数理解析研究所講究録
第 1551 巻 2007 年 117-122

Algebraic methods for genetics(I) (Some useful algebras in genetics)

A.Micali[1) and O. Suzuki[2)]

[1)] *Department of Mathematical Sciences, University of Montpellier II, Place Eugene Bataillon ,Montpellier, France*
[2)] *Department of Computer Sciences and System Analysis, Nihon University, Sakurajosui, Setagaya-ku, Tokyo, Japan.*
(A.Micali)micali@math.univ-montp2.fr (O.Suzuki) osuzuki@cssa.chs-nihon.ac.jp
Key words: population genetics, molecular genetics, gametic algebra.

Abstract

We give some algebras that are useful in genetics especially in pulation genetics and molecular genetics. At first we introduce a concept of evolution algebra which is connected to genetics intimately and give several examples. Next we proceed the normal form of weighted algebra and to the construction of Mendelian gametic algebra and non Mendelian gametic algebras for a finite or infinite number of alleles.

1. Evolution algebras

We take a **R**-vector space A with a R-linear mapping $\omega : A \to R(\omega \neq 0)$. We define the **simplex** associated to (A_V, ω) by $\Delta = \{x \mid x \in A \mid \omega(x) = 1\}$. We suppose that a **R**-quadratic mapping $V : A \to A$ is given. It satisfiesthe conditions(1) $V(\lambda x) = \lambda^2 V(x)$ for $\forall \lambda \in R, \forall x \in V$ and (2)putting $b_V(x,y) = V(x+y) - V(x) - V(y)$, we have a symmetric **R**-bilinear mapping $b_V : A \times A \to A$. Introducing the product on A by

$$xy = \frac{1}{2} b_V(x,y) \quad \text{for} \quad \forall x, \forall y \in A ,$$

we have an algebra A_V which we call **evolution algebra** associated to the operator V. We see that $A_V = A$ as a **R**-vector space and that the algebra A_V is commutative but not necessarily associative. We notice that $x^2 = V(x)\ (\forall x \in A_V)$. This shows that $x \in A_V$ is an idempotent if and only if $V(x) = x$: in other words, the restriction of V to the set $I_p(A_V)\ (= \{x \in A_V \mid x^2 = x\})$ of idempotents of of A_V is the identical mapping. We can prove the following lemma:

Lemma 1.1. If $V(\Delta) \subset \Delta$,then $\omega(V(x)) = \omega(x)^2$ for $\forall x \in A_V$ and (A_V, ω) is a **weighted algebra**, that is, $\omega(x,y) = \omega(x)\omega(y)$ for $\forall x, \forall y \in A_V$.

Example 1.2. For a real number δ $(\delta \neq 1/2)$, a weighted algebra (A_V, ω) is called δ**-Bernstein algebra** if it satisfies the equation:

$$(x^2)^2 - 4\delta\omega(x)x^3 + (4\delta^2 + 4\delta - 1)\omega(x)^2 x^2 + 2\delta(1-2\delta)\omega(x)^3 x = 0 (\forall x \in A).$$

A 0-Bernstein algebra ($\delta = 0$) is simply called **Bernstein algebra** and it is defined by the equation $(x^2)^2 = \omega(x)^2 x^2 (\forall x \in A)$. The quadratic operator $V:\ A \to A$ is given here by $x \mapsto \dfrac{1}{1-2\delta}(x^2 - 2\delta\omega(x)x)$. The quadratic operator verifies $V^2 = V$ on the simplex Δ of (A, ω) (see [1], proposition 4.2.6 or [2], proposition 4.4.1). The **R**-bilinear mapping b_V and product structure are given as follows:

$$b_V(x,y) = \frac{2}{1-2\delta}(xy \ - 2\delta\omega(x)y - \delta\omega(y)x)$$

$$x * y = \frac{1}{2}b_V(x,y) = \frac{1}{1-2\delta}(xy \ - \delta\omega(x)y - \delta\omega(y)x)$$

for $\forall x, \forall y \in A_V$. Moreover, $\omega(x * y) = \omega(x)\omega(y)$ for $\forall x, \forall y \in A_V$. Hence we see that (A_V, ω) is a weighted algebra. Putting $x \circ y = \frac{1}{2}(\omega(x)y + \omega(y)x)$ we have

$$xy = 2\delta\, x \circ y + (1-2\delta)x * y.$$

Hence we see that the algebra (A, ω) is a mixture (see [9] for the definition of mixture of algebras) of the algebras $(G(n+1,2),\omega)$ and (A, ω) with parameter 2δ.

Example 1.3A weighted algebra (A_V, ω) is called δ**-Etherington algebra** $(\delta \neq 1/2)$, if

$$(x^2)^2 - (1+2\delta)\omega(x)x^3 + \delta(1+2\delta)\omega(x)^2 x^2 + \delta(1-2\delta)\omega(x)^3 x = 0\ (\forall x, \in A_V).$$

In the case of $\delta = 0$, the algebra is simply called **Etherington algebra.** It is defined by the equation $(x^2)^2 = \omega(x)x^3 (\forall x \in A)$. The quadratic operator $V: A \to A$ is given by $x \mapsto \dfrac{1}{1-2\delta}(2x^2 - (1+2\delta)\omega(x)x)$ (see [1], proposition 4.2.7 or [2], proposition 4.4.2.). The algebraic structure of the evolution algebra (A, ω) is given by

$$x * y = \frac{1}{2}b_V(x,y) = \frac{1}{1-2\delta}(2xy \ - (1+2\delta)(x \circ y)\ \forall x, \forall y \in A_V.$$

Moreover, $\omega(x * y) = \omega(x)\omega(y)$ for $\forall x, \forall y \in A_V$ and then (A_V, ω) is a weighted algebra. We have a mixture structure algebras (G (n+1, 2), ω) and (A_V, ω) with a parameter $\frac{1}{2}(1+\delta)$:

$$xy = \frac{1}{2}(1+2\delta)x \circ y + \frac{1}{2}(1-2\delta)x * y.$$

Note1.4. **A** weighted algebra (A_V, ω) satisfying $(x^2 - \omega(x)x)^2 = 0$ $(\forall x \in A_V)$ is called **Bernstein-Etherington algebra**. (cf. [2], lemma 7.9). As for a complete study of Bernstein-Etherington algebra see [2], paragraph 7.

2.Normal form of commutation table of weighted algebra

We take a weighted algera (A_V, ω) which is generated by $(e_0, e_1, .., e_n)$. Then we have

$$e_i^2 = \sum_{j=0}^{n} \alpha_{ij} e_j,$$

where $\alpha_{ij} \geq 0$ for all i, j. We see that α_{ij} describe the rate of genotype e_j produced by genotype e_i. Next we find the normal form of α_{ij}. We may assume that the mapping ω:A→**R** satisfies $\omega(e_0) = 1$ and $\omega(e_j) = 0 (j = 1,2,..,n)$. Then we have

$$e_0^2 = e_o + \sum_{j=1}^{n} \alpha_{oj} e_j \text{ and } e_i^2 = \sum_{j=1}^{n} \alpha_{ij} e_j (i = 1,2,..,n)$$

We have $A = \mathrm{Re}_0 \oplus N$, where N =Ker (ω). It is an ideal of the algebra A. This supplies a necessary and sufficient condition to have a structure of weighted algebra on A. Next we consider the linear mapping $d_V : A \to Hom_R(A, A)(= End_R(A))$ defined by $x \mapsto (y \mapsto b_V(x, y))$. Then there exist basis $(f_{ij})_{1 \leq i,j \leq n}$ of $End_R(A)$ such that $f_{ij}(e_k) = \delta_{ik} e_j$ A. From $b_V(e_i, e_k) = d_V(e_i)(e_k)$ and $d_V(e_i) = \sum_{j=1}^{n} \lambda_{ij} f_{ij}$, we have

$$e_i^2 = \frac{1}{2} b_V(e_i, e_i) = V(e_i) = \sum_{j=1}^{n} \lambda_{ij} e_j (i = 1,..,n),$$

$$e_i e_k = \frac{1}{2} b_V(e_i, e_k)(i, k = 1,..,n; i \neq k),$$

where λ_{ij} are real values for all i, j. As for diweighted algebras we refer to [7] or [16].

3.Mendelian gametic algebra

We give a general construction of Mendelian gametic algebras on the space of homogenous polynomials of degree m (see[10]). We use the normal form without mentioning it. We consider the algebra of polynomials of $n+1$ variables $R[X_0, X_1, ..., X_n]$. For an integer m $(m \geq 1)$, we put $S^m(R^{n+1})$ the R-vector subspace of homogeneous monomials degree . Putting the product structure on $S^m(R^{n+1})$ by

$$(X_0^{i_0} X_1^{i_1}, .. X_n^{i_n})(X_0^{j_0} X_1^{j_1}, .. X_n^{j_n}) = \binom{2m}{m}^{-1} \binom{i_0 + j_0}{m} X_0^{i_0+j_0-m} X_1^{i_1+j_1} .. X_n^{i_n+j_n}.$$

We can introduce **Mendelian gametic algebra** which is commutative but non associative and without the unit element. This product structure can be realized by use of derivations:

$$fg = \frac{m!}{(2m)!}\frac{\partial^m}{\partial X_0^m}(f \bullet g)$$

where $f \bullet g$ denotes the ordinary multiplication of polynomials. We can obtain the same formula in the case of K vector space, where K is a field of characteristic zero. For a derivation $d(S^p(M)) \subset S^{p-1}(M)$ which derives from a K-linear mapping $d : M \to K$, $d \neq 0$:

$$d(x_1,\ldots,x_p) = \sum_{i=1}^{p} d(x_i)x_1..\hat{x}_i..x_p,$$

where $\hat{x}_i$ implies removed, we see that

$$fg = \binom{2m}{m}^{-1} \sum_{r+s=m} \frac{1}{r!s!} d^r(f)\cdot d^s(g)$$

for all polynomials f, g in $S^m(M)$. We denote the algebra by $S^m(M, \partial/\partial X_0)$ or $S^m(M, d)$ for a general derivation d.

Example 3.1 In the case of $m=1$ in the mendelian gametic algebra, we have $e_j e_i = e_i e_j = 1/2(e_i + e_j)(i, j = 0,1,\ldots,n)$. This might be regarded as an algebraic description of the Mendelian separation law. We can give an algebraic description of Hardy-Weinberg's law in terms of idempotents:

$$(\sum_{j=1}^{n} p_j e_j)^2 = \sum_{j=0}^{n} p_j e_j \quad \text{where} \quad \sum_{j=0}^{n} p_j = 1 (p_j > 0)$$

Example 3.2. We consider the gametic algebra $S^m(M, \partial/\partial X_0)$ where M is a K-vector space of dimension $n+1$. Then we see that the variables $X_0, X_1,\ldots,X_n$ represent the different alleles of our population and m is the degree of the polyploidy. Monomials $X_0^{i_0}X_1^{i_1}\ldots X_n^{i_n}$ represent different **genotypes** of the mentioned population. In the case of two alleles X_0, X_1, the genotypes can be written $e_k = X_0^{m-k}X_1^k \,(k = 0,1,..,m)$ and the multiplication table in this case is as follows:

$$e_i e_j = \binom{2m}{m}^{-1}\binom{2m-i-j}{m} e_{i+j}(i, j = 0,1,..,m) \quad \text{and} \quad e_i e_j = 0, \ \text{if} \ i+j > m.$$

We call $S^m(M, \partial/\partial X_0)$ the **multiplicated algebra with two alleles.**

Example 3.3. In the case of $S^1(M, \partial/\partial X_0))$ and $M = K[e_0, e_1,..,e_n]$ the multiplication table becomes $e_0^2 = e_0, e_0 e_i = e_i e_0 = 1/2 e_i (i = 0,1,\ldots,n)$. We call $S^1(M, \partial/\partial X_0)$ the diploid algebra with $n+1$ alleles. The genotypes are identified with the alleles. We want to mension that in any case the polyploidy is given by 2m and in our case, $m=1$

that give the 2-ploid or diploid. Elsewhere, the algebra $S^1(M,\partial/\partial X_0)$ is denoted by $i \neq k \quad G(n+1,2)$.

4. Non Mendelian genetic algebras

Mendelian genetics is essentially biparental. Non Mendelian inheritance says that the crossing of different genotype is impossible since the inheritance is uniparental. An example of non Mendelian inheritance is given by **phytophthora infectans** which cause the late blight of potatoes (brunissure) and tomatoes. In view to give a mathematical formulation of genetics, we follows Birky's paper (see [6]; see also [17]) .We call a weighted algebra (A,ω) a **non Mendelian weighted algebra,** if the different genotypes $e_0, e_1,..,e_n$ satisfy

$$e_i e_j = 0 \quad \text{for } i,j = 1,2,..,n \text{ and } i \neq j,$$

$$e_i^2 = \sum_{j=1}^n \alpha_{ij} e_j \text{ and } e_0^2 = e_0 + \sum_{j=1}^n \alpha_{0j} e_j \text{ where } \alpha_{ij} \text{ are positive real numbers.}$$

It is evident that this hypothesis does not cover the totality of non Mendelian genetic algebras.

Example 4.1. Putting $(X_0^{i_0} X_1^{i_1},..X_n^{i_n})\,(X_0^{j_0} X_1^{ij},..X_n^{j_n})$=0 for $(i_0,i_1,...,i_n) \neq (j_0,j_1,...,j_n)$,

$$(X_0^{i_0} X_1^{i_1},..X_n^{i_n})^2 = \frac{m!}{(2m)!}\frac{\partial^m}{\partial X_0^m}(X_0^{2i_0} X_1^{2i_1},..X_n^{2i_n}) = \binom{2m}{m}^{-1}\binom{2i_0}{m} X_0^{2i_0-m} X_1^{2i_1},..X_n^{2i_n}$$

we have non Mendelian genetics on $R[X_0,X_1,...,X_n]$. It's evident that $(X_0^{i_0} X_1^{i_1},..X_n^{i_n})^2$ $=0$ for $2i_0 < m$ and that $(X_0^m)^2 = X_0^m$ which says that X_0^m is an idempotent of the algebra $S^m(R^{n+1})$. The **R**-linear $\omega : S^m(R^{n+1}) \to R$ defined by $\omega(x) = (1/m!)$ $\partial^m/\partial X_0^m(x)$ verifies $\omega(xy) = \omega(x)\omega(y)$. Hence $(S^m(R^{n+1}),\omega)$ is a weighted algebra. We give examples:

(1)For $S^1(R^2)$ a basis is $\{X_0, X_1\}$ with multiplication table $X_0^2 = X_0$, $X_0 X_1 =$ $X_1 X_0 = 0$, $X_1^2 = X_1$.

(2)For $S^2(R^2)$ a basis is $\{X_0^2, X_0X_1, X_0X_2, X_1^2, X_1\, X_2, X_2^2\}$ with multiplication table
$X_0^2 = X_0$, $(X_0X_1)^2 = (1/6)X_1^2$, $(X_1X_0)^2 = (1/6)X_2^2$ and the other product are zero.

REFERENCE

[1] J.Bayara, Sur les algèbras d'évolution, Thèse de Doctorat, Université de Ouagadougou, 1999.

[2] J. Bayara, M. Ouattara, A. Micali et J.C. da Motta Ferreira, Sur une classe d'algèbres d'évolution, Algebras, Groups and Geometries 19 (2002), 315-346.

[3] N. Bourbaki, Algèbre I, Chapiter 1 a 3, Hermann, Paris 1970.

[4] N. Bourbaki, Appendice I of chapter 7, in Groups et Algèbres de Lie, Chapitres 7 et 8, Hermannm Paris 1975.

[5] N. Bourbaki, Algebre Commutative, Chapitres 1ét 2, Hermann, Paris 1961.1

[6] C.W. Birky, Jr., The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms and models, Annu. Rev. Gent. 35 (2001), 125-148.

[7] M.A. Couto and J.C. Gutiérrez Fernandez, Dibaric algebras, Projectiones Journal of Mathematics 19(3) (2000), 249-269.

[8] I.M.H. Etherington, Duplication of linear algebras, Proc. Edinburgh Math. Soc. 6(2) (1941),222-230.

[9] P. Holgate, Selfing in genetic algebras, j. Math. Biol. 6 (1978), 197-206.

[10] A. Micali et Ph. Revoy, Sur les algèbres gamétiques, Proc. Edinburgh Math. Soc. 29 (1986), 187-197.

[11] A. Micali et M. Ouattara, Dupliquée d'une algebre et le théorème d'Etherington, Linear Alg. Appl. 153 (1991), 193-207.

[12] A. Micali et M. Ouattara, Sur la dupliquée d'une algèbre, Bull. Soc. Math. Belgique 45(1) Série B (1993), 5-24.

[13] A. Micali et M. Ouattara, Sur la dupliquée d'une algebre II, Bull. Soc. Math. Belgique 43(1,2) Série A (1991), 113-125.

[14] A. Micali, A. Koulibaly et M. Ouattara, Dupliquée d'une algèbre et le théorème de Whitehead, Indagationes Mathematicae N. S. 5(3) (1994), 341-351.

[15] M. Ouattara, Autour de la duploquée, Indagationes Mathematicae N. S. 2(1) (1991), 99-104.

[16] M. Ouattara, A. Micali and I. Katambé, Diweighted algebras and two notes, JP Journal of Algebra,Number Theory and Applications 6(2)(2006),265-277.

[17] J.P. Tian, Algebraic structure of non-Mendelian inheritance, preprint Ohio State University, Mathematical Bioscience Institute.