The Interpretation of Derivations in Genetic Algebras

P. Holgate
Statistics Department
Birkbeck College
Malet Street
London WC1E 7HX, U.K.

Submitted by Richard A. Brualdi

ABSTRACT

Meanings are assigned to a linear transform 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.

Consider a population in which there are n genetic types $A_1, \dots, A_n$. 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 $x = (x_i)$. Let V be the vector space generated by the x's. The elements of V for which $x_i \geqslant 0$, $\Sigma x_i = 1$ will be called *probability vectors*. We will consider only *autosomal* characteristics, namely those that are determined by genes not linked to sex. 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 $\Sigma_k \gamma_{ijk} = 1$ for every pair i, j. We make V into a *genetic algebra* by defining the product of elements x, y *by* $xy = z$ where $z_k = \Sigma\Sigma x_i y_j \gamma_{ijk}$. The probability vectors form a multiplicatively closed set. 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. In particular the property $\gamma_{ijk} \geqslant 0$, essential for the genetic interpretation, is algebraically unimportant. Let us record

INTERPRETATION 1. If x, y are probability vectors representing the distributions 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.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/87/$3.50

The best introduction to genetic algebra is [4]. A detailed survey up to 1980 is contained in [10].

A *derivation* D of an algebra is a linear mapping, $x \to xD$, such that

$$(xy)D = (xD)y + x(yD). \tag{1}$$

There has recently been much work on the subject of derivations of genetic algebras (see [2, 3, 5, 6, 7, 8]). The Lie algebra of derivations of a given algebra is an important tool for studying its structure, particularly in the nonassociative case, so this is a natural development. In [9] multiplication is defined in terms of derivations, while in the study [1] of Bernstein algebras, Sections 3, 4, and 7 deal with their derivations.

So far, however, no explanation has been given of the *genetic* meaning of a derivation of a genetic algebra. This note is an attempt to fill the gap.

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. To fix the ideas we may think of the *fecundity* of the individual, which may reasonably depend on the type of the mate. We have

INTERPRETATION 2. If x is the probability vector of X, and R a matrix as above, the jth component of xR 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.

Now 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_k. We look for an explanation of the algebra product xq ($= t$) in terms of genetic phenomena. Its kth component is

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

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 $p_{ij}/\Sigma_s p_{sj}$. 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

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

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

$$t_k = n m_k u_k.$$

Hence we have

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.

Now return to Equation (1). Let X be male and Y female.

(i) The left hand side of (1) 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 xD 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 xD, 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

$(xD)y$. If male and female are interchanged, we use $x(yD)$. The sum of these estimates gives the right hand side of (1).

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 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 (1) which are 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, as studied in [6, 7, 8].

The above constitutes *Interpretation 4.*

If D, G are derivations, then it follows from (1) that so are $\alpha D + \beta G$, and $[D, G] = DG - GD$. If two traits are such that the arrays of their values give rise to derivations, their weighted sum will have the same property, since (1) is linear in D. Note that $\Sigma_i \Sigma_j n^{-1} x_i d_{ij} g_{jk}$ 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_i with the conditional D value of his mate, in a uniform breeding trial.

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.

REFERENCES

1 M. T. Alcalde, C. Burgueño, A. Labra, and A. Micali, Sur les algèbres de Bernstein, submitted for publication.
2 R. Costa, On the derivation algebra of gametic algebras for polyploidy with multiple alleles, *Bol. Soc. Brasil. Mat.* 13:69–81 (1982).
3 R. Costa, On the derivation algebra of zygotic algebras for polyploidy with multiple alleles, *Bol. Soc. Brasil. Mat.* 14:63–80 (1983).
4 I. M. H. Etherington, Genetic algebras, *Proc. Roy. Soc. Edinburgh* 59:242–258 (1939).

5 H. Gonshor, Derivations in genetic algebras, *Comm. Algebra*, to appear.
6 A. Micali, T. M. M. Campos, M. C. Costa e Silva, S. M. M. Ferreira, and R. C. F. Costa, Dérivations dans les algèbres gamétiques, *Comm. Algebra* 12:239–243 (1984).
7 A. Micali, T. M. M. Campos, M. C. C. Costa e Silva, and S. M. M. Ferreira, Dérivations dans les algèbres gamétiques. II, *Linear Algebra Appl.* 64:175–184 (1985).
8 A. Micali, T. M. M. Campos, M. C. C. Costa e Silva, and S. M. M. Ferreira, Dérivations dans les algèbres gamétiques. III, *Linear Algebra Appl.*, to appear.
9 A. Micali and P. Revoy, Sur les algèbres gamétiques, *Proc. Edinburgh Math. Soc.*, to appear.
10 A. Wörz-Busekros, *Algebras in Genetics*, Lecture Notes in Biomathematics 36, Springer, 1980.

Received 22 October 1985; revised 2 December 1985