EVOLUTION ALGEBRA

Freshman Seminar
University of California, Irvine

Bernard Russo

University of California, Irvine

Winter 2015

Table of Contents

January 6—Linear Algebras

(Ref.—Dickson, "Linear Algebras", 1914, pp.1–6)

Two couples of real numbers (a, b) and (c, d) are called equal if $a = c, b = d$.

Addition, subtraction and multiplication of two couples are defined by

- $(a, b) + (c, d) = (a + c, b + d)$
- $(a, b) - (c, d) = (a - c, b - d)$
- $(a, b)(c, d) = (ac - bd, ad + bc)$

Addition is seen to be commutative and associative:

- $x + x' = x' + x$, $(x + x') + x'' = x + (x' + x'')$

where x, x', x'' are any couples, $x = (a, b)$, $x' = (a', b')$, $x'' = (a'', b'')$.

Multiplication is commutative, associative, and distributive:

- $xx' = x'x$, $(xx')x'' = x(x'x'')$ $\quad$ (xx is denoted by x^2)
- $x(x' + x'') = xx' + xx''$, $(x' + x'')x = x'x + x''x$

Division is defined as the operation inverse to multiplication. Division except by (0,0) is possible and unique:

$$\frac{(c,d)}{(a,b)} = \left(\frac{ac+bd}{a^2+b^2}, \frac{ad-bc}{a^2+b^2}\right)$$

In particular we have

- $(a,0) \pm (c,0) = (a \pm c, 0)$, $(a,0)(c,0) = (ac,0)$, $\frac{(c,0)}{(a,0)} = \left(\frac{c}{a}, 0\right)$

Hence the couples $(a,0)$ combine under the above defined addition, multiplication, etc. exactly as the real numbers a combine under ordinary addition, multiplication, etc.

Thus, there is no danger in identifying the couple $(a,0)$ with the real number a, just as we identify the natural numbers among the signed integers, the integers among the rational numbers, and the latter among the real numbers

If, for brevity, you write $i = (0,1)$, then $i^2 = (0,1)(0,1) = (-1,0) = -1$ and you get the **complex numbers**: $(a,b) = (a,0) + (0,b) = a + (b,0)(0,1) = a + bi$

A set of complex numbers is called a **number field** if the sum, difference, product, and quotient (the divisor not being zero) of any two equal or distinct numbers of the set must be numbers belonging to the set.

Examples: complex numbers, real numbers, rational numbers.
(The set of integers is not a number field)

The concept of **matrix** affords an excellent example of a **linear algebra**. We can consider square matrices of n rows and n columns. For convenience, we take $n = 2$. Let

$$m = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ and } \mu = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

be two matrices, where the elements $a, b, c, d, \alpha, \beta, \gamma, \delta$ belong to a fixed number field F, which will usually be the real numbers.

We say that m and μ are equal if their corresponding elements are equal, $a = \alpha$, etc. Addition and multiplication are defined by

$$m + \mu = \begin{bmatrix} a+\alpha & b+\beta \\ c+\gamma & d+\delta \end{bmatrix}, \ m\mu = \begin{bmatrix} a\alpha + b\gamma & a\beta + b\delta \\ c\alpha + d\gamma & c\beta + d\delta \end{bmatrix}$$

Addition is commutative and associative

$$x + x' = x' + x \ , \ (x + x') + x'' = x + (x' + x'')$$

where x, x', x'' are any matrices of the same size.

Multiplication is associative and distributive

- $(xx')x'' = x(x'x'')$
- $x(x' + x'') = xx' + xx'' \ , \ (x' + x'')x = x'x + x''x$

However, multiplication of matrices is not commutative, and division m/μ is not always possible, even if

$$\mu \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Consider the four special matrices

$$e_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \ , \ e_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \ , \ e_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \ , \ e_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Their sixteen possible products by twos can be summarized as

$$e_{ij}e_{tk} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ if } t \neq j \qquad \text{and} \qquad e_{ij}e_{jk} = e_{ik} \tag{1}$$

Table 1 $X \times Y$

		Y			
	X $\times$ **Y**	e_{11}	e_{12}	e_{21}	e_{22}
X	e_{11}	e_{11}	e_{12}	0	0
	e_{12}	0	0	e_{11}	e_{12}
	e_{21}	e_{21}	e_{22}	0	0
	e_{22}	0	0	e_{21}	e_{22}

Table 2 $Y \times X$

		Y			
	Y $\times$ **X**	e_{11}	e_{12}	e_{21}	e_{22}
X	e_{11}	e_{11}	0	e_{21}	0
	e_{12}	e_{12}	0	e_{22}	0
	e_{21}	0	e_{11}	0	e_{21}
	e_{22}	0	e_{12}	0	e_{22}

If $m = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is a matrix and e is a number, we define the product em to be

$$em = e\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ea & eb \\ ec & ed \end{bmatrix}$$

We now have

- $m = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ae_{11} + be_{12} + ce_{21} + de_{22}$
- $\mu = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \alpha e_{11} + \beta e_{12} + \gamma e_{21} + \delta e_{22}$
- $m + \mu = (a+\alpha)e_{11} + (b+\beta)e_{12} + (c+\gamma)e_{21} + (d+\delta)e_{22}$
- $m\mu = (a\alpha + b\gamma)e_{11} + (a\beta + b\delta)e_{12} + (c\alpha + d\gamma)e_{21} + (c\beta + d\delta)e_{22}$

The set of **hyper-complex numbers** $ae_{11} + be_{12} + ce_{21} + de_{22}$, in which a, b, c, d range independently over a field F, and for which addition and multiplication are defined as above is an example of a **linear associative algebra** over F with the four units $e_{11}, e_{12}, e_{21}, e_{22}$ subject to the multiplication table (1)

Consider the set of n-tuples $(x_1, \ldots, x_n)$, whose coordinates $x_1 \ldots, x_n$ range independently over a given number field F.

Two n-tupes are called equal if their corresponding coordinates are equal. Addition and subtraction of n-tuples are defined by

$$(x_1, \ldots, x_n) \pm (y_1, \ldots, y_n) = (x_1 \pm y_1, \ldots, x_n \pm y_n) \tag{2}$$

The product of any number ρ of the field F and any n-tuple $x = (x_1, \ldots, x_n)$ is defined to be

$$\rho x = x\rho = (\rho x_1, \ldots, \rho x_n) \tag{3}$$

The n units are defined to be

$$e_1 = (1, 0, \ldots, 0)\ , e_2 = (0, 1, \ldots, 0),\ \ldots\ , e_n = (0, \ldots, 0, 1)$$

Hence any n-tuple can be expressed in the form

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

A linear algebra is obtained by assuming that any two *n*-tuples

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n \text{ and } y = y_1 e_1 + y_2 e_2 + \cdots + y_n e_n$$

can be combined by an operation called multiplication, which is subject to the distributive laws

$$x(y+z) = xy + xz \ , \ (y+z)x = yx + zx$$

Thus

$$xy = x_1 y_1 e_1 e_1 + x_1 y_2 e_1 e_2 + \cdots + x_i y_j e_i e_j + \cdots + x_n y_n e_n e_n$$

The product xy is determined once we know the particular products among the units, that is, for fixed i and j, the coordinates $\gamma_{ij1}, \gamma_{ij2}, \ldots, \gamma_{ijn}$ of $e_i e_j$;

$$e_i e_j = \gamma_{ij1} e_1 + \gamma_{ij2} e_2 + \cdots + \gamma_{ijn} e_n$$

Properties (2) and (3) of *n*-tuples give

$$x \pm y = (x_1 \pm y_1) e_1 + \cdots + (x_n \pm y_n) e_n \text{ and } \rho x = x \rho = (\rho x_1) e_1 + \cdots + (\rho x_n) e_n$$

Every linear algebra of dimension n is nothing more than a set of n^3 numbers γ_{ijk}, where i, j, k range independently over the integers $1, 2, \ldots, n$.

$n = 2, n^3 = 8$

- $e_1^2 = e_1 e_1 = \gamma_{111} e_1 + \gamma_{112} e_2$
- $e_1 e_2 = \gamma_{121} e_1 + \gamma_{122} e_2$
- $e_2 e_1 = \gamma_{211} e_1 + \gamma_{212} e_2$
- $e_2^2 = e_2 e_2 = \gamma_{221} e_1 + \gamma_{222} e_2$

$$\text{If } x = x_1 e_1 + x_2 e_2 \text{ and } y = y_1 e_1 + y_2 e_2$$

then the product xy has coordinates z_1, z_2 given by

$$z_1 = x_1 y_1 \gamma_{111} + x_1 y_2 \gamma_{121} + x_2 y_1 \gamma_{211} + x_2 y_2 \gamma_{221}$$

$$z_2 = x_1 y_1 \gamma_{112} + x_1 y_2 \gamma_{122} + x_2 y_1 \gamma_{212} + x_2 y_2 \gamma_{222}$$

That is,

$$xy = z_1 e_1 + z_2 e_2$$

January 13—Genetic motivation

(Ref.—Reed, "Algebraic Structure of Genetic Inheritance", 1997, pp. 107-108)

Before we discuss the mathematics of genetics, we need to acquaint ourselves with the necessary language from biology.

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**. E.g., the gene which determines blood type in humans has three different alleles, A, B, and O.

Since humans are **diploid** organisms (meaning we carry a double set of chromosomes—one from each parent), blood types are determined by two alleles.

Haploid cells (or organisms) carry a single set of chromosomes.

When diploid organisms reproduce, a process called **meiosis** produces **gametes** (sex cells) which carry a single set of chromosomes.

When these gamete cells fuse (e.g., when sperm fertilizes egg), the result is a **zygote**, which is again a diploid cell, meaning it carries its hereditary information in a double set of chromosomes.

When gametes fuse (or reproduce) to form zygotes a natural “multiplication” operation occurs.

As a natural first example, we consider simple Mendelian inheritance for a single gene with two alleles A and a.

In this case, two gametes fusing (or reproducing) to form a zygote gives the multiplication table shown in the following Table, which in freshman biology class might be called a Punnett square.

Table 3. Alleles passing from gametes to zygotes

	A	a
A	AA	Aa
a	aA	aa

The zygotes AA and aa are called **homozygous**, since they carry two copies of the same allele.

In this case, simple Mendelian inheritance means that there is no chance involved as to what genetic information will be inherited in the next generation; i.e., AA will pass on the allele A and aa will pass on a.

However, the zygotes Aa and aA (which are equivalent) each carry two different alleles. These zygotes are called **heterozygous**.

The rules of simple Mendelian inheritance indicate that the next generation will inherit either A or a with equal frequency. So, when two gametes reproduce, a multiplication is induced which indicates how the hereditary information will be passed down to the next generation.

This multiplication is given by the following rules:

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

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.

These rules are an algebraic representation of the rules of simple Mendelian inheritance. This multiplication table is shown in Table 4.

Table 4. Multiplication table of the gametic algebra 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 **phenotypes** (gene expression). Hence we have made no mention of the dominant or recessive properties of our alleles.

Now that we've defined a multiplication on the symbols A and a we can mathematically dedine the two dimensional algebra over R with basis $\{A, a\}$ and multiplication table as in Table 4. This algebra is called the **gametic algebra** for simple Mendelian inheritance with two alleles.

But gametic multiplication is just the beginning! In order for actual diploid cells (or organisms) to reproduce, they must first go through a reduction division process so that only one set of alleles is passed on.

For humans this occurs when males produce sperm and females produce eggs. When reproduction occurs, the hereditary information is then passed on via the gametic multiplication we've already defined.

Therefore, when two zygotes reproduce, another multiplication operation is formed taking into consideration both the reduction division process and gametic multiplication.

In our example of simple Mendelian inheritance for one gene with the two alleles A and a, zygotes have three possible genotypes: AA, aa,and Aa.

Let's consider the case of two zygotes both with genotype Aa reproducing. The reduction division process splits the zygote and passes on one allele for reproduction.

In the case of simple Mendelian inheritance the assumption is that both alleles will be passed on with equal frequency. Thus, half the time A gets passed on and half the time a does.

We represent this with the "frequency distribution" $\frac{1}{2}A + \frac{1}{2}a$. Therefore, symbolically $Aa \times Aa$ becomes

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

Formally multiplying these two expressions together results in

$$\frac{1}{4}AA + \frac{1}{2}Aa + \frac{1}{4}aa$$

using the notion that aA = Aa.

In this way, zygotic reproduction produces the multiplication table shown in Table 5. So we can define the three dimensional algebra over R with basis {AA, Aa, aa} and multiplication table as in Table 5. It is called the **zygotic algebra** for simple Mendelian inheritance with two alleles.

Table 5. Multiplication table of the zygotic algebra for simple Mendelian inheritance

	AA	Aa	aa
AA	AA	$\frac{1}{2}$(AA+Aa)	Aa
Aa	$\frac{1}{2}$(AA+Aa)	$\frac{1}{4}$AA+$\frac{1}{2}$Aa+$\frac{1}{4}$aa	$\frac{1}{2}$(Aa+aa)
aa	Aa	$\frac{1}{2}$(Aa+aa)	aa

The process of constructing a zygotic algebra from the original gametic algebra is called commutative duplication of algebras. We will discuss this process from a mathematical perspective later.

Now that we've seen how the gametic and zygotic algebras are formed in the most basic example, we shall begin to consider the mathematical (and indeed, algebraic) structure of such algebras.

The Nonassociativity of Inheritance

Depending on the "population" you are concerned with, a general element $\alpha A+\beta a$ of the gametic algebra which satisfies $0\leq\alpha,\beta\leq 1$ and $\alpha+\beta=1$ can represent a population, a single individual of a population, or a single gamete.

In each case, the coefficients α and β signify the percentage of frequency of the associated allele. I.e., if the element represents a population, then α is the percentage of the population which carries the allele A on the gene under consideration. Likewise, β is the percentage of the population which has the allele a.

For those elements of the gametic and zygotic algebras which represent populations, multiplication of two such elements represents random mating between the two populations.

It seems logical that the order in which populations mate is significant. I.e., if population P mates with population Q and then the resulting 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, $(P \times Q) \times R$ is not equal to $P \times (Q \times R)$.

So, we see that from a purely biological perspective, we should expect that the algebras which arise in genetics will not satisfy the associative property.

Now, if we study the multiplication tables of both the gametic and zygotic algebras for simple Mendelian inheritance, we will notice immediately that the algebras are commutative.

From a biological perspective, if populations P and Q are mating, it makes no difference whether you say P mates with Q or Q mates with P!

However, as we should expect, these algebras do not satisfy the associative property.

E.g., in the gametic algebra apply the rules of multiplication and the distributive property to see that $A \times (A \times a) = \frac{3}{4}A + \frac{1}{4}a$. However,
$(A \times A) \times a = A \times a = \frac{1}{2}A + \frac{1}{2}a$

Hence, the associative property does not hold for the gametic algebra.

The same is true for the zygotic algebra.

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

January 20—Genetic algebras and derivations

"There is nothing like going to the original sources"
(Ref.—Etherington, "Genetic Algebras," 1939, §1 pp.242–243,§6,7 249–251)

§1 pp.242–243

"The mechanism of chromosome inheritance, in so far as it determines the probability distributions of genetic types in families and filial generations, and expresses itself through their frequency distributions, may be represented conveniently by algebraic symbols."

A population (i.e. a distribution of genetic types) is represented by a normalized hypercomplex number in one or another algebra.

If P and Q are populations, the filial generation $P \times Q$ (i.e. the statistical population of offspring resulting from the random mating of individuals of P with individuals of Q) is obtained by multiplying representations of P and Q.

A population may mean a single individual, or rather the information which we may have concerning him in the form of a probability distribution.

Gametic algebras

§6 249–250

Consider the inheritance of traits depending on any number of gene differences at any number of loci on any number of chromosomes in a diploid or generally autopolyploid species.

Let $G_1, \ldots, G_n$ denote the set of gametic types determined by these gene differences. There will be $n(n+1)/2$ zygotic types G_iG_j $(= G_jG_i)$

The formulae giving the series of gametic types by each individual (zygote), and their frequencies may be written $G_iG_j = \gamma_{1ij}G_1 + \cdots + \gamma_{kij}G_k + \cdots + \gamma_{nij}G_n$, with the normalizing conditions $\gamma_{1ij} + \cdots + \gamma_{kij} + \cdots + \gamma_{nij} = 1$ and $0 \leq \gamma_{kij} \leq 1$

γ_{kij} is the probability that an arbitrary gamete produced by an individual of zygotic type G_iG_j is of type G_k

A population P which produces gametes G_k in proportions α_k may be represented by writing $P = \alpha_1G_1 + \cdots + \alpha_kG_k + \cdots + \alpha_nG_n$, with the normalizing condition $\alpha_1 + \cdots + \alpha_k + \cdots + \alpha_n = 1$

A population may also be described by the proportions of the zygotic types which it contains, so we may write $P = \alpha_{11}G_1G_1 + \cdots + \alpha_{ij}G_iG_j + \cdots + \alpha_{nn}G_nG_n$, with the normalizing condition $\alpha_{11} + \cdots + \alpha_{ij} + \cdots + \alpha_{nn} = 1 \qquad (\alpha_{ij} = \alpha_{ji})$

If two populations $P = \alpha_1G_1 + \cdots + \alpha_kG_k + \cdots + \alpha_nG_n$, and $Q = \beta_1G_1 + \cdots + \beta_kG_k + \cdots + \beta_nG_n$ intermate at random, representations of the first filial generation are obtained by multiplying P and Q

The population of offspring is then

$$PQ = \alpha_1\beta_1G_1G_1 + \cdots + \alpha_i\beta_jG_iG_j + \cdots + \alpha_n\beta_nG_nG_n$$

The linear algebra with basis $G_1, \ldots, G_k, \ldots G_n$ and multiplication table

$$G_iG_j = \gamma_{1ij}G_1 + \cdots + \gamma_{kij}G_k + \cdots + \gamma_{nij}G_n$$

is called the **gametic algebra** for the type of inheritance considered

Zygotic Algebras

§7 250–251

When individuals of type G_iG_j and G_lG_m mate, the probability distribution of zygotic types in their offspring can be obtained by multiplying the gametic representations given by $G_iG_j = \gamma_{1ij}G_1 + \cdots + \gamma_{kij}G_k + \cdots + \gamma_{nij}G_n$ and $G_lG_m = \gamma_{1lm}G_1 + \cdots + \gamma_{klm}G_k + \cdots + \gamma_{nlm}G_n$

We obtain $G_iG_j \times G_lG_m = \gamma_{1ij}\gamma_{1lm}G_1G_1 + \cdots + \gamma_{\sigma ij}\gamma_{\tau lm}G_\sigma G_\tau + \cdots + \gamma_{nij}\gamma_{nlm}G_nG_n$

Or, writing $Z_{ij} = G_iG_j$ to emphasize the union of paired gametes into single individuals,

$$Z_{ij}Z_{lm} = \gamma_{1ij}\gamma_{1lm}Z_{11} + \cdots + \gamma_{\sigma ij}\gamma_{\tau lm}Z_{\sigma\tau} + \cdots + \gamma_{nij}\gamma_{nlm}Z_{nn} \tag{4}$$

The linear algebra with basis $Z_{11}, \ldots, Z_{ij}, \ldots Z_{nn}$ and multiplication table (4) is called the **zygotic algebra** for the type of inheritance considered

(Ref.—Etherington, "Non-associative algebra and the symbolism of genetics," 1941, §1 p.24, §2 pp.25–26, §5 pp.29–30, §8 pp.34–35)

§1 p.24

"The statistical material of genetics usually consists of frequency distributions—of genes, zygotes and mating couples—from which new distributions referring to their progeny arise."

"Combination of distributions by random mating is usually symbolized by the mathematical sign for multiplication; but this sign is not taken literally for the simple reason that the general laws connecting the distributions of progenitors and progeny are inconsistent with the laws governing multiplication in algebra."

"However, there is no insuperable reason why the genetical sign of multiplication should not be taken literally; for it is possible with any particular type of inheritance to construct an 'algebra'—distinct from ordinary algebra but of a type well known to mathematicians—such that the laws governing multiplication shall represent exactly the underlying genetical situation."

"These 'genetic algebras' are of a kind known as 'linear algebras.' "

Genetical Multiplication

§2 pp.25–26

P denotes a frequency distribution or a probability distribution of a population, a single individual, or a single gamete.

- $P = DD =$ homozygous dominant individual, or population consisting of such
- $P = \alpha DD + \beta DR + \gamma RR =$ population with assigned frequencies α, β, γ of genotypes, or individual with assigned probabilities α, β, γ of belonging to one of the genotypes
- $P = \delta D + \rho R =$ population which produces D and R gametes in the given numerical ratio, or gamete which has probability δ of containing D and probability ρ of containing R

The multiplication of populations (individuals, gametes) means the calculation of progeny distribution resulting from random mating (mating, fusion).

The distributive ($P(Q + R) = PQ + PR$) and commutative ($PQ = QP$) laws are valid in genetical multiplication. The associative law is not: $(P(QR) \neq (PQ)R)$.

Mendelian Gametic and Zygotic Algebras—Revisited

§5 pp.29–30

Consider a pair of autosomal allelomorphs D, R and the corresponding genotypes of zygotes $A = DD \quad , \quad B = DR \quad , \quad C = RR$

In accordance with Mendelian principles, we have

- gametes produced by each type of zygote (**gametic algebra**)

  $$D^2 = DD = D \quad , \quad DR = \frac{1}{2}D + \frac{1}{2}R \quad , \quad R^2 = RR = R$$

  (heterozygote DR produces D and R gametes in equal numbers)
- zygotes produced by each type of mating couple (**zygotic algebra**)

  $$A^2 = A \quad , B^2 = \frac{1}{4}A + \frac{1}{2}B + \frac{1}{4}C \quad , \quad C^2 = C$$

  $$BC = \frac{1}{2}B + \frac{1}{2}C \quad , \quad CA = B \quad , \quad AB = \frac{1}{2}A + \frac{1}{2}B$$

  (the offspring of a mating $DR \times DR$ are 25%DD, 50%DR, 25%RR)

Self Fertilization

§8 pp.34–35 (See also Reed, §5.2 pp.121–122)

Starting from the zygotic distribution $P = \alpha A + \beta B + \gamma C$, where $a = DD, B = DR, C = RR$, if mating proceeds in successive generations by self-fertilization, or by each individual mating with another of the same type, the first filial generation F_1 will consist of the offspring of $A \times A, B \times B, C \times C$, occurring in proportions $\alpha : \beta : \gamma$ so that

$$\begin{aligned} F_1 &= \alpha A^2 + \beta B^2 + \gamma C^2 \\ &= \alpha A + \beta(\frac{1}{4}A + \frac{1}{2}B + \frac{1}{4}C) + \gamma C \\ &= (\alpha + \frac{1}{4}\beta)A + \frac{1}{2}\beta B + (\frac{1}{4}\beta + \gamma)C \end{aligned}$$

The second filial generation $F_2 = F_1 \times F_!$ calculates to

$$F_2 = (\alpha + \frac{3}{8}\beta)A + \frac{1}{4}\beta B + (\frac{3}{8}\beta + \gamma)C$$

QUESTION: What is the n^{th}-filial generation under self-fertilization?

$$F_n = \alpha_n A + \beta_n B + \gamma_n C$$

- $\alpha_1 = \alpha + \frac{1}{4}\beta,\ \alpha_2 = \alpha + \frac{3}{8}\beta$
- $\beta_1 = \frac{1}{2}\beta.\ \beta_2 = \frac{1}{4}\beta$
- $\gamma_1 = \frac{1}{4}\beta + \gamma,\ \gamma_2 = \frac{3}{8}\beta + \gamma$

ANSWER: $\alpha_n = \alpha + \frac{1}{2}\beta - \frac{1}{2^{n+1}}\beta, \qquad \beta_n = \frac{1}{2^n}\beta, \qquad \gamma_n = \frac{1}{2}\beta + \gamma - \frac{1}{2^{n+1}}\beta$

Or, $$F_n = (\alpha + \frac{1}{2}\beta - \frac{1}{2^{n+1}}\beta)A + \frac{1}{2^n}\beta B + (\frac{1}{2}\beta + \gamma - \frac{1}{2^{n+1}}\beta)C$$

The equilibrium distribution is thus

$$F_n = (\alpha + \frac{1}{2}\beta)A + (\frac{1}{2}\beta + \gamma)C$$

Repeated self fertilization kills off the heterozygotes!

Derivations on Linear Algebras

(Ref.—Russo, “Playing havoc with the product rule” Transfer Seminar Fall 2012)

Much of the algebra taught in the undergraduate curriculum, such as linear algebra (**vector spaces, matrices**), modern algebra (**groups, rings, fields**), number theory (**primes, congruences**) is concerned with systems with one or more associative binary products.

For example, addition and multiplication of matrices is associative:

$$A + (B + C) = (A + B) + C$$

$$A(BC) = (AB)C.$$

In the early 20th century, physicists started using the product A.B for matrices, defined by

$$A.B = AB + BA,$$

and called the Jordan product (after the physicist **Pascual Jordan 1902-1980**), to model the observables in quantum mechanics.

Also in the early 20th century both mathematicians and physicists used the product [A,B], defined by

$$[A, B] = AB - BA$$

and called the Lie product (after the mathematician **Sophus Lie 1842-1899**), to study differential equations.

Neither one of these products is associative, so they each give rise to what is called a nonassociative algebra, in these cases, called **Jordan algebras** and **Lie algebras** respectively.

Sophus Lie (1842–1899)

Marius Sophus Lie was a Norwegian mathematician. He largely created the theory of continuous symmetry, and applied it to the study of geometry and differential equations.

Pascual Jordan (1902–1980)

Pascual Jordan was a German theoretical and mathematical physicist who made significant contributions to quantum mechanics and quantum field theory.

Abstract theories of these algebras and other nonassociative algebras were subsequently developed and have many other applications, for example to **cryptography** and **genetics**, to name just two.

Lie algebras are especially important in **particle physics**.

The derivative

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

DIFFERENTIATION IS A LINEAR PROCESS

$$(f+g)' = f' + g'$$

$$(cf)' = cf'$$

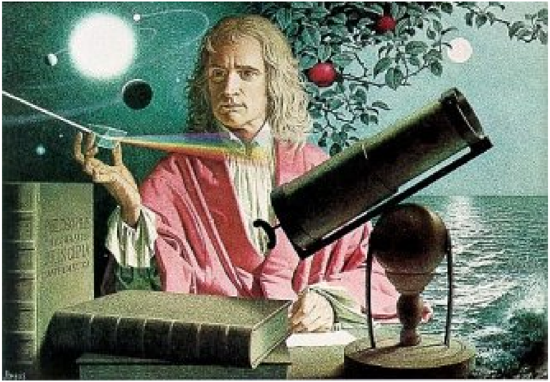
THE SET OF DIFFERENTIABLE FUNCTIONS FORMS AN ALGEBRA $\mathcal{D}$

$$(fg)' = fg' + f'g$$

(product rule)

HEROS OF CALCULUS

#1 Sir Isaac Newton (1642-1727)



Isaac Newton was an English physicist, mathematician, astronomer, natural philosopher, alchemist, and theologian, and is considered by many scholars and members of the general public to be one of the most influential people in human history.

#2 Gottfried Wilhelm Leibniz (1646-1716)

Gottfried Wilhelm Leibniz was a German mathematician and philosopher. He developed the infinitesimal calculus independently of Isaac Newton, and Leibniz's mathematical notation has been widely used ever since it was published.

LEIBNIZ RULE

$$(fg)' = f'g + fg'$$

(order changed)

More generally

$$(fgh)' = f'gh + fg'h + fgh'$$

$$(f_1 f_2 \cdots f_n)' = (f_1' f_2 \cdots f_n) + \cdots + (f_1 \cdots f_i' \cdots f_n) + \cdots + (f_1 f_2 \cdots f_n')$$

The chain rule,

$$(f \circ g)'(x) = f'(g(x))g'(x)$$

plays no role in this seminar. Neither does the quotient rule

$$(f/g)' = \frac{gf' - fg'}{g^2}$$

CONTINUITY: $x_n \to x \Rightarrow f(x_n) \to f(x)$

THE SET OF CONTINUOUS FUNCTIONS FORMS AN ALGEBRA $\mathcal{C}$
(sums, constant multiples and products of continuous functions are continuous)

$\mathcal{D}$ and $\mathcal{C}$ ARE EXAMPLES OF ALGEBRAS WHICH ARE BOTH **ASSOCIATIVE** AND **COMMUTATIVE**

PROPOSITION

EVERY DIFFERENTIABLE FUNCTION IS CONTINUOUS
$\mathcal{D}$ is a subalgebra of $\mathcal{C}$; $\mathcal{D} \subset \mathcal{C}$

DIFFERENTIATION IS A LINEAR PROCESS. LET US DENOTE IT BY D AND WRITE $D(f)$ (or Df) for f'

$$D(f+g) = Df + Dg$$

$$D(cf) = cDf$$

$$D(fg) = (Df)g + f(Dg)$$

$$D(f/g) = \frac{g(Df) - f(Dg)}{g^2}$$

DEFINITION

A DERIVATION ON AN ALGEBRA $\mathcal{A}$ IS A LINEAR PROCESS $\delta : \mathcal{A}$ SATISFYING THE LEIBNIZ RULE:

$$\delta(x+y) = \delta(x) + \delta(y)$$

$$\delta(cx) = c\delta(x)$$

$$\delta(xy) = \delta(x)y + x\delta(y)$$

THEOREM

There are no (non-zero) derivations on $\mathcal{C}$.

In other words, every derivation of $\mathcal{C}$ is identically zero

COROLLARY

$\mathcal{D} \neq \mathcal{C}$

(NO DUUUH! $f(x) = |x|$)

DERIVATIONS ON THE SET OF MATRICES

THE SET $M_n(\mathbb{R})$ of n by n MATRICES IS AN ALGEBRA UNDER
MATRIX ADDITION $A + B$
MATRIX MULTIPLICATION $A \times B$
WHICH IS ASSOCIATIVE BUT NOT COMMUTATIVE.

DEFINITION

A DERIVATION ON $M_n(\mathbb{R})$ WITH RESPECT TO MATRIX MULTIPLICATION IS A LINEAR PROCESS δ WHICH SATISFIES THE LEIBNIZ RULE $\delta(A \times B) = \delta(A) \times B + A \times \delta(B)$.

EXAMPLE

FIX A MATRIX A in $M_n(\mathbb{R})$ AND DEFINE

$$\delta_A(X) = A \times X - X \times A.$$

THEN δ_A IS A DERIVATION WITH RESPECT TO MATRIX MULTIPLICATION (WHICH CAN BE NON-ZERO)

THEOREM

EVERY DERIVATION ON $M_n(\mathbb{R})$ WITH RESPECT TO MATRIX MULTIPLICATION IS OF THE FORM δ_A FOR SOME A IN $M_n(\mathbb{R})$.

CLOSING REMARKS (for today)

- If A is any ASSOCIATIVE algebra, and a is any element of A, then the linear process δ_a defined by $\delta_a(x) = ax - xa$ is a derivation of the algebra A.
- In many (but not all) associative algebras, these are the only derivations.
- What about non-associative algebras? In particular, Lie algebras, Jordan algebras, genetic algebras?

January 27—Lie Algebras; Bracket Product

DEFINITION THE BRACKET PRODUCT ON THE SET $M_n(\mathbb{R})$ OF MATRICES IS DEFINED BY $[X, Y] = X \times Y - Y \times X$

THE SET $M_n(\mathbb{R})$ of n by n MATRICES IS AN ALGEBRA UNDER MATRIX ADDITION AND BRACKET MULTIPLICATION, WHICH IS NOT ASSOCIATIVE AND NOT COMMUTATIVE.

A DERIVATION ON $M_n(\mathbb{R})$ WITH RESPECT TO BRACKET MULTIPLICATION IS A LINEAR PROCESS δ WHICH SATISFIES THE LEIBNIZ RULE $\delta([A, B]) = [\delta(A), B] + [A, \delta(B)]$

EXAMPLE FIX A MATRIX A in $M_n(\mathbb{R})$ AND DEFINE $\delta_A(X) = [A, X] = A \times X - X \times A$. THEN δ_A IS A DERIVATION WITH RESPECT TO BRACKET MULTIPLICATION

THEOREM EVERY DERIVATION ON $M_n(\mathbb{R})$ WITH RESPECT TO BRACKET MULTIPLICATION IS OF THE FORM δ_A FOR SOME A IN $M_n(\mathbb{R})$.

THE CIRCLE PRODUCT ON THE SET OF MATRICES

DEFINITION THE CIRCLE PRODUCT ON THE SET $M_n(\mathbb{R})$ OF MATRICES IS DEFINED BY $X \circ Y = (X \times Y + Y \times X)/2$

THE SET $M_n(\mathbb{R})$ of n by n MATRICES IS AN ALGEBRA UNDER MATRIX ADDITION AND CIRCLE MULTIPLICATION, WHICH IS COMMUTATIVE BUT NOT ASSOCIATIVE.

A DERIVATION ON $M_n(\mathbb{R})$ WITH RESPECT TO CIRCLE MULTIPLICATION IS A LINEAR PROCESS δ WHICH SATISFIES THE LEIBNIZ RULE

$$\delta(A \circ B) = \delta(A) \circ B + A \circ \delta(B)$$

EXAMPLE FIX A MATRIX A in $M_n(\mathbb{R})$ AND DEFINE $\delta_A(X) = A \times X - X \times A$. THEN δ_A IS A DERIVATION WITH RESPECT TO CIRCLE MULTIPLICATION

THEOREM EVERY DERIVATION ON $M_n(\mathbb{R})$ WITH RESPECT TO CIRCLE MULTIPLICATION IS OF THE FORM δ_A FOR SOME A IN $M_n(\mathbb{R})$.

IT IS TIME FOR A SUMMARY OF THE PRECEDING

Table 6

matrix	bracket	circle
$ab = a \times b$	$[a, b] = ab - ba$	$a \circ b = ab + ba$
$\delta_a(x)$ $=$ $ax - xa$	$\delta_a(x)$ $=$ $ax - xa$	$\delta_a(x)$ $=$ $ax - xa$

AXIOMATIC APPROACH

AN ALGEBRA IS DEFINED TO BE A SET (ACTUALLY A VECTOR SPACE) WITH TWO BINARY OPERATIONS, CALLED ADDITION AND MULTIPLICATION

ADDITION IS DENOTED BY $a + b$ AND IS REQUIRED TO BE COMMUTATIVE AND ASSOCIATIVE

$$a + b = b + a, \quad (a + b) + c = a + (b + c)$$

MULTIPLICATION IS DENOTED BY ab AND IS REQUIRED TO BE DISTRIBUTIVE WITH RESPECT TO ADDITION

$$(a + b)c = ac + bc, \quad a(b + c) = ab + ac$$

AN ALGEBRA IS SAID TO BE ASSOCIATIVE (RESP. COMMUTATIVE) IF THE **MULTIPLICATION** IS ASSOCIATIVE (RESP. COMMUTATIVE) (RECALL THAT ADDITION IS ALWAYS COMMUTATIVE AND ASSOCIATIVE)

THE ALGEBRAS $\mathcal{C}$, $\mathcal{D}$ AND $M_n(\mathbb{R})$ ARE EXAMPLES OF ASSOCIATIVE ALGEBRAS.

$\mathcal{C}$ AND $\mathcal{D}$ ARE COMMUTATIVE, AND $M_n(\mathbb{R})$ IS NOT COMMUTATIVE.

THE AXIOM WHICH CHARACTERIZES ASSOCIATIVE ALGEBRAS IS $a(bc) = (ab)c$. THESE ARE CALLED **ASSOCIATIVE ALGEBRAS**

THE AXIOM WHICH CHARACTERIZES COMMUTATIVE ALGEBRAS IS $ab = ba$. THESE ARE CALLED (you guessed it) **COMMUTATIVE ALGEBRAS**

HOWEVER, THESE TWO CONCEPTS ARE TOO GENERAL TO BE OF ANY USE BY THEMSELVES

THE AXIOMS WHICH CHARACTERIZE BRACKET MULTIPLICATION ARE

$$a^2 = 0 \text{ and } (ab)c + (bc)a + (ca)b = 0$$

THESE ARE CALLED **LIE ALGEBRAS**

THE AXIOMS WHICH CHARACTERIZE CIRCLE MULTIPLICATION ARE

$$ab = ba \text{ and } a(a^2b) = a^2(ab)$$

THESE ARE CALLED **JORDAN ALGEBRAS**

Sophus Lie (1842–1899)

Marius Sophus Lie was a Norwegian mathematician. He largely created the theory of continuous symmetry, and applied it to the study of geometry and differential equations.

Pascual Jordan (1902–1980)

Pascual Jordan was a German theoretical and mathematical physicist who made significant contributions to quantum mechanics and quantum field theory.

LET'S SUMMARIZE AGAIN

Table 7—ALGEBRAS

commutative algebras
$ab = ba$

associative algebras
$a(bc) = (ab)c$

Lie algebras
$a^2 = 0$
$(ab)c + (bc)a + (ca)b = 0$

Jordan algebras
$ab = ba$
$a(a^2b) = a^2(ab)$

Closing Remark

Given any algebra A of any kind (associative, Lie, Jordan, genetic, you name it) , the set of all derivations on A is a Lie algebra with the bracket given by

$$[\delta_1, \delta_2] = \delta_1\delta_2 - \delta_2\delta_1$$

For the record, if A is an algebra with product denoted by xy, δ is a derivation if

- $\delta(x+y) = \delta(x) + \delta(y)$
- $\delta(xy) = x\delta y + (\delta x)y$

For any two linear transformations S and T on A, their product ST is defined by

$$ST\ (x) = S(T(x))$$

In particular, $\delta_1\delta_2(x) = \delta_1(\delta_2(x))$

To convince yourself that the remark is true, you have to show that

$$[\delta_1, \delta_2](xy) = x\left([\delta_1, \delta_2](y)\right) + \left([\delta_1, \delta_2](x)\right) y$$

February 3—Evolution Algebras

First, a digression: Proof of an earlier theorem

Matrix units

Let $E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

LEMMA

- $E_{11} + E_{22} = I$
- $E_{ij}^t = E_{ji}$
- $E_{ij}E_{kl} = \delta_{kl}E_{il}$

THEOREM

Let $\delta : M_2 \to M_2$ be a derivation: δ is linear and $\delta(AB) = A\delta(B) + \delta(A)B$. Then there exists a matrix K such that $\delta(X) = XK - KX$ for X in M_2.

PROOF OF THEOREM

$$\begin{aligned}
0 &= \delta(1) = \delta(E_{11}+E_{22}) = \delta(E_{11}) + \delta(E_{22}) \\
&= \delta(E_{11}E_{11}) + \delta(E_{21}E_{12}) \\
&= E_{11}\delta(E_{11}) + \delta(E_{11})E_{11} + E_{21}\delta(E_{12}) + \delta(E_{21})E_{12} \\
&= E_{11}\delta(E_{11}) + E_{21}\delta(E_{12}) + \delta(E_{11})E_{11} + \delta(E_{21})E_{12}.
\end{aligned}$$

Let $K = E_{11}\delta(E_{11}) + E_{21}\delta(E_{12}) = -\delta(E_{11})E_{11} - \delta(E_{21})E_{12}$. Then

- $KE_{11} = -\delta(E_{11})E_{11} \quad , \quad E_{11}K = E_{11}\delta(E_{11})$
- $KE_{12} = -\delta(E_{11})E_{12} \quad , \quad E_{12}K = E_{11}\delta(E_{12})$
- $KE_{21} = -\delta(E_{21})E_{11} \quad , \quad E_{21}K = E_{21}\delta(E_{11})$
- $KE_{22} = -\delta(E_{21})E_{12} \quad , \quad E_{22}K = E_{21}\delta(E_{12})$

- $E_{11}K - KE_{11} = E_{11}\delta(E_{11}) + \delta(E_{11})E_{11} = \delta(E_{11}E_{11}) = \delta(E_{11})$
- $E_{12}K - KE_{12} = E_{11}\delta(E_{12}) + \delta(E_{11})E_{12} = \delta(E_{11}E_{12}) = \delta(E_{12})$
- $E_{21}K - KE_{21} = E_{21}\delta(E_{11}) + \delta(E_{21})E_{11} = \delta(E_{21}E_{11}) = \delta(E_{21})$
- $E_{22}K - KE_{22} = E_{21}\delta(E_{12}) + \delta(E_{21})E_{12} = \delta(E_{21}E_{12}) = \delta(E_{22})$ Q.E.D.

Asexual propagation (Tian LNIM §2.1.1, p. 9)

- Prokaryotes are nonsexual reproductive organisms.
- Prokaryotic cells, unlike eukaryotic cells, do not have nuclei.
- In prokaryote inheritance, there is no mitosis and meiosis.
- Instead, prokaryotes reproduce by binary fission.

- The genetic information passed from one generation to the next should be conserved because of the strictness of DNA self-replication.
- However, there are still many possible factors in the environment that can induce the change of genetic information from generation to generation.
- The inheritance of prokaryotes is then not Mendelian.

- Now, lets mathematically formulate the asexual reproduction process. Suppose that we have n genetically distinct prokaryotes, denoting them by $p_1, \ldots, p_n$.
- We also suppose that the same environmental conditions are maintained from generation to generation. We look at changes in gene frequencies over two generations. To this end, we can set the following relations:

$$p_i \cdot p_i = c_{i1} p_1 + \cdots + c_{in} p_n \quad , \quad p_i \cdot p_j = 0,\ i \neq j$$

Gametic algebras in asexual inheritance (Tian LNIM §2.1.2, pp. 10–11)

- Consider an infinitely large, randomly mating population of diploid individuals.
- Let $a_1, \ldots, a_n$ be the genetically distinct gametes produced by the population.
- By random union of gametes a_i and a_j, zygotes of type $a_i a_j$ are formed.
- Assume that a zygote $a_i a_j$ produces a number γ_{ijk} of gametes of type a_k, which survive in the next generation,

- Now, the gamete algebra is defined on the linear space spanned by these gametes $a_1, \ldots, a_n$ by the following multiplication table: $a_i a_j = \gamma_{ij1} a_1 + \cdots \gamma_{ijn} a_n$
- In asexual inheritance, the interpretation $a_i a_j$ as a zygote does not make sense biologically if $a_i \neq a_j$, but $a_i a_i$ can still be interpreted as self-replication.
- Mathematically, we set $a_i a_j = 0$. Therefore, in asexual inheritance, we can use the following relations to define an algebra:

$$a_i \cdot a_i = \gamma_{i1} a_1 + \cdots + \gamma_{in} a_n \quad , \quad a_i \cdot a_j = 0,\ i \neq j$$

- Of course, this case is not of Mendelian inheritance.

Evolution algebras

Definition (Tian LNIM §3.1.1, p.20)

Let A be an algebra. If it admits a basis $x_1, x_2, \ldots, x_n$ such that $x_i \cdot x_i = a_{i1}x_1 + \cdots + a_{in}x_n \quad , \quad x_i \cdot x_j = 0,\ i \neq j$ we then call this algebra an **evolution algebra**. We call the basis a **natural basis**.

Some basic Properties (Tian LNIM §3.1.1, pp.20–21)

- Evolution algebras are not associative, in general.
- Evolution algebras are commutative
- Evolution algebras are not power-associative, in general.
- The direct sum of evolution algebras is also an evolution algebra.

More definitions (Tian LNIM §3.1.3, p.23)

Let A be a commutative algebra and a an element of A

- The **principal powers** of a are:
  $a^1 = a, \quad a^2 = a \cdot a, \quad a^3 = a^2 \cdot a, \quad \ldots \qquad a^n = a^{n-1} \cdot a$
- The **plenary powers** of a of A are:
  $a^{[1]} = a^2 = a \cdot a, \quad a^{[2]} = a^2 \cdot a^2, \quad a^{[3]} = a^{[2]} \cdot a^{[2]}, \quad \ldots \qquad a^{[n]} = a^{[n-1]} \cdot a^{[n-1]}$

Example (Tian LNIM §3.1.3, p.24)

Let E be an evolution algebra with basis e_1, e_2, e_3 and multiplication defined by $e_1 e_1 = e_1 + e_2, \quad e_2 e_2 = -e_1 - e_2, \quad e_3 e_3 = -e_2 + e_3.$

Let $u_1 = e_1 + e_2, \quad u_2 = e_1 + e_3$ and let F be the set of all $\alpha u_1 + \beta u_2$. Then F is a subalgebra of E but it is not an evolution algebra, since one can construct distinct elements v_1 and v_2 of F such that $v_1 v_2 \neq 0$.

History of general genetic algebras (Tian LNIM §5.1, p.92)

General genetic algebras are the product of interactions between biology and mathematics. The study of these algebras reveals the algebraic structures of Mendelian genetics, which can simplify and shorten the way to understand genetic and evolutionary phenomena.

However, after Baur [1909] and Correns [1909] first detected that chloroplast inheritance departed from Mendels rules, and much later, mitochondrial gene inheritance were also identified in the same way, non-Mendelian inheritance of organelle genes became manifest with two features—uniparental inheritance and vegetative segregation.

Non-Mendelian genetics is now a basic language of molecular geneticists.

When we try to formulate non-Mendelian genetics as algebras, we at least need a new idea to formulate reproduction in non-Mendelian genetics as multiplication in algebras. Evolution algebras stems from this new idea.

Terms in population genetics (Tian LNIM §5.2.1, p.93)

Organisms with a double set of chromosomes are called diploid organisms. For example, humans are diploid. Organisms with one set of chromosomes are called haploid organisms. For instant, most fungi and a few algae are haploid organisms.

The different variants of a gene are referred to as alleles. Biologists refer to individuals with two identical copies of a gene as being homozygous; and individuals with two different copies of the same gene as being heterozygous.

Reproduction of organisms can take place by asexual or sexual processes.

Asexual reproduction involves the production of a new individual(s) from cells or tissues of a preexisting organism. This process is common in plants and in many microorganisms. It can involve simple binary fission in unicellular microbes or the production of specialized asexual spores.

Sexual reproduction differs, in that it involves fusion of cells (gametes) derived from each parent, to form a zygote.

Asexual reproduction allows some genetic changes in offspring by chance. The genetic processes involved in the production of gametes also allow for some genetic changes from generation to generation. Sexual reproduction is limited to species that are diploid or have a period of their life cycle in the diploid state.

The division of somatic cells is called mitosis; and the division of meiotic cells is called meiosis.

Prokaryote chromosomes consist of a single DNA, which is usually circular, with only a small amount of associated protein. Eukaryotes have several linear chromosomes, and the DNA is tightly associated with large amounts of protein.

Organelle (http://en.wikipedia.org/wiki/Organelle)

• In cell biology, an organelle is a specialized subunit within a cell that has a specific function.
• The name organelle comes from the idea that these structures are to cells what an organ is to the body (the suffix -elle being a diminutive).
• There are many types of organelles, particularly in eukaryotic cells. While prokaryotes do not possess organelles per se, some do contain protein-based microcompartments, which are thought to act as primitive organelles.

Mendelian vs. non-Mendelian (Tian LNIM §5.2.2, pp.94–95)

Although most of heredity of nuclear genes obeys Mendels laws, the inheritance of organelle is not Mendelian. There are five aspects in comparison of Mendelian genetics and non-Mendelian genetics:

1 During asexual reproduction, alleles of nuclear genes do not segregate: heterozygous cells produce heterozygous daughters. In contrast, alleles of organelle genes in heteroplasmic cells segregate during mitotic and meiotic divisions to produce homoplasmic cells.

2 Alleles of a nuclear gene always segregate during meiosis, with half of the gametes receiving one allele and half the other. Alleles of organelle genes may or may not segregate during meiosis; the mechanisms are the same as for vegetative segregation.

3 Inheritance of nuclear genes is biparental. Organelle genes are often inherited from only one parent, uniparental inheritance.

4 Alleles of different nuclear genes segregate independently. Organelle genes are nearly always on a single chromosome and recombination is often severely limited by uniparental inheritance or failure of organelles to fuse and exchange genomes.

5 Fertilization is random with respect to the genotype of the gametes. This is the only part of Mendels model that applies to organelle as well as nuclear genes.

(For the basics of organelle biology, see Tian LNIM pp. 94–95)

Algebraic formulation of non-Mendelian genetics (Tian LNIM §5.2.3, pp.95–96)

Let us consider a population of organelles in a cell or a cell clone, and suppose that there are n different genotypes in this organelle population. Denote these genotypes by $g_1, g_2, \ldots, g_n$.

According to the point (3) in Subsection 5.2.2, the crossing of genotypes is impossible since it is uniparental inheritance. Mathematically, we set $g_i \cdot g_j = 0$ for $i \neq j$

According to the point (2) in Subsection 5.2.2, alleles of organelle genes may or may not segregate during meiosis following vegetative segregation, so the frequency of each gene in the next generation could be variant.

According to the point (4) in Subsection 5.2.2, intramolecular and intermolecular recombination within a lineage provides evidence that one organelle genotype could produce other different genotypes.

Therefore, we can mathematically define, $g_i^2 = \alpha_{i1}g_1 + \cdots + \alpha_{in}g_n$ where α_{ij} is a positive number that can be interpreted as the rate of genotype g_j produced by genotype g_i.

Now, we have the algebra defined by generators $g_1, g_2, \ldots g_n$ which are subject to these relations.

Obviously, this is a very general definition. But it is general enough to include all non-Mendelian inheritance phenomena.

February 10—Algebras with Genetic Significance

(Ref. REED, pp.111–114)

Mathematically, the algebras that arise in genetics (via gametic, zygotic, or copular algebras) are very interesting structures. They are generally commutative but nonassociative, yet they are not necessarily Lie, Jordan, or alternative algebras.

In addition, many of the algebraic properties of these structures have genetic significance. Indeed, it is the interplay between the purely mathematical structure and the corresponding genetic properties that makes this subject so fascinating.

We turn our attention now to the more formal mathematical study of the underlying algebraic structure of genetics. The most general algebra which could have genetic significance is an "algebra with genetic realization."

An **algebra with genetic realization** is an algebra A over the real numbers R which has a basis $a_1, \dots a_n$ and a multiplication table $a_i a_j = \sum_{k=1}^n \gamma_{ijk} a_k$ such that $0 \le \gamma_{ijk} \le 1$ and $\sum_{k=1}^n \gamma_{ijk} = 1$. Such a basis is called the natural basis for A.

It is easy to see that our earlier examples of gametic and zygotic algebras for simple Mendelian inheritance, as well as the general gametic and zygotic algebras, are all algebras with genetic realization.

In a general algebra A with genetic realization, an element x in A represents a population, or a gene pool for a population, if its expression as a linear combination of the basis elements

$$x = \xi_1 a_1 + \cdots + \xi_n a_n$$

satisfies $0 \leq \xi_i \leq 1$ and $\sum_{i=1}^{n} \xi = 1$.

Then ξ_i is the percentage of the population x which carries the allele a_i.

The class of all algebras with genetic realization is too large to say much about. However, since all gametic algebras (and their commutative duplicates) satisfy the definition, it provides a solid framework for what constitutes an algebra with genetic significance.

As we have seen, algebras with genetic realization are not necessarily associative algebras. They do belong to a rather special class of nonassociative algebras.

A general nonassociative algebra need not possess a matrix representation. Yet, algebras with genetic realization do. In fact, they possess the simplest possible matrix representation—a scalar representation.

Definition

An algebra A is called a **baric algebra** if it admits a non-trivial algebra homomorphism ω into the algebra consisting of the real numbers.

In other words, a baric algebra is an algebra with a one-dimensional representation. The homomorphism ω is called the weight function (or baric function).

Proposition An n-dimensional algebra with genetic realization is a baric algebra.

Weight functions are not unique in general: see Example 3.1 p.113

Even though the above example shows that not all baric algebras have a unique weight function, many of them do.

In order to exhibit at least a sufficient condition for a baric algebra to have a unique weight function, we must first discuss the issue of powers in a nonassociative algebra.

In a commutative, nonassociative algebra, there are several ways to define and interpret the powers of an element. There are two main types of powers which have genetic significance.

Let x be an element of a commutative nonassociative algebra A.The **principal powers** of x are defined to be $x, x^2, x^3, \ldots$, where $x^i = x^{i-1}x$.

If A is an algebra with genetic realization and an element P represents a population, then each element P^i of the sequence of principal powers represents a population which resulted from the previous population P^{i-1} mating back with the original population P.

On the other hand, the **plenary powers** $x, x^{[2]}, x^{[3]}, \ldots$ are defined as $x^{[i]} = x^{[i-1]}x^{[i-1]}$.

When P is an element representing a population, the sequence of plenary powers contains the successive generations resulting from random mating within the population, beginning with P.

$P^{[2]}$ is the result of the population P mating within itself, and $P^{[3]}$ is the result of the population $P^{[2]}$ mating within itself.

Both the principal and plenary powers are of biological as well as mathematical interest.

The following proposition provides a sufficient condition for a baric algebra to have a unique weight function.

Proposition Let A be a baric algebra with weight function ω. If $N = \ker \omega$ is nil (i.e., all elements of N are nilpotent), then ω is uniquely determined.

A non-zero element e in an algebra which satisfies the relationship $e = e^2$ is called an **idempotent**. In addition to their mathematical importance, idempotents also have genetic significance.

If a population P satisfies the equation $P^2 = P$, this means that genetic equilibrium has been achieved after one generation of random mating within the population P. I.e., the population P^2 has the same genetic pool as the initial population P.

Mathematically, the existence of an idempotent in an algebra provides a direct sum decomposition of the algebra. Hence, idempotents play a crucial role in describing the general structure of an algebra.

Let A be a baric algebra. The existence of an idempotent in A is not guaranteed. However, if A does contain an idempotent e, then $\omega(e) = \omega(e^2) = \omega(e)^2$ so either $\omega(e) = 0$ or $\omega(e) = 1$.

Proposition Let A be a baric algebra with weight function ω. Suppose A contains an idempotent e such that $\omega(e) = 1$.Then $A = Re \oplus \ker \omega$ (R=real numbers)

February 17—Bernstein Algebras

(Ref. REED, pp.118–120)

From a mathematical point of view, the principal powers are much more tractable than their cousins, the plenary powers.

However, from a geneticist's point of view, the sequence of plenary powers is of greater interest since the plenary powers more accurately model the way most populations reproduce.

In the early 1920's, S. Bernstein studied a quadratic evolutionary operator Ψ, which mapped the simplex Δ_n of genetic frequency distributions $(x_1, \ldots, x_n)$ with $x_i > 0$ and $\sum_{i=1}^n x_i = 1$ into itself and represented the passage of generations.

Bernstein set out to classify all such operators which satisfied the condition $\Psi^2 = \Psi$. This condition is known as Bernstein's stationarity principle.

The condition $\Psi^2 = \Psi$ indicates that the population is in equilibrium after one generation. Ψ is thus an idempotent evolutionary operator.

Let A be a finite dimensional, commutative baric algebra and let ω denote the weight function. Then A is called a **Bernstein algebra** if the plenary powers of any element x satisfy $x^{[3]} = \omega(x)^2 x^{[2]}$

Those elements x of a Bernstein algebra with baric weight 1 satisfy the equation $x^{[3]} = x^{[2]}$, which means that they reach genetic equilibrium after one generation of random mating within the population.

Therefore, complete knowledge of the structure of these abstract nonassociative algebras would provide a great deal of genetically significant information.

Mathematically, one of the most useful facts about the structure of a Bernstein algebra is the direct sum decomposition it possesses.

Proposition

Let A be a Bernstein algebra with weight function ω. .Let e denote an idempotent of A. Let $Z = \ker \omega$, let U be the set of elements x in Z with $ex = x/2$ and let V be the set of elements x in Z with $ex = 0$. Then $A = Re \oplus U \oplus V$.

Another remarkable (and extremely useful) fact about a Bernstein algebra is the classification of all of its idempotents.

Proposition

Let A be a Bernstein algebra and let e denote one idempotent in A. Then all idempotents of A have the form $e + u + u^2$, where u in Z satisfies $eu = u/2$.

A very natural generalization of a Bernstein algebra arises when we alter the condition on the plenary powers to reflect not equilibrium after one generation of mating, but instead after k generations of mating.

Definition

Let A be a finite dimensional commutative baric algebra, and let ω denote the weight function. Then A is called a k-th order Bernstein algebra if the plenary powers of any element x satisfy $x^{[k+2]} = \omega(x)^{2^k} x^{[k+1]}$

February 24—Derivations on Genetic Algebras

(Ref. for pp.78-87: Holgate, P. The Interpretation of Derivations in Genetic Algebras Linear Algebra and its Applications 85; 75-79 1987)

Abstract

Meanings are assigned to a linear transformation of an element in a genetic algebra representing a probability distribution over the possible genetic types, and to products of elements where only one of the factors is such a probability element. These lead on to a characterization of a derivation on a genetic algebra in terms of the equality of two genetically meaningful expressions.

Remark

In the literature, the term genetic algebra is often used of algebras that have certain mathematical properties, even in cases that do not represent a biological situation.

Interpretation 1—Preparation

Consider a population in which there are n genetic types $A_1, \ldots, A_n$. We will consider only *autosomal* characteristics, namely those that are determined by genes not linked to sex.

An individual X, for whom the probability that he or she is of genetic type A_i is x_i, will be represented by the vector $x = (x_1, \ldots, x_n)$.

Let V be the vector space generated by the x's, The elements of V for which $x_i > 0, \sum_i x_i = 1$ will be called *probability vectors*.

Let γ_{ijk} be the probability that the offspring of a mating between an A_i and an A_j will be an A_k. Thus $\sum_k \gamma_{ijk} = 1$ for every pair i, j.

We make V into a genetic algebra by defining the product of elements x, y by

$$\text{xy} = \text{z where } z_k = \sum_{i,j} x_i y_j \gamma_{ijk}.$$

The probability vectors form a multiplicatively closed set.

Interpretation 1

If x, y are probability vectors representing the distribution on the genetic types of individuals X, Y, then the product xy represents the probability distribution of the genetic type of the offspring of X and Y.

Let $R = (r_{ij})$ be a matrix whose (i, j)th element is the value of some genetically determined trait, also denoted by R, for a male of genetic type A_i when he is coupled with a female of type A_j, and vice versa.

Interpretation 2

If x is the probability vector of X, and R a matrix as above, the jth component of Rx gives the mean value of the trait for X when it is mated to an A_j.

These components will be called the **conditional means** of R for X.

Preparation for Interpretation 3

Now let let x be a probability vector representing a male, and q a vector whose jth component is the unconditional value of some trait Q for the genetic type A_j.

We look for an explanation of the algebra product xq (= t) in terms of genetic phenomena.

Its kth component is

$$t_k = \sum_i x_i \sum_j q_j \gamma_{ijk} = \sum_j q_j \sum_j x_i \gamma_{ijk} = \sum_j q_j \pi_{jk}$$

where $\pi_{jk} = \sum_i x_i \gamma_{ijk}$ is the probability that the offspring of a mating between X and an A_j, will be an A_k.

If an individual X is mated with a partner chosen by assigning equal probability to each of the genetic types, we say it has taken part in a *uniform breeding trial*.

We will need a simple case of the concept of *posterior probability*.

Let (p_{ij}) be the matrix of probabilities that the event j will occur given the condition i. Suppose that each i has equal prior probability.

Then, with a straightforward frequency interpretation, the probability that the condition was i when the event j has been observed is obtained by normalizing the entries in column j of the matrix to sum to 1.

Thus the posterior probability of i given the occurrence of j is $p_{ij}/\sum_s p_{sj}$. The ratio of posterior probability to prior probability is the likelihood of the condition on the evidence of the event.

We now define the quantities

$$u_k = n^{-1}\sum \pi_{jk}, \qquad \nu_{jk} = \frac{\pi_{jk}}{\sum_j \pi_{jk}} = \frac{\pi_{jk}}{nu_k}, \qquad m_k = \sum_j \nu_{jk} q_j.$$

Thus u_k is the probability that the offspring of X is an A_k in a uniform breeding trial, and ν_{jk} is the posterior probability that the mate of X was an A_j, given that the offspring of a uniform trial is observed to be an A_k.

Further, m_k is the corresponding posterior mean or *Bayes estimate* of the trait Q for X's mate, after the observation that the offspring of mating with X in such a trial is an A_k.

Then $t_k = nm_k u_k$.

Interpretation 3

The kth component of the vector arising as the product of an individual vector x and a general vector q is the product of (i) the ratio of the probability that X will produce an offspring A_k in a uniform breeding trial to the probability that his mate is an A_k, and (ii) the Bayes estimate of trait Q for his mate in such a trial, after the observation that the offspring is an A_k.

We call xq the vector of weighted Bayes estimates of Q.

Note that the average of the n terms obtained by multiplying (i) and (ii) is simply the mean of the trait for X's mate.

Preparation for Interpretation 4: D(xy)=xD(y)+yD(x)

Let X be male and Y female.

(i) D(xy) is the vector of conditional means of the trait D for their offspring given its mate's type.

(ii) Suppose that x represents the proportional distribution of genetic types in the male part of a population. Then Dx is the vector of conditional means for the male population, or what is equivalent, for an individual randomly selected from it with each individual having equal probability.

If x is not known, we cannot compute Dx, but Interpretations 2 and 3 together show that on the basis of a uniform breeding trial with a female Y whose probability vector y is known it can be estimated by the Bayes estimate yD(x).

If male and female are interchanged, we use xD(y). The sum of these estimates gives the right hand side of the equation.

Interpretation 4

Now (i) is a vector of means for an offspring, computed from the probability vectors of its parents, while (ii) is the sum of vectors of means for the parents, estimated from observations on the offspring.

For general traits characterized by a matrix R, the two sides will not be equal.

Equality for a particular trait with array of values D reflects, subject to the precise meanings of the quantities appearing on the two sides of the equation discussed above, a kind of symmetry in respect of the time direction.

Algebraically this corresponds to a set of relationships between the components of a matrix D and the constants γ_{ijk} that define the mechanism of heredity.

If two traits are such that the arrays of their values give rise to derivations, so do their weighted sum $\alpha D + \beta G$ and commutator product [D,G]=DG-GD.

Note that $\sum_i \sum_j n^{-1} x_i d_{ij} g_{jk}$ is the product moment of the D value of X and the G value of his mate, if she were to mate with an A_k in a uniform breeding trial.

Interpretation 5

The (i, k)th element of the commutator product [D,G] is n times the difference between the product moment of the D value of A_i with the conditional G value of his mate, and that of the G value of A_k with the conditional D value of his mate, in a uniform breeding trial.

Conclusion

The matrix of conditional trait values is a derivation in those cases where there is a certain relation between the outcome of a generation of random mating and the estimates obtained from a uniform trial. The closure of the derivation algebra with respect to commutator multiplication implies that it contains "higher order" traits which relate to the interaction, in a statistical sense, between the traits of an individual and those of his mate.

(Ref. for pp.88-90: Bingham, N. H. Obituary: Philip Holgate (1934-1993). Bull. London Math. Soc. 32 (2000), no. 4, 484-492)

Although Mendel's work on genetics dates back to 1865, it came to widespread attention only in 1900.

One of the earliest mathematical results on Mendelian genetics was the Hardy-Weinberg law of 1908, showing that under broad conditions, the stationary distribution of genotypes is achieved in one generation of random mating.

This work was taken further by the great Russian probabilist S. N. Bernstein (1880-1968) in 1922, leading to the Bernstein stationarity principle.

Our genetic inheritance comes to us symmetrically from our two parents (apart from sex-linked characteristics and aspects such as mitochondrial DNA), so the operation of multiplication in a genetic algebra is symmetric.

It is not, however, associative: complete information about who has mated with whom in a family tree is crucially relevant, and thus all brackets need to be retained.

Non-associative algebras in general lack enough structure to be mathematically tractable, and one needs additional structural properties to be able to make progress.

The two best-known types of non-associative algebras are Lie algebras (dating from work of Sophus Lie in the 1870s and Jordan algebras, which date from work of P. Jordan on quantum mechanics in 1933.

After Bernstein, and J. B. S. Haldane in 1930, the next serious attempt to study mathematical genetics via non-associative algebras was made by I. M. H. Etherington in 1939, and later work up to 1951. Subsequent workers included R. D. Shafer in 1949, and H. Gonshor in 1960.

There resulted a number of classes of non-associative algebras with genetic significance, including (in increasing order of generality) ‘special train algebras’ (Etherington), genetic algebras (Shafer, Gonshor), ‘train algebras’, baric algebras, and algebras with genetic realisation.

To one side of this chain of inclusions, and motivated by Bernstein's stationarity principle, are Bernstein algebras.

The study of these classes of algebras, and their genetic significance, has developed into an important area. One of P. Holgate's key contributions was to give a characterisation of Schafer's genetic algebras alternative to the original one.

Schafer's approach was algebraically motivated, and its genetic significance was not transparent; Holgate gave an approach which made the genetic meaning clear.

The operations of genetics, involving the mating of two parents, are naturally quadratic. However, Haldane showed in 1930 that the methods of linear algebra can sometimes still be used.

"After R. D. Shafer's seminal paper of 1949, all the best new ideas on the subject have been introduced either by Professor Holgate himself, or by his disciples."

(Ref. for pp. 91-96: Costa, R. On the derivations of gametic algebras for polyploidy with multiple alleles. Bol. Soc. Brasil. Mat. 13 (1982), no. 2, 69-81)

A genetic algebra is a commutative algebra for which there exists a basis $C_0, C_1, \ldots, C_n$, with a multiplication table satisfying the following conditions: If $C_i Cj = \sum_{k=0}^{n+1} \lambda_{ijk} C_k$ then:

- 1. $\lambda_{000} = 0$ 2. $\lambda_{0jk} = 0$ if $k < j$
- 3. For $i > 0, j > 0$, $\lambda_{ijk} = 0$ if $k \leq \max\{i, j\}$

Any basis of A, satisfying conditions 1,2,3 is called a canonical basis of A.

Let us indicate by G(n + 1, 2m) the gametic algebra of a 2m-ploid population with n + 1 alleles, which we shall denote here by $A_0, A_1, \ldots, A_n$.

This algebra has a natural basis consisting of all monomials of degree m in the "variables" $A_0, A_1, \ldots, A_n$. Each one of these monomials represents one of the gametic types of the population. The dimension of G(n+ 1,2m) is $(n + m)!/m!n!$.

The multiplication of two of these monomials is an algebraic representation of the distribution of gametic types obtained by the mating of the gametic types corresponding to the given monomials.

Of special importance in genetics are the algebras G(n + 1, 2) and G(2, 2m). They correspond to multiallelism and polyploidy respectively.

The algebras G(n + 1, 2) describe the gametic population of a diploid and multiallelic population. The natural basis of G(n + 1, 2) is the set of monomials of degree 1, namely $A_0, A_1, \ldots, A_n$.

The multiplication table is $A_iA_j = 1/2A_i + 1/2Aj$, which reads genetically as "the gametes produced by a zygote resulting from the mating of gametes A_i and A_j will be A_i and A_j with equal probability".

One canonical basis is defined by $C_0 = A_0, C_i = A_0 - A_i$ $(i > 1)$ and now the multiplication table is $C_0^2 = C_0, C_0C_i = 1/2\,C_i, C_iC_j = 0$ if $i \geq 1, j \geq 1$.

The weight function w of G(n + 1, 2) is given by $w(C_0) = 1, w(C_i) = 0(i > 1)$ or by $w(A_i) = 1(i = 0, 1, \ldots, n)$

It is well known that G(n + 1, 2) satisfies the polynomial equation $x^2 = w(x)x$ for every x in G(n + 1, 2). This identity may be linearized to give the two variables identity 2xy = w(x)y + w(y)x.

Theorem

Let A be a baric algebra having a unique weight function w. For every derivation d of A we have w(d(x))=0 for every element x of A.

Corollary

Let A be a baric algebra of dimension n + 1 with unique weight homomorphism. Then the dimension of the derivation algebra of A is not greater than n(n + 1). In particular for every genetic algebra of dimension n + 1, its derivation algebra has dimension not greater than n(n + 1).

Proposition

The derivations of G(n + 1,2) are exactly those linear mappings d such that w(d(x))=0 for all x.

Proposition

Let $A_0, A_1, \ldots, A_n$ be the natural basis of G(n + 1, 2) and let d_{ij} $(i \neq j)$ be defined by $d_{ij}(A_k) = A_i - A_j$ if $k = i$ and zero otherwise. Then the elements d_{ij} $(i \neq j)$ form a basis of the derivation algebra of G(n + 1, 2).

Theorem

Let A be a commutative baric algebra of dimension n + 1, with weight function w. Suppose that every linear mapping d on A such that w(d(x))=0 for every x, is a derivation of A. Then A is (isomorphic to) G(n + l, 2).

The algebras G(2,2m) describe the gametic population corresponding to a 2m-ploid and diallelic population.

The natural basis of G(2, 2m) is the set of monomials of degree m in the two variables A_0 and A_1. They are $A_0^m, A_0^{m-1}A_1, \ldots, A_0A_1^{m-1}, A_1^m$ so the dimension of G(2, 2m) is m + 1.

The product of two of such monomials is given by $(A_0^iA_1^{m-i})(A_0^jA_1^{m-j}) = \binom{2m}{m}^{-1}\sum_{k=0}^{m}\binom{i+j}{k}\binom{2m-i-j}{m-k}A_0^kA_1^{m-k}$ which is an algebraic way of expressing the distribution of probability for the gametes produced by the zygote obtained by mating the gametes $A_0^iA_1^{m-i}$ and $A_0^jA_1^{m-j}$.

A canonical basis for G(2, 2m) is the set of monomials $A_0^i(A_0 - A_1)^{m-i},\ (0 \leq i \leq m)$, with multiplication given by $[A_0^i(A_0 - A_1)^{m-i}][A_0^j(A_0 - A_1)^{m-j}] = \binom{2m}{m}^{-1}\binom{i+j}{m}A_0^{i+j-m}(A_0 - A_1)^{2m-i-j}$ if $i + j \geq m$ and zero otherwise. If we call $A_0^m = c_0, A_0^{m-1}(A_0 - A_1) = c_1, \ldots, (A_0 - A_1)^m = c_m$ then we have $c_ic_j = \binom{2m}{i+j}^{-1}\binom{m}{i+j}c_{i+j}$ if $i + j \leq m$ and zero otherwise.

We construct a derivation δ of G(2, 2m) by defining δ to be the linear mapping $\delta(c_i) = ic_i, (i = 0, 1, \ldots, m)$. It is easy to verify that δ is indeed a derivation.

A second derivation η is defined in the following way: $\eta(c_i) = \frac{t_{i+1}}{t_i - t_{i+1}} c_{i+1}$ for $0 \leq i \leq m-1$ and $\eta(c_m) = 0$, where $t_i = \binom{2m}{i}^{-1}\binom{m}{i}$ are the "train roots" of G(2,2m). We have $1 = t_0 > t_1 = 1/2 > t_2 > \cdots > t_m$. It is not so easy to see that η is a derivation.

Theorem

Every derivation d of G(2, 2m) is a linear combination of δ and η.

Corollary

For any $m > 1$, the derivation algebra of G(2, 2m) is isomorphic to the non abelian Lie algebra of dimension 2.

Theorem

The derivation algebra of G(n+1,2m) has dimension n(n+1).

(Ref. for pp.97-: Camacho, L. M.; Gomez, J. R.; Omirov, B. A.; Turdibaev, R. M. The derivations of some evolution algebras. Linear Multilinear Algebra 61 (2013), no. 3, 309-322)

The concept of evolution algebras lies between algebras and dynamical systems. Dynamically, they represent discrete dynamical systems.

In fact, evolution algebras are closely related with graph theory, group theory, stochastic processes, mathematical physics, genetics etc.

Let E be a finite dimensional evolution algebra with natural basis $e_1, \ldots, e_n$. Then by definition, $e_i \cdot e_i = \sum_{j=1}^{n} a_{ij} e_j$ with all other products zero. The matrix $A = (a_{ij})$ is called the matrix of the algebra E in the natural basis $e_1, \ldots, e_n$.

Let d be a derivation on the evolution algebra E with natural basis $e_1, \ldots, e_n$ and $d(e_i) = \sum_j d_{ij} e_j$ Then the space of derivations for the evolution algebra E is described as follows. It is the set of those linear operators d such that

$$a_{kj} d_{ij} + a_{ki} d_{ij} = 0, \text{ for } i \neq j; \text{ and } 2a_{ji} d_{ii} = \sum_k a_{ki} d_{jk}.$$

Theorem

Let d be a derivation on the evolution algebra E of dimension n with non-singular evolution matrix A, i.e. rank of A is n. Then this derivation d is zero.

Theorem

Let d be a derivation on the evolution algebra E of dimension n with evolution matrix A having rank $n-1$. Then this derivation d is zero or it has a matrix of the explicit form $D_1, \ldots, D_5$.

The matrices $D_1, \ldots, D_5$ are described explicitly in the paper.

(Ref. for pp. 97-: Costa, R. On the derivation algebra of zygotic algebras for polyploidy with multiple alleles. Bol. Soc. Brasil. Mat. 14 (1983), no. 1, 63-80)