k$ e, portanto, $0 \leq \|A^k\|_2 \leq \|B^k\|_2$. Como a função $x \mapsto x^{\frac{1}{k}}$ é crescente em $(0, +\infty)$, vemos que

$$0 \leq (\|A^k\|_2)^{\frac{1}{k}} \leq (\|B^k\|_2)^{\frac{1}{k}}$$

Tomando o limite $k \rightarrow \infty$ e, pela Fórmula de Gelfand, obtemos que:

$$0 \leq \rho(A) \leq \rho(B)$$

E portanto temos o enunciado. □

B Teorema de Perron-Frobenius

Nesse apêndice, iremos enunciar e provar o Teorema de Perron-Frobenius, baseado no trabalho de [REFERÊNCIA]. Inicialmente, iremos provar alguns resultados preliminares.

B.1 Teorema dos Discos de Gershgorin

Seja $A = (a_{i,j}) \in \mathbb{C}^{n \times n}$. Para cada $i = 1, \dots, n$, defina

$$R_i = \sum_{\substack{j=1, \dots, n \\ j \neq i}} |a_{i,j}|$$

Os conjuntos da forma $D[a_{i,i}, R_i] = \{x \in \mathbb{C} : |x - a_{i,i}| \leq R_i\}$ são chamados de discos de Gershgorin da matriz A .

Teorema B.1 (Teorema dos Discos de Gershgorin). Seja $A = (a_{i,j}) \in \mathbb{C}^{n \times n}$. Então todo autovalor de A pertence a pelo menos um disco de Gershgorin de A .

Demonstração. Sejam (λ, x) um par de autovalor e autovetor correspondentes de A e escolha $i \in \{1, \dots, n\}$ tal que $|x_i| \geq |x_j|$ para $j = 1, \dots, n, j \neq i$, em que $x = (x_1, \dots, x_n)$. Como x é autovetor, então x é não-nulo e portanto $|x_i| > 0$. Da igualdade $Ax = \lambda x$, temos que:

$$\begin{aligned} \sum_{j=1}^n a_{i,j}x_j &= \lambda x_i \\ \Rightarrow \sum_{j=1}^n a_{i,j}x_j - a_{i,i}x_i &= \lambda x_i - a_{i,i}x_i = (\lambda - a_{i,i})x_i \end{aligned}$$

$$\Rightarrow |\lambda - a_{i,i}| |x_i| = \left| \sum_{\substack{j=1, \dots, n \\ j \neq i}} a_{i,j}x_j \right|$$

Pela desigualdade triângular e como $x_i \neq 0$, temos que:

$$\Rightarrow |\lambda - a_{i,i}| \leq \sum_{\substack{j=1, \dots, n \\ j \neq i}} |a_{i,j}| \frac{|x_j|}{|x_i|}$$

Como $\frac{|x_j|}{|x_i|} \leq 1$ para $j = 1, \dots, n, j \neq i$, temos que:

$$|\lambda - a_{i,i}| \leq R_i$$

Portanto $\lambda \in D[a_{i,i}, R_i]$. □

Corolário B.1.1. Seja $A = (a_{i,j}) \in \mathbb{C}^{n \times n}$. Então $\rho(A) \leq \max_{i=1, \dots, n} \sum_{j=1}^n |a_{i,j}|$

Demonstração. Seja λ um autovalor de A . Pelo Teorema dos Discos de Gershgorin, existe um índice $i \in \{1, \dots, n\}$ tal que:

$$|\lambda - a_{i,i}| \leq R_i$$

Pela desigualdade triangular, temos que:

$$|\lambda| - |a_{i,i}| \leq |\lambda - a_{i,i}| \leq R_i$$

$$\therefore |\lambda| \leq R_i + |a_{i,i}| = \sum_{j=1}^n |a_{i,j}| \leq \max_{i=1, \dots, n} \sum_{j=1}^n |a_{i,j}|$$

Como λ é um autovalor arbitrário de A , temos o enunciado. □

B.2 Fórmula de Collatz-Wielandt

Teorema B.2 (Fórmula de Collatz-Wielandt). Sejam A matriz quadrada não-negativa e irredutível, $X = \{x \in \mathbb{R}^n : x > 0\}$ e considere a função $f : X \rightarrow \mathbb{R}$ definida como:

$$f(x) = \min_{\substack{i=1,\dots,n \\ (x)_i \neq 0}} \frac{(Ax)_i}{(x)_i}$$

Então as seguintes afirmações são verdadeiras:

- (a) f é contínua em $\{x \in \mathbb{R}^n; x \gg 0\}$
- (b) para $x \in X, i \in \{1, \dots, n\}$,

$$(Ax)_i \geq f(x) \cdot (x)_i$$

- (c) $f(x) = \max\{\rho \in \mathbb{R} : \rho x \leq Ax\}$ para todo $x \in X$
- (d) $f(x) = f(\alpha x)$ para todo $x \in X$ e $\alpha > 0$
- (e) Existe $z \in X$ tal que:

$$f(z) = \sup_{x \in X} f(x)$$

Demonstração. (a) Como o operador linear de \mathbb{R}^n dado por $x \mapsto Ax$ é contínuo, as funções projeções $\mathbb{R}^n \ni x \mapsto (x)_i \in \mathbb{R}$ são contínuas para todo $i = 1, \dots, n$, a função $x \in \mathbb{R}, 0 < x \mapsto \frac{1}{x}$ é contínua, a função min de coleção finita de funções contínuas é contínua e a composição de funções contínuas é uma função contínua, então f é função contínua em $\{x \in \mathbb{R}^n : x \gg 0\}$.

(b) De fato, se $(x)_i = 0$ para algum índice $i \in \{1, \dots, n\}$, então vemos que, como $A \geq 0$ e $x > 0$, então $(Ax)_i \geq 0 = f(x) \cdot (x)_i$. Por outro lado, se $(x)_i \neq 0$ para todo índice, então:

$$\frac{(Ax)_i}{(x)_i} \geq \min_{\substack{i=1,\dots,n \\ (x)_i \neq 0}} \frac{(Ax)_i}{(x)_i} = f(x) \Rightarrow (Ax)_i \geq f(x) \cdot (x)_i$$

(c) Inicialmente, temos que $f(x) \in \{\rho \in \mathbb{R} : \rho x \leq Ax\}$ pois já mostramos que $f(x) \cdot (x)_i \leq (Ax)_i$ para todo índice $i \in \{1, \dots, n\}$.

Seja $c \in \{\rho \in \mathbb{R} : \rho x \leq Ax\}$ qualquer e tome índice $i \in \{1, \dots, n\}$ tal que $\frac{(Ax)_i}{(x)_i} = f(x)$. Daí, temos que:

$$c \in \{\rho \in \mathbb{R} : \rho x \leq Ax\} \Rightarrow c(x)_i \leq (Ax)_i$$

Portanto,

$$c \leq \frac{(Ax)_i}{(x)_i} \Rightarrow c \leq f(x)$$

Logo, $f(x) \in \{\rho \in \mathbb{R} : \rho x \leq Ax\}$ é uma cota superior, então $f(x) = \max \{\rho \in \mathbb{R} : \rho x \leq Ax\}$.

(d) Sejam $x \in X$ e $\alpha > 0$ quaisquer. Como $(\alpha x)_i \neq 0$ se, e somente se $(x)_i \neq 0$, então:

$$f(\alpha x) = \min_{\substack{i=1, \dots, n \\ (\alpha x)_i \neq 0}} \frac{(A(\alpha x))_i}{(\alpha x)_i} = \min_{\substack{i=1, \dots, n \\ (x)_i \neq 0}} \frac{\alpha(Ax)_i}{\alpha(x)_i} = f(x)$$

(e) Inicialmente, considere $X_1 = X \cap S_1 = \{x \in X : \|x\| = 1\}$. Temos que, como S_1 é fechado e limitado e $\bar{X} = X \cup \{0 \in \mathbb{R}^n\}$, então X_1 é fechado e limitado, e pelo teorema de Bolzano-Weierstrass, é compacto. Como o operador $x \mapsto (A + I)^{n-1}x$ é contínuo, então o conjunto $Y = \{y \in \mathbb{R}^n : y = (A + I)^{n-1}x \text{ para algum } x \in X_1\}$ é compacto. Como $X_1 \subset X$ e pelos Lema 2.3 e pelo Teorema 2.7, temos que $Y \subset \{y \in \mathbb{R}^n : y \gg 0\}$. Por (a), segue que f é contínua em Y e portanto, pelo Teorema de Weierstrass, existe vetor $z \in Y$ tal que

$$f(z) = \max_{y \in Y} f(y)$$

Agora, seja $x \in X$ qualquer e escreva $y = (A + I)^{n-1}x$ e tome $\rho_x \in \{\rho \in \mathbb{R} : \rho x \leq Ax\}$. Como $(A + I)^{n-1}$ é matriz positiva e as matrizes A e $(A + I)^{n-1}$ comutam, então

$$\rho_x y = (A + I)^{n-1} \rho_x x \leq (A + I)^{n-1} Ax = A(A + I)^{n-1} x = Ay$$

e portanto $\rho_x \in \{\rho \in \mathbb{R} : \rho y \leq Ay\}$. Por (c), segue que $f(x) \leq f(y)$. Além disso, para todo $x \in X$, temos por (d) que $f(x) = f(\frac{x}{\|x\|})$, em que $\frac{x}{\|x\|} \in X_1$, que é compacto. Assim,

$$\sup_{x \in X} f(x) = \max_{x \in X_1} f(x) \leq \max_{y \in Y} f(y) = f(z)$$

E portanto temos o enunciado. □

B.3 Teorema de Perron-Frobenius

Teorema B.3 (Teorema de Perron-Frobenius). Seja $M = (m_{i,j}) \in \mathbb{R}^{n \times n}$ matriz irreduzível e não-negativa com raio espectral $r = \rho(M)$. Então as seguintes afirmações são verdadeiras:

- (a) O número r é um número real positivo e é autovalor da matriz M , chamado de autovalor de Perron-Frobenius.
- (b) A matriz M tem um autovetor z correspondente ao autovalor r que tem todas as componentes positivas.
- (c) O autovalor de Perron-Frobenius satisfaz:

$$\min_i \sum_j m_{ij} \leq r \leq \max_i \sum_j m_{ij}$$

Demonstração. (a) Sejam f a fórmula de Collatz-Wielandt e $z \in X = \{x \in \mathbb{R}^n : x > 0\}$ tal que $r := f(z) = \max_{x \in X} f(x)$. Vamos mostrar que r é o autovalor de Perron-Frobenius de M .

Para todo $x \in X$, temos que, como $M \geq 0$, então $f(x) > 0$ para todo $x \in X$. Portanto $r > 0$. Além disso, considere $y = (M + I)^{n-1}z \in X$. Pelo item (c) do Teorema B.2, temos que $rz \leq Mz$ e suponha por absurdo que não a igualdade não é válida. Daí, $Mz - rz > 0$ e, pelos Lema 2.3 e Teorema 2.7, temos que $(M + I)^{n-1}(Mz - rz) \gg 0$. Como as matrizes M e $(M + I)^{n-1}$ comutam, temos que:

$$(M + I)^{n-1}(Mz - rz) \gg 0 \Rightarrow My - ry \gg 0 \Rightarrow My \gg ry$$

E portanto é possível tomar $\varepsilon > 0$ tal que $My \gg (r + \varepsilon)y$. Pelo item (c) do Teorema B.2, então $f(y) > r + \varepsilon = f(z) + \varepsilon$, um absurdo pois $f(z)$ é máximo. Deste modo, $rz = Mz$ e portanto z é um autovetor de M com respectivo autovalor r .

(b) Escreva $y = (M + I)^{n-1}z$. Pelo Lema 2.3, temos que $y \gg 0$, e, por (a), z é autovetor de M . Pela expansão de termos de $(M + I)^{n-1}$, temos que:

$$\begin{aligned} y &= (M + I)^{n-1}z = \left(I + M + \binom{n-1}{2}M^2 + \dots M^{n-1} \right) z = \\ &= \left(1 + r + \binom{n-1}{2}r^2 + \dots r^{n-1} \right) z = (r + 1)^{n-1}z \end{aligned}$$

Como $z = \frac{1}{(r+1)^{n-1}}y$ e $y \gg 0$, então $z \gg 0$.

(c) Vamos provar que $r = \rho(M)$, onde $\rho(M)$ é o raio espectral da matriz M . De fato, seja $\alpha \in \mathbb{C}$ autovalor de M com respectivo autovetor v . Defina o vetor v^+ como $(v^+)_i = |(v)_i|$ para todo $v \in \mathbb{R}^n$ e todo índice $i \in \{1, \dots, n\}$. Portanto, para todo índice,

$$((\alpha v)^+)_i = |\alpha v_i| = |\alpha| \cdot |v_i| = (|\alpha| \cdot v^+)_i$$

Por outro lado,

$$((Mv)^+)_i = \left| \sum_{j=1}^n M_{i,j}v_j \right| \leq \sum_{j=1}^n |M_{i,j}v_j| = \sum_{j=1}^n M_{i,j}|v_j| = (Mv^+)_i$$

E portanto, como (α, v) são um par de autovalor e autovetor de M , então temos que $(Mv)^+ = (\alpha v)^+$ e segue que:

$$|\alpha|v^+ = (\alpha v)^+ = (Mv)^+ \leq (Mv^+)$$

Além disso, $v^+ > 0$ e portanto $v^+ \in X$. Pelo item (c) do Teorema B.2, temos que $|\alpha| \leq f(v^+)$ e portanto

$$|\alpha| \leq f(v^+) \leq f(z) = r$$

Como α é autovalor arbitrário e $|\alpha| \leq r$ que também é autovalor, segue que $r = \rho(M)$. Pelo Corolário B.1.1, segue imediatamente que:

$$r = \rho(M) \leq \max_{i=1, \dots, n} \sum_{j=1}^n |m_{i,j}| = \max_{i=1, \dots, n} \sum_{j=1}^n m_{i,j}$$

Por outro lado, tome o vetor $(1, \dots, 1) \in X$. Vemos que, para todo índice:

$$(M \cdot (1, \dots, 1))_i = \sum_{j=1}^n m_{i,j}$$

E como $f((1, \dots, 1)) \leq f(z) = r$, temos que:

$$\min_{i=1, \dots, n} \sum_{j=1}^n m_{i,j} = f((1, \dots, 1)) \leq f(z) = r$$

E assim obtemos que:

$$\min_i \sum_{j=1}^n m_{i,j} \leq r \leq \max_i \sum_{j=1}^n m_{i,j}$$

□

Corolário B.3.1. Sejam $M = (m_{i,j}) \in \mathbb{R}^{n \times n}$ matriz não-negativa e irredutível com raio espectral $r = \rho(M)$ e $x \in \mathbb{R}^n$ vetor positivo. Se $Mx < \beta x$ para algum $\beta \in \mathbb{R}$, então $r < \beta$. Além disso, se $M' = (m'_{i,j}) \in \mathbb{R}^{n \times n}$ é também matriz não-negativa e irredutível com $0 \leq M < M'$, então $\rho(M) < \rho(B)$.

Demonstração. Do item (c) da Fórmula de Collatz-Wietandt, vemos que $f(v) = \max\{\rho \in \mathbb{R} : \rho v \leq Mv\}$ para todo $v \in X = \{v \in \mathbb{R}^n : v > 0\}$. Como $x \in X$ mas $\beta \notin \{\rho \in \mathbb{R} : \rho v \leq Mv\}$, então $f(x) < \beta$. Seja $\rho(M) = \sup_{v \in X} f(v)$. Como β é cota superior, segue que $\rho(M) < \beta$.

Agora suponha que $0 \leq M < M'$. Tome $z' \in X$ um autovetor de Perron da matriz B . Pelo Corolário 2.2.1, vemos que:

$$0 \leq Mz' < Bz' = \rho(M')z'$$

E assim concluímos o enunciado usando a primeira parte desta demonstração. □

Referências Bibliográficas

- [1] A. Berman e R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Society for Industrial e Applied Mathematics, 1994. ISBN: 0-89871-321-8.
- [2] F. A. C. C. Chalub e M. O. Souza. «On the stochastic evolution of finite populations». Em: *Journal of Mathematical Biology* 75 (2017), pp. 1735–1774. DOI: 10.1007/s00285-017-1135-4.
- [3] Y.-J. Cheng e C.-w. Weng. «A matrix realization of spectral bounds of the spectral radius of a nonnegative matrix». Em: (2017). DOI: 10.48550/arXiv.1711.03274.
- [4] C. M. Grinstead e J. L. Snell. *Grinstead and Snell's Introduction to Probability: The CHANCE Project*. https://chance.dartmouth.edu/teaching_aids/books_articles/probability_book/pdf.html. Acesso em 08/05/2022. American Mathematical Society, 2006.
- [5] R. A. Horn e C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 2. ed., 2013. ISBN: 978-0-521-83940-2.
- [6] L. A. Imhof e M. A. Nowak. «Evolutionary game dynamics in a Wright-Fisher process». Em: *Journal of Mathematical Biology* 52 (2006), pp. 667–681. DOI: 10.1007/s00285-005-0369-8.
- [7] J. Díaz e D. Mitsche. «A survey of the modified Moran process and evolutionary graph theory». Em: (2021). DOI: 10.1016/j.cosrev.2020.100347.
- [8] J. Díaz, L. A. Goldberg, D. Richerby e M. Serna. «Absorption Time of the Moran Process». Em: (2018). DOI: 10.48550/arXiv.1311.7631.
- [9] S. Karlin e H. M. Taylor. *A first course in stochastic processes: Second Edition*. Academic Press, 1975. DOI: 10.1016/C2009-1-28569-8.
- [10] E. L. Lima. *Álgebra Linear*. IMPA, 9. ed., 2016. ISBN: 978-85-244-0420-7.
- [11] P. A. P. Moran. «Random processes in genetics». Em: (1957). DOI: 10.1017/S0305004100033193.
- [12] M. A. Nowak. *Evolutionary Dynamics: Exploring the equations of life*. Harvard University Press, 1. ed., 2006. ISBN: 978-0-674-02338-3.

- [13] H. A. Orr. «Fitness and its role in evolutionary genetics». Em: *Nature Reviews: Genetics* 10 (2009), pp. 531–539. DOI: 10.1038/nrg2603.
- [14] P. H. da Silva e M. O. Souza. «Continuous approximations for the fixation probability of the Moran processes on star graphs». Em: (2018). DOI: 10.48550/arXiv.2209.04572.