Uzbek Mathematical Journal
2026, Volume 70, Issue 2, pp. 68–73
DOI: 10.29229/uzmj.2026-2-8

Absorbing quadratic stochastic operators in genetics

Ganikhodjaev N.

Abstract. In this paper, we present exact solution of haploid Markov chain model that describes the fluctuation of gene frequency under the influence of mutation. Also generalizing Markov chain model we introduce and investigate a model that describes hybridization's process by quadratic stochastic operator.

Keywords: Markov chain, quadratic stochastic operator, gene, allele, hybrid
MSC (2020): 37N25, 60J10, 92D25

1. Genetic Motivation

Sewall Wright pointed out in an article written in 1949 [1] that the elementary evolutionary process in a reasonably large homogeneous population maybe considered to be change in gene frequency. This idea is central to population genetics, which studies the genetic makeup of populations and how it changes over time.

Recall the necessary notions from population genetics. A vague, but nevertheless informative, definition of a gene is simply a unit of hereditary information. The genetic code of an organism is carried on chromosomes. Each gene on a chromosome has different forms that it can take. These forms are called alleles. Haploid cells (or organisms) carry a single set of chromosomes and diploid organisms carry a double set of chromosomes one from each parent. The genotypes are gene composition and phenotypes are gene expression. Below we will consider haploid cells with a fixed population size of $2N$ genes composed of type-a and type-A individuals. Also we have made no mention of the dominant or recessive properties of the alleles.

2. Absorbing Markov chains

An absorbing Markov chain is a type of Markov chain where at least one state is absorbing, meaning once entered, it cannot be left. All other states are considered transient, meaning it's possible to transition to an absorbing state from them. In essence, the system will eventually reach and stay in one of the absorbing states.

The following idealized genetics model was introduced by S. Wright [2] to investigate the fluctuation of gene frequency under the influence of mutation and selection. Firstly we consider so-called simple haploid model of random reproduction, disregarding mutation pressures. Assume that we are dealing with a fixed population size of $2N$ genes composed of type-a and type-A individuals. The makeup of the next generation is determined by $2N$ independent Bernoulli trials as follows: if the parent population consists of j a-genes and $(2N-j)$ A-genes, then each trial results in a or A with probabilities

$$p_j = \frac{j}{2N}, q_j = 1 - \frac{j}{2N}$$

respectively.

Repeated selections are done with replacement. By this procedure one can generate a Markov chain X_n, where X_n is the number of a-genes in the nth generation among a constant population size of $2N$ individuals. The state space contains the $(2N+1)$ values $\{0, 1, 2, ..., 2N\}$. The transition probability matrix is computed according to the binomial distribution as

$$\Pr\{X_{n+1} = k | X_n = j\} = P_{jk} = \binom{2N}{k} p_j^k q_j^{2N-k} \tag{2.1}$$

where $j, k = 0, 1, \cdots, 2N$. Since $P_{00} = 1$ and $P_{2N,2N} = 1$, the states 0 and $2N$ are absorbing states, one of the questions of interest is to determine the probability, under the condition $X_0 = i$, that the population will attain fixation, i.e., that it will be come a pure population composed only of a-genes or A-genes.

In [3] using first step analysis presented exact solution for absorbing Markov chains with 1 and 2 transient states.

Below we will discuss exact solution for Markov chain (2.1) with $2N = 4$. In this case we have the following transition matrix

$$\Pi = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{81}{256} & \frac{27}{64} & \frac{54}{256} & \frac{3}{64} & \frac{1}{256} \\ \frac{1}{16} & \frac{4}{16} & \frac{6}{16} & \frac{4}{16} & \frac{1}{16} \\ \frac{1}{256} & \frac{3}{64} & \frac{54}{256} & \frac{27}{64} & \frac{81}{256} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Absorption occurs in states 0 and 4, and states 1,2 and 3 are transient. The probability of ultimate absorption in state 0, say, now depends on the transient state in which the process began. The first step analysis [3] proceeds by analysing, or breaking down, the possibilities that can arise at the end of the first transition, and then invoking the law of total probability coupled with the Markov property to establish a characterizing relationship among the unknown variables. Let

$$T = \min\{n \geq 0 : X_n = 0 \text{ or } X_n = 4\}$$

be the time of absorption of the process. Assume $u_i = \Pr\{X_T = 0|X_0 = i\}$ and $\nu_i = E[T|X_0 = i]$ for $i = 1, 2, 3$.

One can extend the definitions for u_i and ν_i in a consistent and common sense manner by prescribing $u_0 = 1, u_4 = 0$, and $\nu_0 = \nu_4 = 0$. We consider separately the five contingencies $X_1 = i$ with respective probabilities P_{0i} where $i = 0, 1, 2, 3, 4$. Let

$$\begin{aligned} \Pr\{X_T = 0|X_0 = 0\} &= 1 \\ \Pr\{X_T = 0|X_0 = 1\} &= u_1 \\ \Pr\{X_T = 0|X_0 = 2\} &= u_2 \\ \Pr\{X_T = 0|X_0 = 3\} &= u_3 \\ \Pr\{X_T = 0|X_0 = 4\} &= 0 \end{aligned}$$

We consider the three possible starting states $X_0 = 1, X_0 = 2$ and $X_0 = 3$ separately. For $X_0 = 1$ applying a first step analysis to $u_1 = \Pr\{X_T = 0|X_0 = 1)$, the law of total probability gives

$$\begin{aligned} u_1 &= \Pr\{X_T = 0|X_0 = 1\} \\ &= \sum_{k=0}^{4} \Pr\{X_T = 0|X_0 = 1, X_1 = k\} \Pr\{X_1 = k|X_0 = 1\} \\ &= \sum_{k=0}^{4} \Pr\{X_T = 0|X_1 = k\} \Pr\{X_1 = k|X_0 = 1\} \\ &= P_{10} + P_{11}u_1 + P_{12}u_2 + P_{13}u_3 \end{aligned}$$

Similarly one can produce equations for the cases $X_0 = 2$ and $X_0 = 3$. Finally we have the following system of equations

$$\begin{aligned} u_1 &= P_{10} + P_{11}u_1 + P_{12}u_2 + P_{13}u_3 \\ u_2 &= P_{20} + P_{21}u_1 + P_{22}u_2 + P_{23}u_3 \\ u_3 &= P_{30} + P_{31}u_1 + P_{32}u_2 + P_{33}u_3 \end{aligned}$$

For considered transition matrix Π we have the following system of equations

$$\begin{aligned} 148u_1 &= 81 + 54u_2 + 12u_3 \\ 10u_2 &= 1 + 4u_1 + 4u_3 \\ 148u_3 &= 1 + 12u_1 + 54u_2 \end{aligned}$$

The three equations in u_1, u_2 and u_3, are now solved simultaneously. By simple calculations we have $u_1 = 3/4; u_2 = 1/2; u_3 = 1/4$.
The result $u_3 = 1/4$ means that once begun in state $X_0 = 3$, the Markov chain ultimately end up in state 0 with probability $u_3 = 1/4$, and alternatively, will be absorbed in state 4 with probability $1 - u_3 = 3/4$.

The mean time to absorption also depends on the starting state. The first step analysis equations for $\nu_i = E[T|X_0 = i]$ are

$$\begin{aligned}\nu_1 &= 1 + P_{11}\nu_1 + P_{12}\nu_2 + P_{13}\nu_3\\ \nu_2 &= 1 + P_{21}\nu_1 + P_{22}\nu_2 + P_{23}\nu_3\\ \nu_3 &= 1 + P_{31}\nu_1 + P_{32}\nu_2 + P_{33}\nu_3\end{aligned}$$

For considered transition matrix, the equations are

$$\begin{aligned}148\nu_1 &= 256 + 54\nu_2 + 12\nu_3\\ 10\nu_2 &= 16 + 4\nu_1 + 4\nu_3\\ 148\nu_3 &= 256 + 12\nu_1 + 54\nu_2\end{aligned}$$

with solution $\nu_1 = 859/232; \nu_2 = 132/29; \nu_3 = 893/232$. Thus we have proved the following statement.

Theorem 2.1. *For $2N = 4$:*
i) a process that begins in state $X_0 = 3, (X_0 = 2$ or $X_0 = 1)$ the Markov chain ultimately end up in state 0 with probability $u_3 = 1/4($respectively $u_2 = 1/2$ or $u_1 = 3/4)$ and alternatively, will be absorbed in state 4 with probability $1 - u_i$ for $i = 1, 2, 3$;
ii) a process that begins in state $X_0 = 3, (X_0 = 2$ or $X_0 = 1)$ on the average $\nu_3 = 893/232 \approx 3,85($respectively $\nu_2 = 132/29 \approx 4.55$ or $\nu_1 = 859/232 = 3,7$ steps will transpire prior to absorption.

2.1. **Model with mutation pressure.** A more complete model takes account of mutation pressures. We assume that prior to the formation of the new generation each gene has the possibility to mutate, that is, to change into a gene of the other kind. Specifically, we assume that for each gene the mutation $a \to A$ occurs with probability α, and $A \to a$ occurs with probability β. Again we assume that the composition of the next generation is determined by 2N independent binomial trials. The relevant values of p_j and q_j when the parent population consists of j a-genes are now taken to be

$$p_j = \frac{j}{2N}(1-\alpha) + \left(1 - \frac{j}{2N}\right)\beta \tag{2.2}$$

$$q_j = \frac{j}{2N}\alpha + \left(1 - \frac{j}{2N}\right)(1-\beta) \tag{2.3}$$

Assume that the mutation pressures operate first, after which a new gene is chosen by random selection from the population. Now, the probability of selecting an a-gene after the mutation forces have acted is just $1/(2N)$ times the number of a-genes present; hence the average probability (averaged with respect to the possible mutations) is simply $1/(2N)$ times the average number of a-genes after mutation. But this average number is clearly $j(1-\alpha) + (2N - j)\beta$, which leads at once to (2.2-2.3). The transition probabilities of the associated Markov chain are calculated by (2.1) using the values of p_j (2.2) and q_j (2.3).
If $\alpha\beta > 0$, then $0 < p_j < 1$ for all j, that is fixation will not occur in any state. Instead, as $n \to \infty$, the distribution function of X_n will approach a steady-state distribution of a random variable ξ, where $\Pr\{\xi = k\} = \pi_k$ for $k = 0, \cdots, 2N$ with $\sum_{k=0}^{2N} \pi_k = 1$. The distribution function of ξ is called the steady state gene frequency distribution.

For brevity we consider the case $2N = 4$. Then we have the following transition matrix

$$\Pi = \begin{pmatrix} (1-\beta)^4 & 4\beta(1-\beta)^3 & 6\beta^2(1-\beta)^2 & 4\beta^3(1-\beta) & \beta^4\\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14}\\ p_{20} & p_{21} & p_{22} & p_{23} & p_{24}\\ p_{30} & p_{31} & p_{32} & p_{33} & p_{34}\\ \alpha^4 & 4\alpha^3(1-\alpha) & 6\alpha^2(1-\alpha)^2 & 4\alpha(1-\alpha)^3 & (1-\alpha)^4 \end{pmatrix}$$

where all entries of second, third and fourth rows are strictly positive. If $\alpha\beta > 0$ then Π is regular matrix,i.e. all entries strictly positive, therefore there is a single limit distribution, that is this matrix describes the steady-state distribution. Thus we have proved the following statement

Theorem 2.2. *For model with mutation pressure if $\alpha\beta > 0$ then one can reach the steady-state distribution.*

For $\alpha\beta = 0$ with $\alpha + \beta > 0$ absorption occurs only in state 0 or 4.

3. HYBRIDIZATION

A hybrid of two genotypes refers to an individual resulting from the cross-breeding of two organisms that have different genetic makeups (genotypes). This typically involves one parent being heterozygous (carrying different alleles for a trait). The resulting offspring inherits a mix of genetic material from both parents.

Hybridization is frequently used in plant breeding to create varieties with desirable characteristics like increased yield or pest resistance. As before we assume that we are dealing with a fixed population size of $2N$ genes composed of type-a and type-A individuals. We consider a hybrid of two genotypes $X_n = i$ and $\tilde{X}_n = j$. Then one can consider genotype of size $4N$ that consists of $i+j$ a-genes and $(4N-(i+j))$ A-genes. In this case the make up of the next generation is determined by $2N$ independent Bernoulli trials as follows: If the hybrid parent population consists of $i+j$ a-genes and $(4N-(i+j))$ A-genes, then each trial results in a or A with probabilities

$$p_{ij} = \frac{i+j}{4N}, q_{ij} = 1 - \frac{i+j}{4N}$$

respectively.

Repeated selections are done with replacement. Thus the transition probability that the next generation $X_{n+1} = k$ be the result of hybrid of two genotypes $X_n = i$ and $\tilde{X}_n = j$ is computed according to the binomial distribution as

$$\Pr\{X_{n+1} = k | X_n = i, \tilde{X}_n = j\} = P_{ij,k} = \binom{2N}{k} p_{ij}^k q_{ij}^{2N-k} \tag{3.1}$$

where $i, j, k = 0, 1, \cdots, 2N$.

By cubic matrix $(P_{ij,k})_{i,j,k=0}^{2N}$ we can specify a quadratic stochastic operator $V : S^{2N} \to S^{2N}$ that is defined as follows

$$(V\mathbf{x})_k = \sum_{i,j=0}^{2N} P_{ij,k} x_i x_j, \qquad k = 0, 1, 2, \cdots, 2N, \tag{3.2}$$

where

$$S^{2N} = \{\mathbf{x} = (x_0, x_1, x_2, \cdots, x_{2N}) \in R^{2N+1} : \forall i \ x_i \geq 0, \sum_{i=0}^{2N} x_i = 1\}.$$

Such operators frequently arise in many models of mathematical genetics, namely, theory of heredity [4], [5], [6], [7],[8]. Since $p_{00} = 1$ and $p_{(2N)(2N)} = 1$, then $P_{00,0} = 1$ and $P_{(2N)(2N),2N} = 1$, the states 0 and $2N$ are absorbing states, one of the questions of interest is to determine the probability, under the condition $X_0 = i$, that the population will attain fixation, i.e., that it will be come a pure population composed only of a-genes or A-genes.

Definition 3.1. A quadratic stochastic operator with few absorbing states is called absorbing qso.

A qso (3.2) is absorbing. Let us consider example of absorbing qso in the case $2N = 4$. Then we have the following qso $V : S^4 \to S^4$

$$\begin{aligned}
x_0' &= x_0^2 + \frac{81}{256}x_1^2 + \frac{1}{16}x_2^2 + \frac{1}{256}x_3^2 \\
&+\frac{2401}{2048}x_0x_1 + \frac{81}{128}x_0x_2 + \frac{625}{2048}x_0x_3 + \frac{1}{8}x_0x_4 + \frac{625}{2048}x_1x_2 \\
&+\frac{1}{8}x_1x_3 + \frac{81}{2048}x_1x_4 + \frac{81}{2048}x_2x_3 + \frac{1}{128}x_2x_4 + \frac{1}{2048}x_3x_4
\end{aligned} \tag{3.3}$$

$$\begin{aligned}
x_1' &= \frac{27}{64}x_1^2 + \frac{1}{4}x_2^2 + \frac{3}{64}x_3^2 \\
&+\frac{343}{512}x_0x_1 + \frac{27}{32}x_0x_2 + \frac{375}{512}x_0x_3 + \frac{1}{2}x_0x_4 + \frac{375}{512}x_1x_2 \\
&+\frac{1}{2}x_1x_3 + \frac{135}{512}x_1x_4 + \frac{135}{512}x_2x_3 + \frac{3}{32}x_2x_4 + \frac{7}{512}x_3x_4
\end{aligned} \tag{3.4}$$

$$\begin{aligned}
x_2' &= \frac{27}{128}x_1^2 + \frac{3}{8}x_2^2 + \frac{27}{128}x_3^2 \\
&+\frac{147}{1024}x_0x_1 + \frac{27}{64}x_0x_2 + \frac{675}{1024}x_0x_3 + \frac{3}{4}x_0x_4 + \frac{675}{1024}x_1x_2 \\
&+\frac{3}{4}x_1x_3 + \frac{675}{1024}x_1x_4 + \frac{675}{1024}x_2x_3 + \frac{27}{64}x_2x_4 + \frac{147}{1024}x_3x_4
\end{aligned} \tag{3.5}$$

$$\begin{aligned}
x_3' &= \frac{3}{64}x_1^2 + \frac{1}{4}x_2^2 + \frac{27}{64}x_3^2 \\
&+\frac{7}{512}x_0x_1 + \frac{3}{32}x_0x_2 + \frac{135}{512}x_0x_3 + \frac{1}{2}x_0x_4 + \frac{135}{512}x_1x_2 \\
&+\frac{1}{2}x_1x_3 + \frac{375}{512}x_1x_4 + \frac{375}{512}x_2x_3 + \frac{27}{32}x_2x_4 + \frac{343}{512}x_3x_4
\end{aligned} \tag{3.6}$$

$$\begin{aligned}
x_4' &= \frac{1}{256}x_1^2 + \frac{1}{16}x_2^2 + \frac{81}{256}x_3^2 + x_4^2 \\
&+\frac{1}{2048}x_0x_1 + \frac{1}{128}x_0x_2 + \frac{81}{2048}x_0x_3 + \frac{1}{8}x_0x_4 + \frac{81}{2048}x_1x_2 \\
&+\frac{1}{8}x_1x_3 + \frac{625}{2048}x_1x_4 + \frac{625}{2048}x_2x_3 + \frac{81}{128}x_2x_4 + \frac{2401}{2048}x_3x_4
\end{aligned} \tag{3.7}$$

Starting some initial point $(\tilde{x_0}, \tilde{x_1}, \tilde{x_2}, \tilde{x_3}, \tilde{x_4})$ we consider limit behaviour of corresponding trajectory. If limit exists, it will be fixed point.

For

$$S^4 = \{\mathbf{x} = (x_0, x_1, x_2, x_3, x_4) \in R^5 : \forall i \ x_i \geq 0, \sum_{i=0}^{4} x_i = 1\}$$

assume $intS^4 = \{\mathbf{x} \in S^4 : \prod_{i=0}^{4} x_i > 0\}$ is interior and $\partial S^4 = \{\mathbf{x} \in S^4 : \prod_{i=0}^{4} x_i = 0\}$ is boundary of the simplex S^4.

Let $M_0(1,0,0,0,0); M_1(0,1,0,0,0); M_2(0,0,1,0,0); M_3(0,0,0,1,0); M_4(0,0,0,0,1)$ are the vertices of the simplex.

Proposition 3.2. *For any* $\boldsymbol{x} \in S^4 \setminus \{M_0, M_4\}$ *we have* $V(\boldsymbol{x}) \in intS^4$.

Proof. It is evident $V(M_0) = M_0$ and $V(M_4) = M_4$. By simple calculations we have that eigenvalues of Jacobian qso V at the points M_0 and M_4 are $\lambda_1 = 2$; $\lambda_2 = 1; \lambda_3 = 3/8; \lambda_4 = 3/32$ and $\lambda_5 = 3/256$. Therefore M_0 and M_4 are unstable fixed points. Starting with $X_0 = 1 (X_0 = 2$, or $X_0 = 3$ we reach $V(M_1) = (81/256; 27/64; 27/128; 3/64; 1/256)$(respectively $V(M_2) = (1/16; 1/4; 3/8; 1/4; 1/16)$or $V(M_3) = (1/256; 3/64; 27/128; 27/64; 81/256)$, i.e. $V(M_1), V(M_2), V(M_3) \in intS^4$, then we can conclude validity of the proposition. □

Corollary 3.3. *For arbitrary* $\boldsymbol{x} \in S^4 \setminus \{M_0, M_4\}$ *we never reach pure population.*

Now we investigate the limit behaviour of the trajectories qso V.

Theorem 3.4. *Qso* V *is the regular transformation.*

Proof. As noted above the fixed points M_0 and M_4 are unstable fixed points. By numerical calculations one can show that $M(0.4499, 0.2634, 0.1593, 0.0895, 0.0379)$ is the unique fixed point in interior of simplex S^4. A Jacobian of qso at this fixed point is defined as follows

$$\Pi = \begin{pmatrix} 1.3415 & 0.7417 & 0.3888 & 0.1772 & 0.0679 \\ 0.3954 & 0.6950 & 0.6793 & 0.5121 & 0.3106 \\ 0.2776 & 0.3971 & 0.5579 & 0.4448 & 0.5911 \\ 0.0611 & 0.1453 & 0.2888 & 0.4679 & 0.6123 \\ 0.0144 & 0.0313 & 0.0851 & 0.2004 & 0.4181 \end{pmatrix}$$

and its eigenvalues are
$\lambda_1 = 2; \lambda_2 = 0.3786; \lambda_3 = 0.9374; \lambda_4 = 0.0705 - 0.0294i; \lambda_5 = 0.0705 + 0.0294i.$
Since all eigenvalues except $\lambda_1 = 2$ by modulo less then 1 we conclude that M is attracting fixed point, i.e V is the regular transformation. $\square$

REFERENCES

[1] Wright S., Evolution and the Genetics of Populations, Volume 2: The Theory of Gene Frequencies. Universiy of Chikago, (1969).

[2] Fisher R. A., The Genetical Theory of Natural Selection. Clarendon Press, London and New York, (1962).

[3] Taylor H. M., Karlin S., An Introduction to Stochastic Modelling. Academic Press, San Diego; London; Boston; New York; Sydney; Tokyo; Toronto, (1998).

[4] Bernstein S. N., Uchn. Zapiski. NI Kaf. Ukr. Otd. Mat. 1, 83. (1924),

[5] Jenks R. D., J. Diff. Eqs. 5, 497. (1969),

[6] Kesten H., Adv. Appl. Prob. 2, 1. (1970),

[7] Lyubich Y. I., Russian Math. Surveys 26, 51. (1978),

[8] Lyubich Y. I., Mathematical Structures in Population Genetics. Springer-Verlag, (1992).

Ganikhodjaev N.N.,
V.I.Romanovskiy Institute of Mathematics,
Uzbekistan Academy of Sciences,
Tashkent, Uzbekistan.
e-mail: nasirgani@yandex.com