Malaysian Journal of Science 27 (2): 131 – 136 (2008)

REPORT

Associativity in Inheritance or are there Associative Populations?

Nasir Ganikhodjaev[1, 2 *] and **Haziana Hartini binti Hisamuddin**[1]

[1] Department of Computational and Theoretical Sciences, Faculty of Science, International Islamic University Malaysia, P. O. Box 141, 25710 Pahang, Malaysia
* gnasir@iiu.edu.my Telephone: 6 09 5716752 Facsimile: 6 09 5716781
(corresponding author)
[2] Department of Mechanics and Mathematics, National University of Uzbekistan, 700095, Tashkent, Uzbekistan
Received in revised form 24th December 2007, accepted 26th March 2008.

ABSTRACT An algebraic structure naturally occurs as genetic information gets passed down through the generations. The existence of associative algebras with genetic realization is proved.

(Associative algebra, genetic realization, quadratic stochastic operator)

HISTORY OF GENERAL GENETIC ALGEBRAS

The theories of heredity attributed to Gregor Mendel (1822 - 1884), were based on his work with pea plants. But his work was so brilliant and unprecedented at the time it appeared that it took thirty-four years for the rest of the scientific community to catch up to it. The short monograph, *Experiments with Plant Hybrids*, in which Mendel described how traits were inherited, has become one of the most enduring and influential publications in the history of science. He saw that the traits were inherited in certain numerical ratios. He then came up with the idea of dominance and segregation of genes and set out to test it in peas. It took seven years to cross and score the plants to the thousand to verify the laws of inheritance. From his studies, Mendel derived certain basic laws of heredity: hereditary factors do not combine, but are passed intact; each member of the parental generation transmits only half of its hereditary factors to each offspring (with certain factors "dominant" over others); and different offspring of the same parents receive different sets of hereditary factors. Mendel's work became the foundation for modern genetics.

MENDEL GENETIC ALGEBRA

Mendel, in his first paper [4] exploited some symbolism, which is quite algebraically suggestive, to express his genetic laws. In fact, it was later termed "Mendelian algebras" by several authors. In the 1920s and 1930s, general genetic algebras were introduced. Apparently, Serebrowsky [6] was the first to give an algebraic interpretation of the multiplication sign "x", which indicated reproduction, and to give a mathematical formulation of the Mendelian laws. The systematic study of algebras occurring in genetics was due to I. M. H. Etherington. In his paper [1], he succeeded in giving a precise mathematical formulation of Mendel's laws in terms of non-associative algebras.

More recent results including evolution in genetic algebras can be found in the book [3]. A very good survey article is Reed's [5].

General genetic algebras indeed are the product of interaction between biology and mathematics. In addition, Mendelian genetics offers a new object to mathematics that is general genetic algebras. The study of these algebras reveals the algebraic structure of Mendelian genetics, which always simplifies and shortens the way to understand the genetic and evolutionary phenomena in the real world.

GENETIC MOTIVATION

Before we discuss the genetic from a mathematical perspective, it is useful to know some of the basic language from biology. A gene is a unit of hereditary information. The genetic code of an organism is carried on chromosomes. In addition, each gene on a chromosome can take different forms that are called alleles. For example, the gene that determines blood type in humans has three different alleles, that are A, B, and O. Blood types for human are determined by two alleles since humans are diploid organisms. This means that we carry a double set of chromosomes, one from each parent. Moreover, when diploid organisms reproduce, a process called meiosis produces sex cells that are called gametes. Gametes carry a single set of chromosomes and when gametes fuse or reproduce, the result is a zygote, which again is a diploid cell.

SIMPLE MENDELIAN INHERITANCE

As a natural first example, we consider simple Mendelian inheritance for a single gene with two alleles A and a. The rules of simple Mendelian inheritance indicate that the next generation will inherit either A or a with equal frequency ½. Therefore, when two gametes reproduce, a multiplication is induced which indicates the way the hereditary information will be passed down to the next generation. This multiplication is given by the following rules:

$A \times A = A;$ (1)

$A \times a = \frac{1}{2} A + \frac{1}{2} a;$ (2)

$a \times A = \frac{1}{2} a + \frac{1}{2} A;$ (3)

$a \times a = a.$ (4)

Rules (1) and (4) are expressions of the fact that if both gametes carry the same allele, then the offspring will inherit it. Rules (2) and (3) indicate that when gametes carrying A and a reproduce, half of the time the offspring will inherit A and the other half of the time it will inherit a.

REMARKS

The rules (2) and (3) have a statistical nature. Let us consider N pairs of gametes carrying A and a respectively. Let N_1 be the number offspring inheriting A and $N_2 = N - N_1$ be the number offspring inheriting a. Then the ratio $\frac{N_1}{N}$ represents the frequency of allele A and it indicates that the next generation will inherit A with frequency $\frac{N_1}{N}$. Similarly the ratio $\frac{N_2}{N} = 1 - \frac{N_1}{N}$ represents the frequency of allele a and it indicates that the next generation will inherit a with frequency $\frac{N_2}{N}$. Thus for simple Mendelian inheritance we have $\frac{N_1}{N} = \frac{N_2}{N} = \frac{1}{2}$, that is the next generation inherit either A or a with equal frequency ½.

The rules (1 - 4) are an algebraic representation of the rules of simple Mendelian inheritance. This multiplication table is shown in Table 1.

Table 1. Multiplication Table for Simple Mendelian Inheritance

	A	a
A	A	$\frac{1}{2}(A + a)$
a	$\frac{1}{2}(a + A)$	a

We should point out that we are only concerning ourselves with genotypes: gene composition and not gene expression or phenotypes. Hence, we have made no mention of the dominant or recessive properties of our alleles.

Now that we have defined a multiplication on the symbols A and a we can mathematically define the two dimensional algebra over the set of real numbers R with basis $\{A, a\}$ and multiplication table as in Table 1. This algebra is called the gametic algebra for simple Mendelian inheritance with two alleles.

THE NONASSOCIATIVITY OF INHERITANCE

For those elements of the gametic and zygotic algebras which represent populations, multiplication of two such elements represents random mating between the two populations. From Table 1, we can see that the genetic algebra is commutative. In biological terms this means that result when population P mates with population Q is the same as when population Q mates with population P. Thus the algebra holds the commutative property.

However, if population P mates with population Q and then the obtained population mates with R, the resulting population is not the same as the population resulting from P mating with the population obtained from mating Q and R originally. Symbolically,

(P x Q) x R is not equal to P x (Q x R). For example,

$$\mathbf{Ax(Axa)=Ax(}\frac{1}{2}(A + \boldsymbol{a})\mathbf{=}\ \frac{1}{2}\mathbf{(AxA)+}\ \frac{1}{2}\mathbf{(Axa)=}$$

$$\frac{1}{2}\mathbf{A}+\frac{1}{4}\mathbf{A}+\frac{1}{4}\mathbf{a}=\frac{3}{4}\mathbf{A}+\frac{1}{4}\ \mathbf{a}\,.$$

However **(AxA)xa =Axa** $= \frac{1}{2}(A + \boldsymbol{a})$,so that

Ax(Axa)≠ (AxA)xa.

In general, the algebras which arise in genetics are commutative but non-associative.

Nevertheless, in this report we will show the existence of associative genetic algebras and discuss the biological meaning of such algebras. We will apply the theory of quadratic stochastic operators [2].

QUADRATIC STOCHASTIC OPERATOR

The set $S^{n-1} = \{x= (x_1, x_2 \ldots, x_n) \in R^n : x_i \geq 0,\ i = 1,2..,n$ and $\sum_{i=1}^{n} x_i = 1\}$ is called (n -1) dimensional simplex in R^n.

If n = 2, the 1-dimensional simplex S^1 in R^2 has the following form:

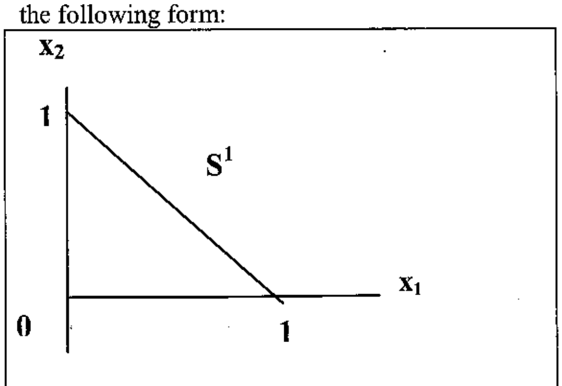

Figure 1. 1-Dimensional Simplex S^1 in R^2

A mapping V: $S^{n-1} \rightarrow S^{n-1}$ is called a quadratic stochastic operator, if for any $x = (x_1, x_2 ,\ldots, x_n) \in S^{n-1}$, $x' = Vx$ is defined as

$$x_k' = \sum_{i,j=1}^{n} P_{ij,k}\ x_i x_j$$

where the coefficients $P_{ij,k}$ (so-called coefficients of heredity) satisfy the following conditions:

(i) $P_{ij,k} \geq 0$;

(ii) $P_{ij,k} = P_{ji,k}$

and

(iii) $\sum_{k=1}^{n} P_{ij,k} = 1$

for all i, j, k ϵ { 1, 2, …, n}.

If n = 2 and $x = (x_1, x_2) \in S^1$, then

$$x_1' = P_{11,1}x_1^2 + 2\,P_{12,1}x_1x_2 + P_{22,1}\,x_2^2$$

$$x_2' = P_{11,2}x_1^2 + 2\,P_{12,2}x_1x_2 + P_{22,2}\,x_2^2$$

As for a numerical example, if we let:

$P_{11,1} = 1$ $P_{12,1} = ½$ $P_{22,1} = 0$

and respectively

$P_{11,2} = 0$ $P_{12,2} = ½$ $P_{22,2} = 1$

then $x_1' = 1 \cdot x_1^2 + 2\,(½)x_1x_2 + 0 \cdot x_2^2$

$= x_1^2 + x_1x_2 = x_1(x_1+x_2) = x_1$

and $x_2' = x_2$.

We shall call a quadratic stochastic operator V a genetic realization of some model of heredity.

ALGEBRA WITH GENETIC REALIZATION

Mathematically, the algebras that arise in genetics are very interesting structures. They are generally commutative but non-associative, yet they are not necessarily Lie, Jordan, or alternative algebras. In addition, many of the algebraic properties of these structures have genetic significance. Indeed, the interplay between the purely mathematical structure and the corresponding genetic properties makes this subject so fascinating.

The most general definition of an algebra $\mathfrak{R}$ which could have genetic significance is that of an algebra with genetic realization.

Let a quadratic stochastic operator V: $S^{n-1} \rightarrow S^{n-1}$ be a genetic realization, where V is defined by cubic matrix $\{P_{ij,k} : i, j, k = 1,...,n\}$ such that

(i) $P_{ij,k} \geq 0$;

(ii) $P_{ij,k} = P_{ji,k}$ and;

(iii) $\sum_{k=1}^{n} P_{ij,k} = 1$

An algebra $\mathfrak{R}$ with genetic realization V is an algebra over the real numbers R which has a basis $\{a_1,\ldots, a_n\}$ and a multiplication table

$$a_i a_j = \sum_{k=1}^{n} P_{ij,k} a_k$$

where $0 \leq P_{ij,k} \leq 1$ for all i, j, k and $\sum_{k=1}^{n} P_{ij,k} = 1$ for all $i, j = 1, \ldots, n$.

Here $P_{ij,k}$ is a frequency that the next generation reproduced by two gametes carrying a_i and a_j will inherit a_k, $k = 1, 2, \ldots, n$.

Such a basis is called the natural basis for $\mathfrak{R}$.
If n = 2, $\{A, a\}$ is the natural basis of $\mathfrak{R}$ and if
$P_{AA,A} = 1$ $P_{Aa,A} = ½$ $P_{aa,A} = 0$,
then we have the simple Mendelian inheritance algebra.

ASSOCIATIVE ALGEBRAS WITH GENETIC REALIZATION

Now we will describe associative algebras with genetic realization. Here we will restrict ourselves to the case n=2.

Let $\{A, a\}$ be the natural basis of $\mathfrak{R}$ as above in the case of simple Mendelian inheritance. Assume $P_{AA,A}$ is a frequency that the next generation reproduced by two gametes carrying **A** will inherit **A** and $P_{AA,a}$ is a frequency that the next generation reproduced by two gametes carrying **A** will inherit **a** ,then

$P_{AA,A} + P_{AA,a} = 1$(see **Remark**).
Similarly $P_{Aa,A} + P_{Aa,a} = P_{aa,A} + P_{aa,a} = 1$,

where $P_{Aa,A}$ (resp. $P_{Aa,a}$) is a frequency that the next generation reproduced by gametes carrying **A** and **a** will inherit **A** (resp. $\boldsymbol{a}$), and $P_{aa,A}$ (resp. $P_{aa,a}$) is a frequency that the next generation reproduced by two gametes carrying **a** will inherit **A** (resp. $\boldsymbol{a}$) (see **Remark**).

A member **u** of algebra $\mathfrak{R}$ is a linear combination of **A** and **a**, that is u=λ**A**+μ**a**, where λ and μ are arbitrary real numbers. Now algebra $\mathfrak{R}$ is associative if for any **u, v, w** ∈ $\mathfrak{R}$ we have: $(\mathbf{u} \cdot \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot (\mathbf{v} \cdot \mathbf{w})$.

It is easy to see that an n-dimensional algebra with a genetic realization is associative if and only if for $i, j, k, s = 1,..., n$

$$\sum_{r=1}^{n} P_{ij,r} P_{rk,s} = \sum_{r=1}^{n} P_{ir,s} P_{jk,r}$$

For n=2 this becomes for $i, j, k, s = 1,2$

$$P_{ij,1}P_{1k,s} + P_{ij,2}P_{2k,s} = P_{i1,s}P_{jk,1} + P_{i2,s}P_{jk,2}$$

This clearly holds if $i = k$, so we may as well assume that i=1 and k=2, giving

$$P_{1j,1}P_{12,s} + P_{1j,2}P_{22,s} = P_{11,s}P_{j2,1} + P_{12,s}P_{j2,2}$$

for $j, s = 1,2$.

Using the fact that $P_{ij,1} + P_{ij,2} = 1$ for i,j=1,2 it is straightforward to check that this holds precisely when

$$P_{11,1}P_{22,1} + P_{12,1} - P_{12,1}^2 - P_{22,1} = 0.$$

THEOREM OF ASSOCIATIVE ALGEBRA

We have proved the following theorem.
Theorem: The two-dimensional algebra with the standard basis $\{A, a\}$ and with genetic realization
$P_{AA,A} = P_1$ $P_{Aa,A} = P_2$ $P_{aa,A} = P_3$
and
$P_{AA,a} = 1 - P_1$ $P_{Aa,a} = 1 - P_2$ $P_{aa,a} = 1 - P_3$
is associative algebra if and only if the coefficients P_1, P_2, P_3 satisfy the following equation

$$P_1P_3 + P_2 - P_2^2 - P_3 = 0 \qquad (5)$$

We denote $P_2 = x$, $P_3 = y$, $P_1 = z$ and rewrite equation (5) in the following form:

$$z = (x^2 - x + y)/y \qquad (6)$$

The variables x, y and the function (6) should to satisfy following conditions:
(a) $0 \leq x \leq 1$,
(b) $0 \leq y \leq 1$,
(c) $0 \leq z \leq 1$

Let us describe all possible values (x, y) such that $0 \leq z \leq 1$. Consider the following two inequalities:

$(x^2 - x + y) / y \leq 1$ (7) and;

$(x^2 - x + y) / y \geq 0$ (8)

From (7) we have:

$x^2 - x + y \leq y$; $x^2 - x \leq 0$; thus $x(x-1) \leq 0$ which valid for all $x \in [0, 1]$

From (8) we have

$x^2 - x + y \leq 0$ or $y \geq x - x^2$

Let us consider the graph of the function $f(x) = x - x^2$

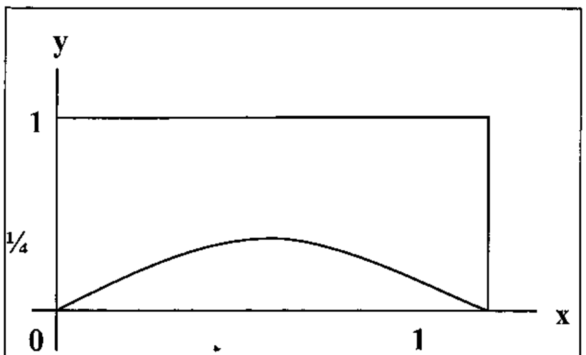

Figure 2. Graph of the function $f(x) = x - x^2$

Thus for any $x \in [0, 1]$, and $x - x^2 \leq y \leq 1$, we have $0 \leq (x^2 - x + y) / y \leq 1$

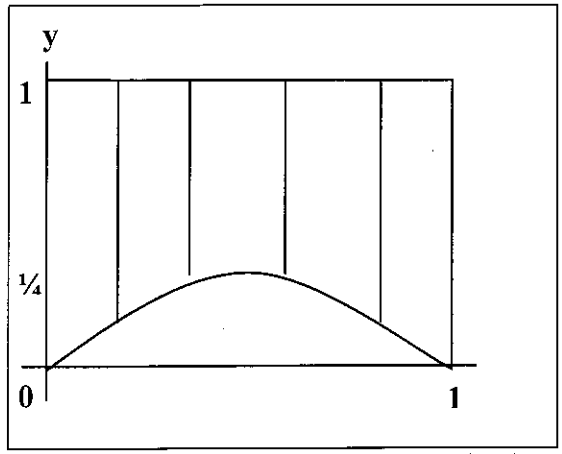

Figure 3. Domain of the function $z = f(x,y)$ with $0 \leq z \leq 1$

The shaded region on Figure 3 is the domain of the function $z = f(x,y)$ such that $0 \leq z \leq 1$.

Let us consider the graph of function $z=f(x,y)$ (Figure 4).

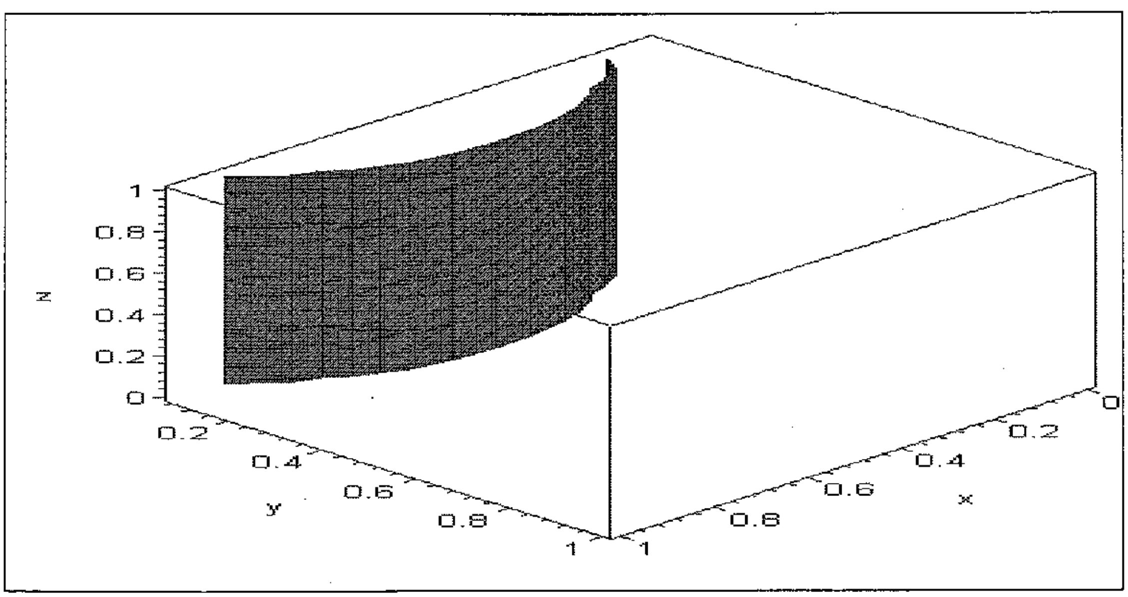

Figure 4. Graph of the function $z=f(x,y)$

The chosen surface on Figure 4 is the area that the algebra generated by point on this surface is associative. Any point that lies outside the surface determines an algebra that is not associative.

CONCLUSION

Let us call a population Q an associative one if the genetic algebra corresponding to is associative. Gregor Mendel showed that the genetic algebra corresponding to a population of pea plants is not an associative algebra. We have proved that there are associative algebras corresponding to associative populations.

There arises a natural question:
Are there in nature associative populations?
We address this problem to specialists in biotechnology and genetic engineering. Probably associative populations have some unexpected properties.

ACKNOWLEDGMENTS

The authors would like to thank referee for helpful suggestions to improve the paper.

REFERENCES

1. Etherington I. M. H. (1941). Non-associative algebra and the symbolism of genetics. *Proc. Roy. Soc. Edinburgh* **61**: 24 - 42.
2. Ganikhodjaev N. N. (1999). An application of the theory of Gibbs distributions to Mathematical Genetics. *Doklady Mathematics* **61** (3): 13 - 16.
3. Lyubich Y. I. (1992). *Mathematical Structures in Population Genetics*. Springer-Verlag, New York
4. Mendel G. (1959). Experiments in Plant-Hybridization. In *Classic Papers in Genetics*, (eds. Peters, J. A.). Prentice-Hall Inc. USA. pp. 1-20.
5. Reed M. L. (1997). Algebraic structure of genetic inheritance. *Bulletin of AMS* **34** (2): 107 - 130.
6. Serebrowsky, A. (1934). On the properties of the Mendelian equations, *Doklady A. N. SSSR*. **2**: 33 - 36.
7. Worz - Busekros A. (1980). *Algebras in Genetics, Lecture Notes in Biomathematics* 36, Springer-Verlag, Berlin.