EVOLUTION ALGEBRA

Freshman Seminar
University of California, Irvine

Bernard Russo

University of California, Irvine

Winter 2015

Linear Algebras

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:

- $xx' = x'x$, $(xx')x'' = x(x'x'')$ (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 4. $X \times Y$

	X×*Y*	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 5. $Y \times X$

	Y×*X*	Y: 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

$$x = x_1e_1 + x_2e_2 + \cdots + x_ne_n \text{ and } y = y_1e_1 + y_2e_2 + \cdots + y_ne_n$$

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

$$xy = x_1y_1e_1e_1 + x_1y_2e_1e_2 + \cdots + x_iy_je_ie_j + \cdots + x_ny_ne_ne_n$$

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_ie_j;

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

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,

$$xy = z_1 e_1 + z_2 e_2$$

Genetic motivation

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 1. 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 2.

Table 2. 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 2. 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 3. So we can define the three dimensional algebra over R with basis {AA, Aa, aa} and multiplication table as in Table 3. It is called the **zygotic algebra** for simple Mendelian inheritance with two alleles.

Table 3. 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.