Theor Appl Genet (1988) 75:701–711

Properties of components of covariance of inbred relatives and their estimates in a maize population *

P. L. Cornelius

Departments of Agronomy and Statistics, University of Kentucky, Lexington, KY 40546-0091, USA

Received February 30, 1987; Accepted October 23, 1987
Communicated by A. R. Hallauer

Summary. Properties of three parameterizations, denoted as the C-model, D-model and Q-model, for covariances of inbred relatives under assumptions of no linkage or epistasis are explored and compared. Additive variance in an inbred population with inbreeding coefficient F, $\sigma^2_{AF} = (1 + F)\sigma^2_A$, where σ^2_A is additive variance in a panmictic population, if Q-model parameters Q_{xx} and Q_{xy} are both zero. Conditions sufficient for this to hold are presented in terms of gene frequencies and dominance contrasts (homozygotes vs. heterozygotes). Some other properties and potential uses of estimates of components in the models are also discussed. Estimates of components in the D-model and Q-model were calculated from a maize (*Zea mays* L.) study from which estimates of components in the C-model were previously published. Of particular interest were the covariance (Q_{xy}) of effects of alleles at complete homozygosity with "inbreeding depression effects", the covariance (D_1) of additive effects at panmixia with inbreeding depression effects and the within-locus variance (D_2, alias Q_{xx}) of inbreeding depression effects. Estimates of Q_{xy}, D_1, and D_2 were small and nonsignificant in most cases. For ear height in the second year of the study, D_2 appeared to be a major component. In some cases, results were obtained which had contradictory implications (negative D_2 coupled with positive Q_{xy} or D_1, and positive D_2 coupled with negative σ^2_D). A negative estimate of one or the other of σ^2_D or σ^2_A was obtained in one of the two within-year analyses for every character. Problems in getting realistic results were thought to be owing to excessive multicollinearity among the coefficients of the components in the expectations of the covariances of the kinds of relatives included in the study. Implications for future studies of this kind are discussed.

Key words: Genetic covariance – Quadratic components – Inbred relatives – Quantitative genetics – Maize

Introduction

A general model for genetic variances and covariances of inbred relatives with dominance and two-locus epistasis contains an inordinately large number of terms and, with linked loci, depends on usually unknown recombination frequencies (Weir and Cockerham 1977). Consequently, practical application must employ models derived under simpler assumptions. Parameterizations for the covariances of inbred relatives under assumptions of no linkage, no epistasis and random mating equilibrium in an original noninbred population from which the inbred relatives were derived have been given by Cornelius and Dudley (1975, 1976), Cockerham (1983), Wright and Cockerham (1986), and Cornelius and VanSanford (1988), the latter two papers developing the same model, but with differences in notation and method of derivation. Any of these models may be readily extended to include terms owing to joint dominance effects of loci and additive × additive epistasis as described by Cockerham and Matzinger (1985) in a theoretical study of response to selection under selfing. Cockerham and Matzinger also briefly discuss some issues concerning estimation of covariances for prediction purposes and put out an appeal for "more experimental information on these issues". Cornelius and VanSanford (1988) showed that the covariance components are not all estimable if the experimental study involves only progenies derived by self-fertilization.

* Journal Article No. 87-3-14 of the Kentucky Agricultural Experiment Station published with the approval of the Director

Cornelius and Dudley (1976) gave estimates of parameters in their model applied to data from a broadbase maize (*Zea mays* L.) synthetic population inbred for several generations by selfing and full-sib mating. Objectives of this paper are to present estimates of parameters in the models of Cockerham (1983) and Cornelius and VanSanford (1988) (or equivalently, of Wright and Cockerham 1986) obtained from the data of Cornelius and Dudley (1976) and to present some theoretical results concerning properties of parameters in these models. The models used by Cornelius and Dudley (1975, 1976), Cockerham (1983), and the model developed by Cornelius and VanSanford (1988) will be referred to as the *C*-model, *D*-model and *Q*-model, respectively, the names chosen because the *C*-model employs a parameter denoted as C, the *D*-model contains parameters D_1 and D_2 and the *Q*-model contains parameters denoted as Q_{yy}, Q_{xy} and Q_{xx}.

Relationships of parameters in the models

Definitions and relationships of parameters in the covariance component models are shown in Table 1. All of the quadratic components of covariance are first defined for individual loci and then summed over loci.

Q-model parameters Q_{yy} and Q_{xx} were developed by Cornelius and VanSanford (1988) as nonnegative quadratic forms in vectors of contrasts $\boldsymbol{y}$ and $\boldsymbol{x}$, respectively, and Q_{xy} as a bilinear form. Let Q_{yyi}, Q_{xyi} and Q_{xxi} denote the contributions of the i^{th} locus to Q_{yy}, Q_{xy}, and Q_{xx}, respectively. If there are n alleles, I_1, $I_2, \ldots, I_n$ at the i^{th} locus, and g_{jk} denotes the genotypic value of genotype I_jI_k, then vector $\boldsymbol{y}$ for the i^{th} locus consists of the $n(n-1)/2$ contrasts among homozygotes, $y_{jk} = g_{kk} - g_{jj}$, $(j < k)$. Similarly, $\boldsymbol{x}$ consists of the dominance contrasts (homozygotes vs. heterozygote), $x_{jk} = g_{jj} + g_{kk} - 2g_{jk}$, $(j < k)$. Expressions for Q_{yyi}, Q_{xyi} and Q_{xxi} in matrix notation are given in the Appendix. When $j > k$, for convenience we will define $y_{jk} = -y_{kj}$ and $x_{jk} = x_{kj}$, but these quantities contribute only once, not twice, to vectors $\boldsymbol{y}$ and $\boldsymbol{x}$, respectively. The expressions for Q_{yy}, Q_{xy} and Q_{xx} in Table 1 follow from the structure of the matrices involved. Q_{xxi} is the variance of $\tilde{x}_{\cdot k} = \sum_{j \neq k} p_j x_{jk}$, Q_{yyi} is the variance of $\tilde{y}_{\cdot k} = \sum_{j \neq k} p_j y_{jk} = g_{kk} - \sum_j p_j g_{jj}$, and Q_{xyi} is the covariance of $\tilde{x}_{\cdot k}$ and $\tilde{y}_{\cdot k}$. In addition to Q_{xx}, σ_D^2 and μ_∞^2 are also owing to dominance and, for completeness, expressions for these parameters in term of the x's is also given (Table 1). The expression for σ_D^2 follows from expressions for the dominance deviations (d_{jk}) given by Kempthorne (1969).

For later use, let us define the degree of dominance (Comstock and Robinson 1948) of I_j over I_k as $a_{jk} = x_{jk}/y_{jk}$, which is positive (negative) if I_j is dominant over (recessive to) I_k. By this definition, $a_{kj} = x_{kj}/y_{kj} = -a_{jk}$.

Properties of Q_{xx} (alias D_2)

Cornelius and VanSanford (1988) showed that the additive variance in an inbred population obtained by a regular system of inbreeding beginning with an initial panmictic population is given by $\sigma_{AF}^2 = [(1+F)/2]\,Q_{yy} - (1-F)\,Q_{xy} + [(1-F)^2/2(1+F)]\,Q_{xx}$. If $Q_{xx} = Q_{xy} = 0$ then $\sigma_A^2 = Q_{yy}/2$ and $\sigma_{AF}^2 =$

Table 1. Definitions and relationships of parameters in models for covariances of inbred relatives

Q-model (Cornelius and VanSanford 1988)

Genotypic contrasts:
$y_{jk} = g_{kk} - g_{jj}$, $x_{jk} = g_{jj} + g_{kk} - 2g_{jk}$,
g_{jk} = genotypic value of i^{th} locus genotype I_jI_k.

Quadratic functions:
$Q_{xx} = \sum_i [\sum_k p_k \tilde{x}_{\cdot k}^2 - \tilde{x}_{\cdot\cdot}^2]$, $Q_{xy} = \sum_i \sum_k p_k \tilde{x}_{\cdot k} \tilde{y}_{\cdot k}$
$Q_{yy} = \sum_i \sum_k p_k \tilde{y}_{\cdot k}^2$, $\mu_\infty^2 = \sum_i \tilde{x}_{\cdot\cdot}^2/4$
$\sigma_D^2 = \sum_i [\sum_k p_k^2 (\tilde{x}_{\cdot k} - \tilde{x}_{\cdot\cdot}/2)^2 + \sum_j \sum_{k>j} p_j p_k (x_{jk} - \tilde{x}_{\cdot j} - \tilde{x}_{\cdot k} + \tilde{x}_{\cdot\cdot})^2/2]$
where $\tilde{x}_{\cdot k} = \sum_{j \neq k} p_j x_{jk}$, $\tilde{x}_{\cdot\cdot} = \sum_k p_k \tilde{x}_{\cdot k}$,
$\tilde{y}_{\cdot k} = \sum_{j \neq k} p_j y_{jk}$

C-model (Cornelius and Dudley 1975)

Genotypic value:
$g_{jk} = \mu_R + (1 - \delta_{jk})(\alpha_{Rj} + \alpha_{Rk} + d_{jk}) + \delta_{jk}(\mu_{\infty(i)} + \alpha_{\infty k})$
where $\delta_{jk} = 0$ if $j \neq k$,
$\delta_{kk} = 1$ if homozygosity is a result of identity by descent,
$\delta_{kk} = 0$ otherwise.

Quadratic functions:
$\sigma_\infty^2 = \sum_i \sum_k p_k \alpha_{\infty k}^2$, $\mu_\infty^2 = \sum_i \mu_{\infty(i)}^2$,
$C = 2\sum_i \sum_k p_k \alpha_{Rk} \alpha_{\infty k}$,
$\sigma_A^2 = 2\sum_i \sum_k p_k \alpha_{Rk}^2$, $\sigma_D^2 = \sum_i \sum_j \sum_k p_j p_k d_{jk}^2$.

Relationship to *Q*-model:
$\sigma_\infty^2 = Q_{yy}$, $\mu_\infty^2 = \mu_\infty^2$, $C = Q_{yy} - Q_{xy}$,
$\sigma_A^2 = (Q_{yy} - 2Q_{xy} + Q_{xx})/2$, $\sigma_D^2 = \sigma_D^2$.

Relationship to *D*-model:
$\sigma_\infty^2 = 2\sigma_A^2 + 4D_1 + D_2$, $\mu_\infty^2 = \ddot{\mathrm{H}}$,
$C = 2(\sigma_A^2 + \mathrm{D}_1)$, $\sigma_A^2 = \sigma_A^2$, $\sigma_D^2 = \sigma_D^2$.

D-model (Cockerham 1983)

Genotypic value:
$g_{jk} = \mu + \alpha_j + \alpha_k + d_{jk}$.

Quadratic functions:
$\sigma_A^2 = \sum_i \sum_k p_k \alpha_k^2$, $\sigma_D^2 = \sum_i \sum_j \sum_k p_j p_k d_{jk}^2$,
$D_1 = \sum_i \sum_k p_k \alpha_k d_{kk}$,
$D_2 = \sum_i [\sum_k p_k d_{kk}^2 - (\sum_k p_k d_{kk})^2]$,
$\ddot{\mathrm{H}} = \sum_i (\sum_k p_k d_{kk})^2$.

Relationship to *Q*-model
$\sigma_A^2 = (Q_{yy} - 2Q_{xy} + Q_{xx})/2$, $\sigma_D^2 = \sigma_D^2$,
$D_1 = (Q_{xy} - Q_{xx})/2$, $D_2 = Q_{xx}$, $\ddot{\mathrm{H}} = \mu_\infty^2$.

Relationship to *C*-model:
$\sigma_A^2 = \sigma_A^2$, $\sigma_D^2 = \sigma_D^2$, $D_1 = 0.5\,C - \sigma_A^2$,
$D_2 = \sigma_\infty^2 - 2C + 2\sigma_A^2$, $\ddot{\mathrm{H}} = \mu_\infty^2$.

Wright and Cockerham (1986) Notation for *Q*-model

Genotypic value:
Same as *D*-model.

Quadratic functions:
$\sigma_{\mathrm{M}}^2 = \sum_i [\sum_k p_k (2\alpha_k + d_{kk})^2 - (\sum_k p_k d_{kk})^2]$,
$\sigma_{\mathrm{MH}} = \sum_i [\sum_k p_k (2\alpha_k + d_{kk}) d_{kk} - (\sum_k p_k d_{kk})^2]$,
σ_D^2, D_2^* and H^* same as D_1, D_2 and $\ddot{\mathrm{H}}$ in *D*-model

Relationship to *Q*-model:
$\sigma_{\mathrm{M}}^2 = Q_{yy}$, $\sigma_{\mathrm{MH}} = Q_{xy}$, $\sigma_D^2 = \sigma_D^2$,
$D_2^* = Q_{xx}$, $H^* = \mu_\infty^2$.

$(1+F)Q_{yy}/2 = (1+F)\sigma_A^2$. If $Q_{xx} = 0$ then it must also be true that $Q_{xy} = 0$. Furthermore, $Q_{xx} = 0$ if and only if $Q_{xxi} = 0$ at all loci. Therefore, situations which give $Q_{xxi} = 0$ are of particular interest.

Whatever the number of alleles per locus, Q_{xxi} is zero if there is no dominance at the i^{th} locus. When there are only two alleles per locus, $Q_{xxi} = p_1 p_2 (p_1 - p_2)^2 x_{12}^2$. Clearly, no dominance and/or $p_1 = p_2 = 0.5$ at every locus are the only circumstances that will give $Q_{xx} = 0$ if there are only two alleles at all loci.

With more than two alleles at a locus, however, other situations which give $Q_{xxi} = 0$ can occur. It is shown in the Appendix that $Q_{xxi} = 0$ if the $\bar{x}_{.k}$ are all equal and this further implies that

$$x_{jk} = [2\textstyle\sum_{l<m\,(l,m\neq j)} p_l p_m x_{lm} - (1-2p_j)\sum_{l\neq j,k} p_l x_{kl}]/p_j(1-2p_j) \tag{1}$$

for some choice of j for which $p_j \neq 0.5$, all $k \neq j$ and x_{lm} and x_{kl} within the brackets are arbitrary. The notation $l, m \neq j$ should be understood to indicate that neither l nor m is equal to j and $l \neq j, k$ indicates that l is not equal to j or k. It is also shown in the Appendix that $Q_{xxi} = 0$ if

$$x_{jk} = c - 2\textstyle\sum_{l\neq j,k} p_l x_{kl} \tag{2}$$

for some choice of j such that $p_j = 0.5$, all $k \neq j$, where c is an arbitrary constant and x_{lm} for $l, m \neq j$ are such that

$$\textstyle\sum_{l<m\,(l,m\neq j)} p_l p_m x_{lm} = 0. \tag{3}$$

It can be shown by direct substitution that if (1) holds for some choice of j for which $p_j \neq 0.5$, then it also holds for all other choices of j for which $p_j \neq 0.5$ and (2) and (3) also hold for any other choice of j for which $p_j = 0.5$. (Note that, if more than two alleles exist, then there cannot be more than one choice of j for which $p_j = 0.5$.) Conversely, if (2) and (3) hold for some choice of j for which $p_j = 0.5$, this is sufficient to assure that (1) holds for all other choices of j. Suppose, for the moment, that $p_j \neq 0.5$ for some choice of j. Then, if $Q_{xxi} = 0$, x_{jk} is given by (1), but, since $x_{jk} = x_{kj}, x_{jk}$ is also given by (1) with subscripts j and k interchanged if $p_k \neq 0.5$ or by (2) with subscripts j and k interchanged if $p_k = 0.5$. To illustrate the expansion of (1), (2) and (3), let $n = 5$ alleles, $j = 2$, $k = 3$. Then (1) gives $x_{23} = [2(p_1 p_3 x_{13} + p_1 p_4 x_{14} + p_1 p_5 x_{15} + p_3 p_4 x_{34} + p_3 p_5 x_{35} + p_4 p_5 x_{45}) - (1-2p_2)(p_1 x_{13} + p_4 x_{34} + p_5 x_{35})]/p_2(1-2p_2)$ provided $p_2 \neq 0.5$ and also $x_{23} = x_{32} = [2(p_1 p_2 x_{12} + p_1 p_4 x_{14} + p_1 p_5 x_{15} + p_2 p_4 x_{24} + p_2 p_5 x_{25} + p_4 p_5 x_{45}) - (1-2p_3)(p_1 x_{12} + p_4 x_{24} + p_5 x_{25})]/p_3(1-2p_3)$ provided $p_3 \neq 0.5$. If $p_3 = 0.5$ then (2) gives $x_{23} = x_{32} = c - 2(p_1 x_{12} + p_4 x_{24} + p_5 x_{25})$ and the constraint (3) is $p_1 p_2 x_{12} + p_1 p_4 x_{14} + p_1 p_5 x_{15} + p_2 p_4 x_{24} + p_2 p_5 x_{25} + p_4 p_5 x_{45} = 0$. Example 1 in Table 2 gives a set of hypothetical values for the g_{jk} and x_{jk} which give $Q_{xxi} = 0$ if $p_1 = p_2 = 0.1$, $p_3 = 0.2$, $p_4 = p_5 = 0.3$. The hypothetical values represent only one out of an infinite number of possibilities which will give $Q_{xxi} = 0$, given the gene frequencies and genotypic values of the homozygotes (g_{jj}).

An important consequence of (3) is that, if one $p_j = 0.5$, and $Q_{xxi} = 0$, then among the x_{lm} (for $l, m \neq j$), provided they are not all zero, there must exist both positive and negative values, i.e., one or more heterozygotes must be such that the favorable allele is dominant, but one or more other heterozygotes must be such that the unfavorable allele is dominant.

Populations synthesized by the intermating of homozygous lines provide important special cases in that, if one supposes that each parent line contributes a different allele to the population, then the gene frequencies are known. If n lines contribute equally to such a population, then the p_j are all equal to $1/n$. In such a case (1) reduces to

$$x_{jk} = [2\textstyle\sum_{l<m\,(l,m\neq j)} x_{lm} - (n-2)\sum_{l\neq j,k} x_{kl}]/(n-2). \tag{4}$$

Example 2 in Table 2 gives a hypothetical example which gives $Q_{xxi} = 0$ with five alleles all having frequency 0.2. The $\bar{x}_{.k}$ are obviously all equal and equation (4) is always satisfied if the alleles are equally frequent and the x_{jk} are all equal for all j and k ($j \neq k$). Example 3 in Table 2 has this property.

Another interesting special case is where only one allele (the j^{th}, say) shows dominance (or recessiveness) in combination with other alleles, all other alleles behaving additively when in combination with one another. In this situation, $x_{jk} \neq 0$, but $x_{kl} = 0$ for all $k, l \neq j$. It is shown in the Appendix that in this situation Q_{xxi} can be zero only if $p_j = 0.5$ and it then follows from (2) that the x_{jk} must be all equal for $k \neq j$. Example 4 in Table 2 has this structure.

In cases of three or four alleles $Q_{xxi} = 0$ requires constraints which are not necessary with higher numbers of alleles. With three alleles with all $p_j \neq 0.5$, (1) requires that

$$x_{jk} = p_l(1-2p_l)w, \tag{5}$$

for j, k and l running through all permutations of the integers $\{1, 2, 3\}$, where w is an arbitrary constant. This may be a reasonable possibility if all $p_j < 0.5$. However, if any $p_j > 0.5$, it would require both positive and negative dominance contrasts at the given locus. Equation (5) implies that if there are three alleles with $p_1 = p_2 = p_3 = 1/3$, then it is not only sufficient, but necessary, that $x_{12} = x_{13} = x_{23}$. Equation (5) further implies that, provided all $p_j \neq 0.5$, if any $x_{jk} = 0$ then Q_{xxi} cannot be zero unless all $x_{jk} = 0$ (i.e., all gene action is additive). If $p_j = 0.5$ for a particular choice of j then (2) and (3) only admit solutions of the form $x_{jk} = x_{jl} = c$ and $x_{kl} = 0$ for $k, l \neq j$. Populations constructed from three-way crosses using homozygous parents provide an interesting application if we suppose that each of the three parents contributes a different allele to the population. If populations are developed from each of the three possible three-way crosses, Q_{xxi} cannot be zero in all three populations unless all gene action at the i^{th} locus is additive.

It is shown in the Appendix that if there are four alleles, then $Q_{xxi} = 0$ if

$$x_{lm} = [(2p_l + 2p_m - 1)w + 2p_j p_k x_{jk}]/2p_l p_m \tag{6}$$

for j, k, l and m running through all permutations of the integers $\{1, 2, 3, 4\}$ and $w = \bar{x}_{.j} = \bar{x}_{.k} = \bar{x}_{.l} = \bar{x}_{.m}$. To find a solution which satisfies (6) it is sufficient to choose $\{j, k, l, m\}$ to be a particular permutation of $\{1, 2, 3, 4\}$, choose arbitrary values for x_{jk}, x_{jl} and x_{jm}, put $w = \bar{x}_{.j}$ and compute (6) letting k, l, m run through all permutations of the subset of three integers obtained from $\{1, 2, 3, 4\}$ by deleting the integer chosen for j. It can be shown by direct substitution that if (6) holds for one choice of j under all permutations of k, l and m, then it also holds for all other choices of j. Also, $\bar{x}_{.j} = \bar{x}_{.k}$ is easily manipulated algebraically to give

$$x_{jk} = [p_l(x_{kl} - x_{jl}) + p_m(x_{km} - x_{jm})]/(p_k - p_j) \tag{7}$$

provided that $p_j \neq p_k$. If $p_j = p_k$ then $Q_{xxi} = 0$ will require that $p_l(x_{kl} - x_{jl}) + p_m(x_{km} - x_{jm}) = 0$. If $p_l + p_m = 0.5$ then (6) reduces to $x_{lm} = p_j p_k x_{jk}/p_l p_m$. Consequently, if $p_1 = p_2 = p_3 = p_4 = 0.25$, then $Q_{xxi} = 0$ requires $x_{12} = x_{34}$, $x_{13} = x_{24}$ and $x_{14} = x_{23}$. Finally, if, for some choice of j, $x_{jk} = 0$ for all $k \neq j$, then (6) shows that all $x_{lm} = 0$. Thus, if there are four alleles and one of them is such that it does not show dominance (or recessiveness) in combination with any of the other alleles, then $Q_{xxi} = 0$ if and only if all gene action at the locus is additive. It is clear from (5) that this also holds for $n = 3$, but it does not necessarily hold for $n > 4$.

Both Q_{xx} and σ_D^2 are nonnegative quadratic functions of $\boldsymbol{x}$ vectors for the various loci. As has been noted, it is possible, though perhaps unlikely, for Q_{xxi} to be zero even if vector $\boldsymbol{x}$ is not

Table 2. Hypothetical examples with five alleles which give $Q_{xxi} = 0$

Examples 1 and 2

Example 1 above diagonal; $Q_{xxi} = 0$ if $p_1 = p_2 = 0.1$, $p_3 = 0.2$, $p_4 = p_5 = 0.3$.
Example 2 below diagonal; $Q_{xxi} = 0$ if all $p_j = 0.2$.

Genotypic values[a] (g_{jk})

Allele	I_1	I_2	I_3	I_4	I_5
I_1	24	39	57	71	79
I_2	46	48	66	84	92
I_3	64	66	72	92	108
I_4	70	84	90	96	120
I_5	76	96	108	120	120

Dominance contrasts[a] (x_{jk})

Allele	I_1	I_2	I_3	I_4	I_5
I_1	–	−6	−18	−22	−14
I_2	−20	–	−12	−24	−16
I_3	−32	−12	–	−16	−24
I_4	−20	−24	−12	–	−24
I_5	−8	−24	−24	−24	–

Examples 3 and 4

Example 3 above diagonal; $Q_{xxi} = 0$ if all $p_j = 0.2$.
Example 4 below diagonal; $Q_{xxi} = 0$ if $p_1 = 0.5$, $p_2 + p_3 + p_4 + p_5 = 0.5$.

Genotypic values[a] (g_{jk})

Allele	I_1	I_2	I_3	I_4	I_5
I_1	24	48	60	72	84
I_2	48	48	72	84	96
I_3	60	60	72	96	108
I_4	72	72	84	96	120
I_5	84	84	96	108	120

Dominance contrasts[a] (x_{jk})

Allele	I_1	I_2	I_3	I_4	I_5
I_1	–	−24	−24	−24	−24
I_2	−24	–	−24	−24	−24
I_3	−24	0	–	−24	−24
I_4	−24	0	0	–	−24
I_5	−24	0	0	0	–

[a] Addition of any constant to the genotypic values and multiplication of the genotypic values and the dominance contrasts by any constant also gives $Q_{xxi} = 0$

zero. On the other hand, σ_D^2 cannot be zero unless $\boldsymbol{x}$ is zero at all loci. Is it possible (or likely) for Q_{xx} to be larger than σ_D^2? This appears to be a difficult question to answer in multiple allelic situations, but with two alleles at the i^{th} locus, the contribution of the i^{th} locus to σ_D^2 is $\sigma_{di}^2 = p_1^2 p_2^2 x_{12}^2$. Comparing this to $Q_{xxi} = p_1 p_2 (p_1 - p_2)^2 x_{12}^2$, it is apparent that Q_{xxi} and σ_{di}^2 are equal if $p_1 p_2 = (p_1 - p_2)^2 = 1 - 4 p_1 p_2$ which is satisfied if either p_1 or p_2 is $(1 + 1/\sqrt{5})/2 = 0.7236$; Q_{xxi} will be the larger if either p_1 or p_2 exceeds this value and σ_{di}^2 will be the larger if both p_1 and p_2 are less than 0.7236. When $Q_{xxi} = \sigma_{di}^2$, both are equal to $0.64\, x_{12}^2$. Set $Q_{xxi}/\sigma_{di}^2 = r$. Then $p_1 = 0.5\,\{1 \pm [r/(r+4)]^{1/2}\}$, p_2 is the same expression with a reversal of the $\pm$ sign, and the absolute value of $p_1 - p_2$ is $[r/(r+4)]^{1/2}$. Thus, if there are only two alleles per locus and $r = Q_{xx}/\sigma_D^2$, then $[r/(r+4)]^{1/2}$ might be viewed as an estimate of the "average" absolute difference in gene frequency between dominant and recessive alleles.

Cornelius and Dudley (1976) noted that, with two alleles per locus, $(\sum p q x^2 / \sum p q y^2)^{1/2}$ where $p = p_2$, $q = p_1$, and the sums are over loci, gives an estimate of average degree of dominance. In Q-model notation, this can be expressed as $[(Q_{xx} - 4\sigma_D^2)/Q_{yy}]^{1/2}$. In the case of two alleles per locus, $\sigma_D^2 = \breve{H}$ (alias μ_∞^2), and in such a situation presumably the best estimate of σ_D^2 for use in this formula or the previous formula for r will be obtained by combining the terms in these components in the expectations of the empirical covariances used for estimation.

Properties of Q_{xy}

Any expression for the covariance of inbred relatives must, except under very special conditions, contain a component which includes the contribution of the products of the x's and the y's. D_1 serves that purpose in the D-model and C does so for the C-model. Both of these are functions of Q_{xy} along with another component and Q_{xy} itself may be the more interesting component since it alone measures the association of x and y values.

If we characterize the quantity $d_{kk} - \sum_k p_k d_{kk}$ (where d_{kk} is the dominance deviation of $I_k I_k$ at panmixia) as the within-locus "inbreeding depression effect" of the i^{th} allele, then Q_{xy} is the covariance of effects of alleles at complete homozygosity with the inbreeding depression effects of those alleles. The correlation of such effects is $\varrho_{d\infty} = Q_{xy}/(Q_{yy} Q_{xx})^{1/2}$. This same quantity is the correlation of the $\tilde{x}_{\cdot k}$ with the $\tilde{y}_{\cdot k}$.

Q_{xy} is necessarily zero if $Q_{xx} = 0$, but Q_{xy} may be zero even if $Q_{xx} \neq 0$. If there are only two alleles at the i^{th} locus $Q_{xyi} = p_1 p_2 (p_1 - p_2)\, x_{12} y_{12} = p_1 p_2 (p_1 - p_2)\, a_{12} y_{12}^2$ which is zero if any one of the following hold: (1) there is no dominance ($x_{12} = 0$), (2) there is overdominance with homozygotes having equal value, ($g_{11} = g_{22} \neq g_{12}$ resulting in $y_{12} = 0, x_{12} \neq 0$), (3) gene frequencies are 0.5 ($p_1 - p_2 = 0$). Q_{xyi} is positive if the dominant gene is the more frequent and is negative if the dominant gene is the allele of lesser frequency. Clearly, both positive and negative values of Q_{xyi} could occur at the various loci even if all loci have similar directions and degrees of dominance. Thus $Q_{xy} = 0$ would not necessarily imply any of the situations which make $Q_{xyi} = 0$.

With respect to a single locus with two alleles the correlation $Q_{xyi}/(Q_{yyi} Q_{xxi})^{1/2}$ is ± 1 with the sign determined by the sign of Q_{xyi}. We might loosely interpret $\varrho_{d\infty}$ as a sort of average of these correlations for segregating loci ("loosely" because pooling the components for all loci and then computing the correlation is not equivalent to a simple averaging of the correlations for the

loci separately). A positive value of $\varrho_{d\infty}$ would be expected if dominant alleles tend to be more frequent than recessive alleles.

If there are two alleles at all loci with the same pair of values for p_1 and p_2 ($p_1 \neq p_2$), a situation for which one could easily develop a population to satisfy, then the correlation $\varrho_{d\infty}$ is $\pm\sum x_{12}y_{12}/(\sum y_{12}^2 \sum x_{12}^2)^{1/2}$ where the sums are over loci and the sign is determined by the sign of $p_1 - p_2$. If $p_1 > p_2$, $\varrho_{d\infty}$ may be interpreted as the correlation of the x contrasts with the y contrasts. The correlation can still be either positive or negative because, for the i^{th} locus, $x_{12}y_{12} = a_{12}y_{12}^2$ which is negative if I_2 is dominant over I_1. Using a population constructed from two homozygous parents represented in different dosages, this statistic might give some insight into which parent carries more of the alleles which are dominant (or at least more of the dominant alleles at loci which have large dominance effects). Presumably, if one wishes to do this, one should choose p_1 and p_2 so as to maximize the absolute value of $p_1p_2(p_1 - p_2)$. This gives $p_1 = 0.5 \pm 0.29$. This is closely approximated in populations derived from backcrosses where $p_1 = 0.75$ and $p_2 = 0.25$ or vice versa. If populations derived from reciprocal backcrosses are studied, Q_{xy} should be of the same absolute magnitude in both populations, but should differ in sign.

With more than two alleles, expansion of $Q_{xyi} = \sum_k p_k \bar{x}_{\cdot k}\bar{y}_{\cdot k}$ gives

$$Q_{xyi} = \sum_{j<k} p_j p_k (p_j - p_k) x_{jk} y_{jk} \qquad (8)$$
$$+ \sum_{j<k<l} p_j p_k p_l [y_{jk}(x_{kl} - x_{jl}) + y_{jl}(x_{kl} - x_{jk}) + y_{kl}(x_{jl} - x_{jk})].$$

The quantities $p_j p_k (p_j - p_k) x_{jk} y_{jk}$ are all zero if gene frequencies are all equal. As in the case of only two alleles, these quantities are positive (negative) if the dominant allele of the pair (I_j, I_k) is the more (less) frequent. All terms in Q_{xyi} are zero if there is no dominance (x_{jk} all zero) or if homozygotes are all of equal value (y_{jk} all zero). The difference $x_{kl} - x_{jl} = (g_{kk} - g_{jj}) - 2(g_{kl} - g_{jl})$. This is a contrast of the double substitution of allele I_k for allele I_j in the homozygous state compared to 2 single substitutions in the presence of allele I_l. The other differences $x_{kl} - x_{jk}$ and $x_{jl} - x_{jk}$ are structured similarly with a permutation of subscripts. The second sum of terms in (8) is zero if such quantities are all zero. Q_{xyi} could be small (or even zero) through the cancellation of positive and negative quantities. Furthermore, as in the case of two alleles, positive and negative Q_{xyi} values could cancel in the summation over loci which gives Q_{xy}.

Clearly, unless one has prior knowledge of the number of alleles and their frequencies, little can be concluded from an experimentally estimated value of Q_{xy}, particularly if the estimate is of small magnitude. A large positive Q_{xy} might suggest that the dominant alleles are present in fairly high frequency (but not extremely high because Q_{xyi} approaches zero as the frequency of any allele at the i^{th} locus approaches unity). Negative Q_{xy} would suggest that such alleles are in low frequency.

Other properties of parameters in the models

If $Q_{xx} = 0$, then $D_1 = 0$, but it is possible to have $D_1 = 0$, even if $Q_{xx} \neq 0$. Let $d_{1i} = (Q_{xyi} - Q_{xxi})/2$ and $\sigma_{ai}^2 = (Q_{yyi} - 2Q_{xyi} + Q_{xxi})/2$ denote the contributions of the i^{th} locus to D_1 and σ_A^2, respectively. With two alleles per locus and letting $x_{12} = a_{12}y_{12}$ we can write $d_{1i} = (1/2)p_1p_2[1 - (p_1 - p_2)a_{12}](p_1 - p_2)a_{12}y_{12}^2$. If the quantity $1 - (p_1 - p_2)a_{12} = 0$ then the additive effects (at panmixia) are zero and d_{1i} is zero as is also σ_{ai}^2. This requires overdominance equal to $a_{12} = 1/(p_1 - p_2)$ and is unlikely to occur except in hypothetical examples. The d_{1i} values for the various loci can be positive or negative so that D_1 can also be zero as a result of cancellation of positive and negative values when summed over loci. Of interest, therefore, are the situations which yield $d_{1i} > 0$ and those where $d_{1i} < 0$. In particular, $d_{1i} > 0$ in two situations: (1) $p_1 > p_2$ and $0 < a_{12} < 1/(p_1 - p_2)$ and (2) $p_1 < p_2$ and $-1/(p_2 - p_1) < a_{12} < 0$. Thus, if the dominant allele is the more frequent, d_{1i} will be non-negative unless there is overdominance of sufficiently large magnitude (degree of dominance greater in absolute value than the reciprocal of the difference between gene frequencies). The exception includes the case where $g_{11} = g_{22} \neq g_{12}$. A negative d_{1i} will always occur if the recessive allele has the higher frequency. (In this regard, note that any situation which gives a negative Q_{xyi} must necessarily give a still more negative d_{1i}).

If or to what extent these properties of D_1 hold in multiple allelic situations is not easily assessed since d_{1i} involves all of the $[n(n-1)/2]^2$ possible products of the form $x_{jk}y_{lm}(j < k, l < m)$ and all squares and pairwise products among the x_{jk}. It is always true, however, that $D_1 \leq (1/2)Q_{xy}$ with equality only if $Q_{xx} = 0$, which will require that $D_1 = Q_{xy} = 0$ (but $D_1 = 0$ or $Q_{xy} = 0$ does not imply that $Q_{xx} = 0$).

Though it is not likely to occur in any experimental or breeding population, $\sigma_\infty^2 (= Q_{yy}) = 0$ would imply that effects of genes at complete homozygosity are all zero. For the i^{th} locus, in D-model notation for gene effects (Table 1), we would have $2\alpha_j + (d_{jj} - h_j) = 0$ where $h_i = \sum_j p_j d_{jj}$. In this very hypothetical situation, $Q_{xx} = 2\sigma_A^2$ and $D_1 = -\sigma_A^2$. Thus, $-\sigma_A^2$ places a lower limit on D_1, which can be reached only if all homozygotes are of equal value.

In any case, D_1 and Q_{xx} (alias D_2) are highly dependent on gene frequencies and their usefulness would be greatly enhanced in situations where one knows the number of alleles and their frequencies a priori (as in the case of a population constructed by the mating of homozygous lines). In the absence of such knowledge, estimates of D_1 and Q_{xx} may, if the estimates are large, give some information (though that information is far from clear) concerning the genetics of the character under investigation. If the estimates are small, they are virtually noninformative, since a myriad of circumstances can operate to make them small.

The correlation of inbreeding depression effects of alleles with additive genetic effects at panmixia is given by $\varrho_{da} = D_1/(D_2\sigma_A^2/2)^{1/2}$. Another correlation which may be interesting is that of additive effects of alleles at panmixia with effects of those alleles at complete homozygosity, i.e., $\varrho_{a\infty} = C/(2\sigma_A^2\sigma_\infty^2)^{1/2}$. For a single locus with two alleles, both of these correlations are ± 1 with the sign of the first depending on whether the quantities $(p_1 - p_2)a_{12}$ and $1 - (p_1 - p_2)a_{12}$ are of like or unlike sign, and the sign of the second determined by the sign of $1 - (p_1 - p_2)a_{12}$. A negative value of $1 - (p_1 - p_2)a_{12}$ results only if the dominant allele is the more frequent and there is overdominance of sufficient degree that $|a_{12}| > 1/|p_1 - p_2|$. These same conditions will cause $1 - (p_1 - p_2)a_{12}$ and $(p_1 - p_2)a_{12}$ to differ in sign. Furthermore, they will always differ in sign if the recessive allele is the more frequent. As with the correlation $\varrho_{d\infty}$ which we have previously discussed, we might cautiously regard a computed correlation as a sort of average of these correlations for separate loci.

Estimates of D_1, Q_{xx} (alias D_2) and Q_{xy} in a maize population

Results of an experiment evaluating various kinds of inbred relatives derived by selfing and full-sib mating in a broad-base maize population (Synthetic O.P.) were reported by Cornelius and Dudley (1974, 1976). Experimental material and methods were described previously. A path diagram of the mating scheme by which families

Table 3. Estimates of quadratic components of covariance in a maize population

Year	$\sigma^2_\infty (= Q_{yy})$	$\mu^2_\infty (= \check{H})$	C	σ^2_A	σ^2_D	Q_{xy}	D_1	$Q_{xx} (= D_2)$
	Plant height (cm)							
1970	458 ± 133[a]	455 ± 511	540 ± 148	312 ± 106	−101 ± 157	−81 ± 175	− 42 ± 95	2 ± 258
1971	454 ± 111	1311 ± 461	473 ± 188	448 ± 184	2 ± 240	−19 ± 187	−212 ± 166	405 ± 381
Combined[c]	451 ± 88	1045 ± 341	484 ± 117	370 ± 99	− 57 ± 139	−33 ± 125	−129 ± 89	225 ± 220
	Ear height (cm)							
1970	143 ± 25	653 ± 371	378 ± 103	210 ± 92	77 ± 131	−236 ± 114	− 20 ± 78	−195 ± 182
1971	291 ± 74	979 ± 330	301 ± 146	771 ± 174	−408 ± 141	− 11 ± 153	−621 ± 171	1231 ± 385
Combined[c]	249 ± 53	430 ± 238	315 ± 85	382 ± 83	−103 ± 98	− 66 ± 92	−225 ± 76	383 ± 179
	Yield (g/plant)							
1970	370 ± 290	3077 ± 1615	111 ± 275	11 ± 282	1082 ± 522	259 ± 458	44 ± 293	170 ± 852
1971	290 ± 125	2961 ± 1170	42 ± 218	−240 ± 218	1676 ± 751	248 ± 265	261 ± 391	−273 ± 859
Combined[c]	280 ± 114	3286 ± 886[b]	79 ± 167	−125 ± 231	1403 ± 436	201 ± 216	165 ± 234	−129 ± 552
	Grain moisture (%)							
1970	12.6 ± 3.3	19.6 ± 10.4	13.2 ± 3.4	7.4 ± 2.1	−3.5 ± 2.9	−0.6 ± 4.2	−0.8 ± 1.9	0.9 ± 5.7
1971	14.1 ± 2.8	10.8 ± 9.3	16.2 ± 3.7	4.6 ± 2.7	0.8 ± 4.5	−2.1 ± 2.9	3.5 ± 2.1	−9.2 ± 4.9
Combined[c]	13.1 ± 2.1	15.3 ± 6.9	14.9 ± 2.5	5.9 ± 1.7	−1.0 ± 2.5	−1.9 ± 2.3	1.6 ± 1.4	−5.0 ± 3.6
	Oil (%)							
1970	0.264 ± 0.083	0.520 ± 0.332	0.308 ± 0.090	0.167 ± 0.067	−0.027 ± 0.106	−0.044 ± 0.111	−0.013 ± 0.061	−0.019 ± 0.166
1971	0.331 ± 0.069	0.205 ± 0.243	0.317 ± 0.128	0.129 ± 0.090	0.123 ± 0.139	0.013 ± 0.097	0.030 ± 0.078	−0.046 ± 0.182
Combined[c]	0.310 ± 0.052	0.311 ± 0.191	0.307 ± 0.068	0.145 ± 0.054	0.050 ± 0.084	0.003 ± 0.070	0.009 ± 0.048	−0.015 ± 0.120
	Kernel weight (g/100)							
1970[d]	33.4 ± 8.6	8.5 ± 32.4	25.6 ± 9.1	20.7 ± 6.3	−6.7 ± 8.7	7.8 ± 12.0	−7.9 ± 6.0	23.6 ± 17.7
1971[d]	26.5 ± 5.5	25.8 ± 17.8	25.0 ± 6.9	6.8 ± 5.6	4.1 ± 9.8	1.5 ± 6.2	5.7 ± 4.8	−10.0 ± 11.3
Combined	30.7 ± 4.8	15.3 ± 16.0	27.8 ± 5.8	14.7 ± 4.2	−2.3 ± 6.3	3.0 ± 6.2	−0.8 ± 3.8	4.5 ± 9.9

[a] Standard error of estimated component
[b] Estimate of μ^2_∞ was erroneously reported as 2386 by Cornelius and Dudley (1976)
[c] Cornelius and Dudley (1976) reported significant heterogeneity between years, not resolved as to whether it was genetic or environmental
[d] Model showed significant lack of fit to within-year data (Cornelius and Dudley 1976)

of progenies were produced is given in the 1976 paper which also gives expressions in C-model parameters for the genetic covariances among the progenies derived by selfing and covariances of the selfed progenies with the progenies derived by sib-mating. Covariances among the sib-mated progenies were derived in C-model parameterization by Cornelius and Dudley (1975). (The coefficient of C in the covariance of generation-1 uncle and generation-2 nephew is erroneously given as 0.625. The correct value is 0.0625.) Families were allocated to sets in a split-plot arrangement and Cornelius and Dudley (1976) presented estimates of σ^2_∞, μ^2_∞, σ^2_A and σ^2_D computed by applying the method described by Cornelius and Byars (1977) to family within set, error and, for characters where individual plants were measured, plants within plots mean square and cross product matrices. The purpose of the present analysis is to present the estimates of D_1, Q_{xx} (alias D_2) and Q_{xy} (Table 3) from these data. These estimates and their standard errors were computed by simple transformations of the previously computed solutions. For example, if we let $\boldsymbol{\Phi} = (\sigma^2_\infty, \mu^2_\infty, C, \sigma^2_A, \sigma^2_D)'$, then $\hat{D}_1 = \boldsymbol{K}\hat{\boldsymbol{\Phi}}$, where $\boldsymbol{K} = (0, 0, 0.5, -1, 0)$, and $V(\hat{D}_1) = \boldsymbol{K}\,V(\hat{\boldsymbol{\Phi}})\,\boldsymbol{K}'$, where $V(\hat{\boldsymbol{\Phi}})$ is the estimated covariance matrix of $\hat{\boldsymbol{\Phi}}$. The estimates of components from the original study are also included in Table 3 for comparison with D_1, Q_{xx} and Q_{xy}.

Cornelius and Dudley (1976) reported that all characters except kernel weight showed significant heterogeneity of variance between years, which, however, was not resolved with respect to whether it was genetic (i.e., genotype × year interaction), environmental or both. It was not feasible with the available material to design the experiment so that estimates of interactions of the quadratic components of genetic covariance with years could be easily obtained. Therefore, the estimates from years combined are a sort of average for the two years, which in some cases may be averaging estimates of unequal parametric values.

The most striking results, in particular the large estimates of the dominance components μ^2_∞ (alias $\bar{H}$) and σ^2_D for yield and the large estimate of μ^2_∞ for plant and ear height in 1971, have been previously discussed (Cornelius and Dudley 1976). Of interest at present are the estimates of Q_{xy}, D_1 and Q_{xx}. Most of these estimates, when compared with their asymptotic standard errors, were small and could hardly be considered statistically significant. In absolute value, Q_{xy} exceeded 1.96 times its standard error (the 0.05-level test criterion based on asymptotic normality of the estimates) only in the case of ear height in 1970. However, this was accompanied by a negative estimate of Q_{xx}. If $Q_{xx} = 0$, this implies that also $Q_{xy} = 0$. It, therefore, makes sense to test the hypothesis $Q_{xy} = Q_{xx} = 0$, or its equivalent in the C-model ($\sigma^2_\infty = C = 2\sigma^2_A$) or the D-model ($D_1 = D_2 = 0$). An approximate chi-square (Table 4) computed from the parameter estimates and their asymptotic covariance matrix indicated that the hypothesis $Q_{xy} = Q_{xx} = 0$ was an acceptable hypothesis in all cases except ear height in 1971 and in the analysis with years combined. $Q_{xy} = Q_{xx} = 0$ implies that $D_1 = 0$. Examined separately, D_1 and D_2 (alias Q_{xx}) (Table 3) were also nonsignificant in all cases where $Q_{xy} = Q_{xx} = 0$ was an acceptable hypothesis.

If there are only two alleles at all loci where there is dominance, $\mu^2_\infty = \sigma^2_D$. Approximate chi-square tests (Table 4) indicated that this hypothesis was acceptable in all cases except plant and ear height in 1971 and in the

Table 4. Approximate chi-square tests of hypotheses concerning covariance components

Year	Chi-square[a] for test of hypothesis			
	H_{01}: $Q_{xy} = Q_{xx} = 0$[b]	H_{02}: $\mu^2_\infty = \sigma^2_D$	H_{01} and H_{02}[c]	No dominance
	Plant height			
1970	0.41	0.82	0.82	0.82
1971	1.63	4.38*	5.59	19.60**
Combined[d]	2.18	6.52*	8.12*	17.43**
	Ear height			
1970	4.28	1.59	4.30	11.75*
1971	13.34**	11.34**	25.60**	30.59**
Combined[d]	9.39**	11.81**	16.80**	35.82**
	Yield			
1970	0.44	1.13	7.11	25.60**
1971	1.53	0.56	8.17*	42.26**
Combined[d]	1.69	2.63	16.37**	75.78**
	Moisture			
1970	0.19	3.52	6.03	6.05
1971	3.55	0.68	9.03*	9.43
Combined[d]	2.01	3.73	12.19**	12.26**
	Oil			
1970	0.21	1.85	2.63	3.66
1971	0.16	0.06	0.67	3.91
Combined[d]	0.03	1.12	2.82	6.94
	Kernel weight			
1970[e]	2.01	0.15	2.15	2.17
1971[e]	1.47	0.79	8.63*	8.71
Combined	0.27	0.77	2.53	2.63

*,** Significant at $p = 0.05$ and 0.01, respectively

[a] Df of chi-square is 2, 1, 3, and 4 for the four hypotheses

[b] Equivalently, $\sigma^2_\infty = C = 2\sigma^2_A$ in the C-model or $D_1 = D_2 = 0$ in the D-model

[c] Equivalently, $p_1 = p_2 = 0.5$ at all loci where there is dominance

[d] Cornelius and Dudley (1976) reported significant heterogeneity between years not resolved as to whether it was genetic or environmental

[e] C-model showed significant lack of fit to within-year data (Cornelius and Dudley 1976)

combined analysis. Cornelius and Dudley (1976) also tested this hypothesis, but they used a different criterion, namely, by comparing goodness of fit of multiple-allele and two-allele models. The conclusions rendered by the two methods were consistent except for yield in the combined analysis where Cornelius and Dudley found the multiple-allele model to fit significantly better. $Q_{xy} = Q_{xx} = 0$ and $\mu_\infty^2 = \sigma_D^2$ jointly imply that gene frequencies are all 0.5 at loci where there is any dominance. However, separate acceptability of these two hypotheses does not necessarily imply that they are jointly acceptable. Indeed, they were separately, but not jointly, acceptable for yield and moisture in 1971 and years combined and for kernel weight in 1971. Where the no epistasis model gave a satisfactory fit (not the case for kernel weight in 1971), separate, but not joint acceptance of the hypotheses $Q_{xy} = Q_{xx} = 0$ and $\mu_\infty^2 = \sigma_D^2$ implies that the data are consistent with either of two conclusions concerning loci where there is dominance: (1) there are two alleles per locus and gene frequencies are not 0.5, or (2) there are multiple-alleles, but $Q_{xy} = Q_{xx} = 0$ (or at least close enough to zero to escape detection). Either conclusion is a distinct possibility, but the first would seem to be the more likely, as it was previously made clear that gene effects must satisfy severe constraints in order for Q_{xx} to be zero in multiple-allelic situations.

Synthetic O.P. is a composite of open-pollinated cultivars and there is little reason to believe gene frequencies are all exactly 0.5 even if such a hypothesis is statistically acceptable. The hypotheses $Q_{xy} = Q_{xx} = 0$ and $\mu_\infty^2 = \sigma_D^2$ will always be jointly acceptable if there is no dominance (or not enough to be statistically detectable). This is, apparently, true for percent oil in both years and for moisture in 1970. They will also be jointly acceptable if there are two alleles per locus, but loci where there is dominance with respect to effects on the character studied do not have gene frequencies sufficiently different from 0.5 to make Q_{xy} and/or Q_{xx} sufficiently large to be statistically detectable. This may be the situation for ear height in 1970.

For ear height in 1971, Q_{xx} was the largest component, Q_{xy} was negligible and D_1 was negative and significant owing, of course, to the contribution of Q_{xx} to D_1. However, a negative estimate was obtained for σ_D^2 and this is inconsistent with large values for Q_{xx} and μ_∞^2. If the negative σ_D^2 is regarded as an estimate of a true value of zero, this would imply that both μ_∞^2 and Q_{xx} must also be zero which would further imply that all gene action is additive. However, an additive model does not adequately model the variances and covariances (Table 4).

With two exceptions estimates of $\varrho_{d\infty}$ (Table 5), where they could be computed, were of small magnitude. In some cases estimates of $\varrho_{d\infty}$ and ϱ_{da} could not be computed owing to a negative estimate of Q_{xx}. If this is regarded as an estimate of $Q_{xx} = 0$, this implies that $\varrho_{d\infty} = \varrho_{da} = 0$. With one exception, estimates of ϱ_{da}, where they could be computed, were negative. Usually they were more negative than $\varrho_{d\infty}$. Estimates of $\varrho_{a\infty}$, wherever they could be computed, were positive and generally large, but they were not as large for plant and ear height (characters which showed inbreeding depression) in 1971 and years combined as for some of the other characters. Moisture and percent oil, for which a strictly additive variance model gave an adequate fit (Cornelius and Dudley 1976) gave results consistently suggesting $\varrho_{a\infty} = 1$. No estimate of $\varrho_{a\infty}$ could be obtained for yield in 1971 and years combined owing to a negative estimate of σ_A^2. The estimate of σ_A^2 for yield in 1970, though positive, was extremely small and it is quite probable that $\varrho_{a\infty} = 1.23$ is spuriously large as a consequence. No attempt was made to compute standard errors for the correlations, but they are no doubt high, especially for ϱ_{da} where there is the greatest intercorrelation among the estimated components involved.

Table 5. Correlations of effects computed from covariance component estimates

Character	Correlation			
	Year	$\varrho_{d\infty}$	$\varrho_{a\infty}$	ϱ_{da}
Plant height	1970	−2.68	1.01	−2.38
	1971	−0.04	0.74	−0.70
	Combined	−0.10	0.84	−1.58
Ear height	1970	–[a]	1.54	–[a]
	1971	−0.02	0.45	−1.11
	Combined	−0.21	0.72	−1.11
Yield	1970	1.03	1.23	1.44
	1971	–[a]	–[a]	–[a]
	Combined	–[a]	–[a]	–[a]
Moisture	1970	−0.18	0.97	−0.44
	1971	–[a]	1.42	–[a]
	Combined	–[a]	1.20	–[a]
Percent oil	1970	–[a]	1.04	–[a]
	1971	–[a]	1.08	–[a]
	Combined	–[a]	1.02	–[a]
Kernel weight	1970	0.28	0.69	−0.51
	1971	–[a]	1.32	–[a]
	Combined	0.26	0.93	−0.14

[a] Negative component in denominator

The results perhaps shed more light on the problems involved in attempting to estimate the quadratic components than on the genetics of the population in which they were obtained. The standard error of Q_{xx} seems in many cases to be large relative to the apparent magnitude of the component being estimated, suggesting that the set of relatives used does not have nice properties for estimating Q_{xx}. This may be owing to multicollinearity in the expectations of the mean squares and products. Cornelius and Dudley (1976) suggested that the correlation of the coeffi-

cients of σ_A^2 and σ_D^2 in the C-model for the covariances among relatives evaluated in the experiment was sufficiently high to result in difficulty in obtaining a realistic separation of the variance attributable to σ_A^2 and to σ_D^2. Such a situation easily triggers negative estimates of one or the other component because it is quite possible for an overallocation of variance to one component to be easily rectified (in terms of adequacy of fit of the model to the data) by negative allocation of variance to the other. A negative estimate of one or the other of these two components (usually of σ_D^2, but of σ_A^2 in one case, namely, yield in 1971) was obtained in one of the two within-year analyses for every character. A similar multicollinearity problem (if this is indeed the cause) exists regardless of whether the C-, D-, or Q-model is used, since all of these models have the same set of coefficients for σ_D^2 and there always exists a linear combination of the other coefficients in the D- or Q-model which is equal to the coefficient of σ_A^2 in the C-model.

In the Q-model the correlation of the coefficients of Q_{xx} (alias D_2) and σ_D^2 is the same as the correlation of the coefficients of σ_A^2 and σ_D^2 in the C-model. Thus, the problem of separating σ_D^2 from σ_A^2 (which is a function of Q_{xx}) appears to be, more fundamentally, a problem of separating σ_D^2 and Q_{xx}. A high correlation of coefficients of σ_D^2 and Q_{xx} in the Q-model manifests itself in the D-model as a high degree of multicollinearity involving the coefficients of σ_D^2, σ_A^2, D_1 and D_2.

Statistical joint significance of the hypotheses $Q_{xy} = Q_{xx} = 0$ and $\mu_\infty^2 = \sigma_D^2$ without significance of either hypothesis when tested alone can also be a consequence of multicollinearity. It is analogous to a multiple regression problem where one or the other of two highly correlated regressor variables is needed in the model, but neither explains a significant amount of variation when added to a model which already contains the other.

The lesson to be learned is that anyone who wishes to devise a study to estimate these components would be well advised to determine whether the set of relatives to be used provide covariances with desirable properties for estimation of the components or at least of the desired linear combinations of components. Besides predicting response of highly inbred progenies resulting from selection of progenies in earlier generations of inbreeding, which could be done quite well from their study, Cornelius and Dudley (1976) attempted to use the estimates of the components to predict improvement in response of outbred performance resulting from selection of inbred performance and vice versa. Cockerham and Matzinger (1985), using an extension of the model to include additive × additive epistasis, made some theoretical predictions of response, in terms of outbred performance to selection based on performance of selfed progenies. They also briefly discussed whether estimated response to selection should be done with the needed variances and covariances directly estimated or indirectly estimated through the use of estimated components. It is to be remembered that estimation of a covariance as a linear combination of estimated components, where the combination is considerably different from any of the experimentally obtained covariances used in estimating the components, presents a situation very much analogous to extrapolation of a regression model to a point outside the range of data to which the model was fit (or, more precisely, outside the hyperellipse which contains the experimental data points). Such projections can be notoriously inaccurate even when the model is correct. Predictions of outbred performance based on inbred performance or vice versa involve linear combinations of σ_A^2 and C (or equivalently, σ_A^2 and D_1 in the D-model, or Q_{yy}, Q_{xy} and Q_{xx} in the Q-model) for the numerator and, for the latter case, σ_A^2, as the only genetic component involved in the denominator of the relevant heritability (b_{xy} of Cockerham and Matzinger 1985). These are going to require a precise estimate of σ_A^2. The experimental results of Cornelius and Dudley (1976), considered again here, suggest that satisfactory separation of σ_A^2 and σ_D^2 (or of Q_{xx} and σ_D^2) is the greatest difficulty to be surmounted in estimating the components from data on inbred relatives. From selfed progenies alone these components are not estimable at all (Cornelius and VanSanford 1988). In the Cornelius and Dudley (1976) study, inclusion of families of non-inbred full sibs which were collateral relatives with the sib-mated and selfed inbred progenies did not seem to satisfactorily alleviate the problem. Inclusion of unilateral relatives would give covariances which depend only on σ_A^2, C and σ_∞^2 (or σ_A^2, D_1, and D_2 in the D-model, or Q_{yy}, Q_{xy} and Q_{xx} in the Q-model) and would help to separate these components from σ_D^2 and μ_∞^2 (alias $\ddot{\mathrm{H}}$). Most importantly, if an estimate of σ_A^2 free of σ_D^2 is needed, the inclusion of noninbred relatives whose covariance depends only on σ_A^2 would be highly desirable.

Acknowledgements. The collection of the data used in this paper was suported by funds of the Illinois Agricultural Experiment Station while the author was a graduate student at the University of Illinois at Urbana-Champaign.

Appendix

The quadratic forms Q_{yyi}, Q_{xyi} and Q_{xxi}, derived by Cornelius and VanSanford (1988) expressed in matrix notation are $\boldsymbol{y'R'P_1Ry}$, $\boldsymbol{y'R'P_1Tx}$ and $\boldsymbol{x'T'P_1Tx}$, respectively, where $\boldsymbol{P_1} = \mathrm{Diag}(p_j)$. In the j^{th} row of matrix $\boldsymbol{R}$, the element which multiplies y_{lm} in the product $\boldsymbol{Ry}$ is $\delta_{jm}p_l - \delta_{jl}p_m$ where $\delta_{jm} = 1$ if $j = m$ and is zero otherwise (δ_{jl} similarly defined). Also, the element in the j^{th} row of $\boldsymbol{T}$ which multiplies x_{lm} in the product $\boldsymbol{Tx}$ is $T_{j(lm)} = \delta_{jl}p_m + \delta_{jm}p_l - 2p_lp_m$. Vector $\boldsymbol{Ry}$ is a vector of additive effects of alleles in a completely homozygous population and vector $\boldsymbol{Tx}$ is an expression for the dominance deviations (at panmixia) of homozygotes, d_{jj}, expressed as a deviation from a mean defined as $h_i = \sum_j p_j d_{jj}$. Consequently, Q_{yy} is equivalent to

σ^2_∞ in the C-model and Q_{xx} is an alternative expression for D_2 in the D-model, but Q_{xy} is unique to the Q-model. Q_{xy} is, in a sense, the covariance of dominance deviations of homozygotes at panmixia with the additive effects of the same alleles in a completely homozygous population.

Another interpretation of Q_{yy}, Q_{xy} and Q_{xx} is obtained by using the relationships $\tilde{y}_{ml} = -y_{lm}$ and $x_{ml} = x_{lm}$ and then noting that the k^{th} element of $\boldsymbol{R}\boldsymbol{y}$ is $\tilde{y}_{\cdot k} = \sum_{j \neq k} p_j y_{jk}$, and the k^{th} element of $\boldsymbol{T}\boldsymbol{x}$ is $\tilde{x}_{\cdot k} - \sum_k p_k \tilde{x}_{\cdot k}$. Consequently, $\tilde{x}_{\cdot k}$ and d_{kk} differ by an additive constant and Q_{yy} and Q_{xx} are shown to be the variances of the $\tilde{y}_{\cdot k}$ and $\tilde{x}_{\cdot k}$, respectively, and Q_{xy} is their covariance.

$Q_{xxi} = 0$ if and only if $\boldsymbol{T}\boldsymbol{x} = \boldsymbol{0}$. But $\boldsymbol{T}\boldsymbol{x} = \boldsymbol{0}$ if and only if the $\tilde{x}_{\cdot j}$ are all equal. Thus, to determine the conditions under which $Q_{xxi} = 0$, it is sufficient to consider any set of $n-1$ linearly independent contrasts among the $\tilde{x}_{\cdot j}$. Let us consider the contrasts $e_j = \tilde{x}_{\cdot j} - \tilde{x}_{\cdot n}$, $j = 1, \ldots, n-1$. Collecting the e_j into a vector, we can write $\boldsymbol{e} = \boldsymbol{S}\boldsymbol{x}$ where the element of the $(n-1) \times n(n-1)/2$ matrix $\boldsymbol{S}$ which multiplies x_{lm} $(l < m)$ in the product $\boldsymbol{S}\boldsymbol{x}$ is $p_m(\delta_{jl} - \delta_{nl}) + p_l(\delta_{jm} - \delta_{nm})$. If the elements x_{lm} in $\boldsymbol{x}$ are sorted by m within l, then

$$\boldsymbol{x} = [x_{12} \quad x_{13} \cdots x_{1(n-1)} \quad x_{1n} \quad x_{23} \cdots x_{2(n-1)} \quad x_{2n} \quad x_{34} \cdots \quad x_{3n} \quad \cdots \quad x_{(n-1)n}]',$$

$$\boldsymbol{S} = \begin{bmatrix} p_2 & p_3 & \cdots & p_{n-1} & (p_n - p_1) & 0 & \cdots & 0 & -p_2 & 0 & \cdots & -p_3 & \cdots & -p_{n-1} \\ p_1 & 0 & \cdots & 0 & -p_1 & p_3 & \cdots & p_{n-1} & (p_n - p_2) & 0 & \cdots & -p_3 & \cdots & -p_{n-1} \\ 0 & p_1 & \cdots & 0 & -p_1 & p_2 & \cdots & 0 & -p_2 & p_4 & \cdots & (p_n - p_3) & \cdots & -p_{n-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & p_1 & -p_1 & 0 & \cdots & p_2 & -p_2 & 0 & \cdots & -p_3 & \cdots & (p_n - p_{n-1}) \end{bmatrix}.$$

Thus, $Q_{xxi} = 0$ for any $\boldsymbol{x}$ which satisfies the set of homogeneous linear equations $\boldsymbol{S}\boldsymbol{x} = \boldsymbol{0}$. Obviously, $\boldsymbol{x} = \boldsymbol{0}$ (the "trivial" solution) is always a solution, i.e., $Q_{xxi} = 0$ if there is no dominance at the i^{th} locus.

If $n = 2$, then $\boldsymbol{S}\boldsymbol{x} = (p_2 - p_1)x_{12}$ so that no dominance $(x_{12} = 0)$ or equal gene frequencies $(p_2 - p_1 = 0)$ are the only conditions which will give $Q_{xxi} = 0$. However, if $n > 2$, $\boldsymbol{S}$ always has more columns than rows and $\boldsymbol{S}\boldsymbol{x} = \boldsymbol{0}$ always has a nontrivial solution for $\boldsymbol{x}$.

If we express $\boldsymbol{S}\boldsymbol{x}$ as $[\boldsymbol{S}_1 \boldsymbol{S}_2]\boldsymbol{x} = \boldsymbol{S}_1\boldsymbol{x}_1 + \boldsymbol{S}_2\boldsymbol{x}_2$ where $\boldsymbol{x}_1 = [x_{12}, x_{13}, \ldots, x_{1n}]'$ and $\boldsymbol{x}_2 = [x_{23}, x_{24}, \ldots, x_{(n-1)n}]'$, then $\boldsymbol{S}\boldsymbol{x} = \boldsymbol{0}$ implies that $\boldsymbol{x}_1 = \boldsymbol{S}_1^{-1}\boldsymbol{S}_2\boldsymbol{x}_2$ provided that $p_1 \neq 0.5$ ($\boldsymbol{S}_1$ is singular if $p_1 = 0.5$).

$$\boldsymbol{S}_1^{-1} \doteq \frac{1}{p_1(1-2p_1)} \begin{bmatrix} p_1 & (1-2p_1-p_2) & -p_3 & \cdots & -p_{n-2} & -p_{n-1} \\ p_1 & -p_2 & (1-2p_1-p_3) & \cdots & -p_{n-2} & -p_{n-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ p_1 & -p_2 & -p_3 & \cdots & (1-2p_1-p_{n-2}) & -p_{n-1} \\ p_1 & -p_2 & -p_3 & \cdots & -p_{n-2} & (1-2p_1-p_{n-1}) \\ p_1 & -p_2 & -p_3 & \cdots & -p_{n-2} & -p_{n-1} \end{bmatrix}$$

and, with algebraic manipulation,

$$x_{1k} = [2\sum_{l<m\,(l,m\neq 1)} p_l p_m x_{lm} - (1-2p_1)\sum_{l\neq 1,k} p_l x_{kl}]/ p_1(1-2p_1),$$

provided $p_1 \neq 0.5$. Equation (1) in the text follows from the fact that the indexing of alleles is arbitrary.

Now suppose that $p_1 = 0.5$. Then $\boldsymbol{S}_1\boldsymbol{x}_1 + \boldsymbol{S}_2\boldsymbol{x}_2 = \boldsymbol{0}$ has a solution for $\boldsymbol{x}_1$ in terms of $\boldsymbol{x}_1$ if and only if rank $[\boldsymbol{S}_1, \boldsymbol{S}_2\boldsymbol{x}_2]$ = rank $(\boldsymbol{S}_1) = n-2$. If $p_1 = 0.5$ then the rows of $\boldsymbol{S}_1$ are such that $\boldsymbol{S}_{1(1)} = 2\sum_{j=2}^{n-1} p_j \boldsymbol{S}_{1(j)}$ where $\boldsymbol{S}_{1(j)}$ is the j^{th} row of $\boldsymbol{S}_1$. The elements of the vector $\boldsymbol{S}_2\boldsymbol{x}_2$ must satisfy the same constraint. The result is that we must have $\sum_{l<m\,(l,m\neq 1)} p_l p_m x_{lm} = 0$. Given this constraint, a solution is then $x_{1k} = c - 2\sum_{l\neq 1,k} p_l x_{kl}$, $k = 2, \ldots, n$, where c is an arbitrary constant. Again, since the indexing of alleles is arbitrary, equations (2) and (3) follow.

Consider again the equation $\boldsymbol{S}_1\boldsymbol{x}_1 + \boldsymbol{S}_2\boldsymbol{x}_2 = \boldsymbol{0}$ and suppose that $\boldsymbol{x}_2 = \boldsymbol{0}$. Then the equation $\boldsymbol{S}\boldsymbol{x} = \boldsymbol{0}$ reduces to $\boldsymbol{S}_1\boldsymbol{x}_1 = \boldsymbol{0}$. A nontrivial solution for $\boldsymbol{x}_1$ exists if and only if $\boldsymbol{S}_1$ is singular. But, since $p_1 \neq 0$ (else the number of alleles is $n-1$ rather than n), $\boldsymbol{S}_1$ is singular only if $p_1 = 0.5$. Thus, if the only heterozygotes which show dominance are those which involve the allele I_1, then Q_{xxi} can be zero only if $p_1 = 0.5$.

With three alleles, equation (1) gives

$$\begin{aligned} x_{12} &= [2p_2p_3x_{23} - (1-2p_1)p_3x_{23}]/p_1(1-2p_1) \\ &= p_3(2p_2 - 1 + 2p_1)x_{23}/p_1(1-2p_1) \\ &= p_3(1-2p_3)x_{23}/p_1(1-2p_1) \end{aligned}$$

and, similarly, $x_{13} = p_2(1-2p_2)x_{23}/p_1(1-2p_1)$, with x_{23} arbitrary, as a solution for $\boldsymbol{x}$ which gives $Q_{xxi} = 0$. In general, we see that $x_{jk}/x_{jl} = p_l(1-2p_l)/p_k(1-2p_k)$ for j, k, l any permutation of the integers 1, 2, 3. This result establishes (5).

The equations $\boldsymbol{S}_1\boldsymbol{x}_1 + \boldsymbol{S}_2\boldsymbol{x}_2 = \boldsymbol{0}$ in the case of four alleles permits a unique solution for $\boldsymbol{x}_2$ in terms of $\boldsymbol{x}_1$ as $\boldsymbol{x}_2 = -\boldsymbol{S}_2^{-1}\boldsymbol{S}_1\boldsymbol{x}_1$, i.e.,

$$\begin{bmatrix} x_{23} \\ x_{24} \\ x_{34} \end{bmatrix} = -\begin{bmatrix} 0 & -p_2 & -p_3 \\ p_3 & (p_4 - p_2) & p_3 \\ p_2 & -p_2 & (p_4 - p_3) \end{bmatrix}^{-1} \begin{bmatrix} p_2 & p_3 & (p_4 - p_1) \\ p_1 & 0 & -p_1 \\ 0 & p_1 & -p_1 \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{13} \\ x_{14} \end{bmatrix}.$$

The solution is $x_{lm} = [2p_l + 2p_m - 1)\tilde{x}_{\cdot 1} + 2p_1p_kx_{1k}]/2p_lp_m$ for k, l, m running through all permutations of the integers 2, 3, 4 and x_{1k}, $k = 2, 3, 4$ are arbitrary. Since the indexing of alleles is arbitrary and the $\tilde{x}_{\cdot j}$ are all equal, equation (6) follows.

References

Cockerham CC (1983) Covariances of relatives from self-fertilization. Crop Sci 23:1177–1180

Cockerham CC, Matzinger DF (1985) Selection response based on selfed progenies. Crop Sci 25:483–488

Comstock RE, Robinson HF (1948) The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics 14:254–266

Cornelius PL, Byars J (1977) Computing iterative weighted least squares estimates of genetic variance components. Biometrics 33:375–382

Cornelius PL, Dudley JW (1974) Effects of inbreeding by selfing and full-sib mating in a maize population. Crop Sci 14:815–819

Cornelius PL, Dudley JW (1975) Theory of inbreeding and covariances between relatives under full-sib mating in diploids. Biometrics 31:169–187

Cornelius PL, Dudley JW (1976) Genetic variance components and predicted response to selection under selfing and full-sib mating in a maize population. Crop Sci 16:333–339

Cornelius PL, VanSanford DA (1988) Quadratic components of covariance of inbred relatives and their estimation in naturally self-pollinated species. Crop Sci 28:1–7

Kempthorne O (1969) An introduction to genetic statistics. Iowa State University Press, Ames. (Reprinting, with corrections, of a 1957 edition published by Wiley and Sons, New York)

Weir BS, Cockerham CC (1977) Two-locus theory in quantitative genetics. In: Pollak E, Kempthorne O, Bailey TB Jr (eds) Proc Int Conf Quanti Genet. Ames, Iowa. Iowa State University Press, Ames, pp 247–249

Wright AJ, Cockerham CC (1986) Covariances of relatives and selection response in generations of selfing from an outcrossed base population. Theoret Appl Genet 72:689–694