Theor Appl Genet (1987) 73:815–820

Ancestral covariance and the Bulmer effect

G.M. Tallis

Department of Statistics, University of Adelaide, Adelaide, SA 5001, Australia

Received March 19, 1986; Accepted November 12, 1986
Communicated by H.F. Linskens

Summary. An n-generation pedigree is set up and selection is carried out generation by generation. The influence of this procedure on the covariances in subsequent generations is assessed and, ultimately, Bulmer's general recursion formula for the reduction in genetic variance due to selection is obtained. The results are extended to assess the effect of selection of one or more characters on the genetic covariance matrix of a number of characters. The concept of ancestral regression is also used to provide a different insight into the selection process and to justify some models used in the analysis of assortative mating.

Key words: Bulmer effect – Ancestral covariance structure – Selection – Assortive mating

Introduction

The effect of artificial selection on the additive genetic variance of a character influenced by many loci was examined by Bulmer (1971). Bulmer showed that, in general, although substantial changes in genetic variance and heritability can be anticipated during the first two or three generations of selection, the effect is transitory and the original levels of variation can be anticipated to return several generations after selection pressure has been relaxed.

It is the purpose of this paper to re-examine the problem by a direct statistical argument using the full ancestral covariance structure.

Selection for a single character

The first task is to set up a suitable notation which leads to easy generalisations. This notation is then illustrated in the recapitulation of Bulmer's derivation of the disequilibrium component of genetic variance due to selection. Subsequently, ancestral pedigrees and covariances are introduced, and the effect of selection on phenotypic and genetic variances is analysed from this viewpoint. Panmixia is assumed throughout.

Let P_{ij} be the j^{th} individual in the pedigree at generation i from the start of selection. Thus, considering one generation of selection only, P_{01}, P_{02} refer to two parents who are subjected to selection and P_{11} to one of their offspring.

When there are two generations of selection, pertinent individuals in a typical pedigree are the four grandparents P_{01}, P_{02}, P_{03}, P_{04}; the two parents P_{11}, P_{12} and finally the offspring P_{21}.

Let p_{ij} and g_{ij} be the phenotype and additive genotype of P_{ij} for a single character where $E[p_{ij}] = E[g_{ij}] = 0$, $V[p_{ij}] = P$ and $V[g_{ij}] = C[p_{ij}, g_{ij}] = G$ in the absence of selection. Then, under the standard normality assumptions (see Bulmer (1980)) an appropriate model for g_{11}, which will be abbreviated to g_1, is

$$g_1 = b_1 p_{01} + b_1 p_{02} + \varepsilon_1 = b_1 p_0 + \varepsilon_1 \tag{1}$$

where $p_0 = p_{01} + p_{02}$, $b_1 = c[g_1, p_0]/V[p_0] = G/2P = h^2/2$ and $V[\varepsilon_1] = G(1 - h^2/2)$.

Suppose selection in the maternal and paternal populations changes P to $(1 + k_0)P$ and $(1 + k_1)P$ respectively, then the induced change in $V[g_1]$ as a result of selection is

$$\Delta G = \frac{h^4}{4}((1 + k_0)P + (1 + k_1)P - 2P)$$

$$= \frac{G h^2}{2}\bar{k}, \bar{k} = (k_0 + k_1)/2 \tag{2}$$

and $V[g_1] = G_1 = G_0 + G_0 h_0^2 \bar{k}/2$, where G_0 and h_0^2 are the genetic variance and heritability prior to selection. For the example considered by Bulmer, $P_0 = 100$, $G_0 = 50$, $h_0^2 = 0.5$ and $\bar{k} = -0.7818$ corresponding to the selection of the top 20% in both the male and female populations.

Suppose now that selection is carried out in the grand-parental generation only. For this situation the appropriate model is

$$g_2 = b_2 p_0 + \varepsilon_2 \tag{3}$$

where $p_0 = p_{01} + p_{02} + p_{03} + p_{04}$, $b_2 = c[g_2, p_0]/V[p_0] = h^2/4$ and $V[\varepsilon_2] = G[1 - h^2/4]$. Hence

$$\Delta G = \frac{h^4}{16}(2P(1+k_0) + 2P(1+k_1) - 4P) = G h^2 \bar{k}/4 \tag{4}$$

which is one half the ΔG of (2).

Bulmer argues as follows. Since Crow and Kimura (1970) have shown that, under panmixia, the additive genetic variance of a polygenic character is practically unaffected by selection, the effect noted above must be transitory and due primarily to the covariance between loci. Thus at the n^{th} generation after the start of selection

$$G_n = G_0 + D_n$$

where D_n is the transitory or disequilibrium component.

Under the assumption that this disequilibrium component is additive across generations

$$D_n = \frac{\bar{k}}{2} h_{n-1}^2 G_{n-1} + \frac{\bar{k}}{4} h_{n-2}^2 G_{n-2} + \ldots + \frac{\bar{k}}{2^n} h_0^2 G_0 \tag{5}$$

with $h_i^2 = G_i/P_i$, $P_i = G_i + E_0$, E_0 being the environmental component which is assumed invariant. After a number of generations an equilibrium situation should be reached, approximately, and

$$\hat{D} = \bar{k} \hat{h}^2 \hat{G}, \quad \hat{G} = G_0 + \hat{D}, \quad \hat{P} = P_0 + \hat{D}, \quad \hat{h}^2 = \hat{G}/\hat{P}. \tag{6}$$

The quadratic equation in $\hat{D}$ determined by (6) can be solved to give

$$\hat{D} = P_0\{2\bar{k} h_0^2 - 1 + [1 - 4\bar{k} h_0^2(1 - h_0^2)]^{1/2}\}/2(1 - \bar{k}) \tag{7}$$

which is expression (19) of Bulmer (1971). The reader is referred to this paper for further details.

This is a very interesting solution to a long standing problem. The results, however, seem to depend on a number of steps and assumptions and it is desirable to compare (5) with formulae obtained by alternative statistical argument.

To achieve this, a well known general result concerning the selection of random variables is needed. Let $X_1, X_2, \ldots, X_m$ be m random variables with zero means and variance matrix V, and suppose that selection is carried out on $X_1, X_2, \ldots, X_p$, $p < m$. If in partitioned form

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$$

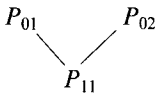
and, under the direct selection of $X_1, X_2, \ldots, X_p$, V_{11} becomes U_{11}, then V_{22} is changed indirectly to

$$U_{22} = V_{22} - V_{21}(V_{11}^{-1} - V_{11}^{-1} U_{11} V_{11}^{-1}) V_{12}. \tag{8}$$

Equation (8) holds under fairly general conditions of distribution of the variables X_i, see Johnson and Kotz (1972, p 70). It certainly applies under the usual assumptions of normality of distribution.

Parent-offspring

Pedigree 1



The vector of phenotypes is (p_{01}, p_{02}, p_{11}) with covariance matrix

$$\begin{bmatrix} P & 0 & G/2 \\ 0 & P & G/2 \\ G/2 & G/2 & P \end{bmatrix}.$$

In this example

$$V_{11} = \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix}, \quad V_{12} = \begin{bmatrix} G/2 \\ G/2 \end{bmatrix}, \quad V_{22} = P$$

and, using the same notation as earlier

$$U_{11} = \begin{bmatrix} (1+k_0)P & 0 \\ 0 & (1+k_1)P \end{bmatrix}$$

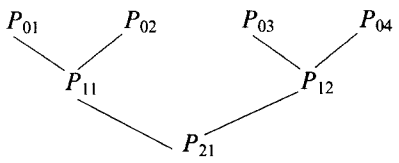
and P in generation 1 changes to

$$U_{22} = P_1 = P_0 + P_0 \frac{h_0^4}{2} \bar{k} \tag{9}$$

where P and G have been labelled P_0 and G_0, and $h_0^2 = G_0/P_0$. This is the same as the result obtained from equation (1), since from (9) $G_1 = G_0 + G_0 \frac{h_0^2}{2} \bar{k}$ assuming $P_1 = G_1 + E_0$.

Grandparent-parent-offspring

Pedigree 2



The vector of phenotypes is $(p_{01}, p_{02}, p_{03}, p_{04}, p_{11}, p_{12}, p_{21})$ with covariance matrix

$$\left[\begin{array}{cccc|ccc} P & 0 & 0 & 0 & G/2 & 0 & G/4 \\ 0 & P & 0 & 0 & G/2 & 0 & G/4 \\ 0 & 0 & P & 0 & 0 & G/2 & G/4 \\ 0 & 0 & 0 & P & 0 & G/2 & G/4 \\ \hline G/2 & G/2 & 0 & 0 & P & 0 & G/2 \\ 0 & 0 & G/2 & G/2 & 0 & P & G/2 \\ G/4 & G/4 & G/4 & G/4 & G/2 & G/2 & P \end{array}\right]$$

This time V_{11} is the (4×4) top principal minor, V_{12} is the neighbouring (4×3) sub-matrix and V_{22} is the bottom (3×3) principal minor. The matrix U_{11} can be written $U_{11} = P(I_4 + K)$ where I_4 is the 4-dimensional identity matrix and $K = \text{diag}(k_0, k_1, k_0, k_1)$. Now

$$V_{11}^{-1} - V_{11}^{-1} U_{11} V_{11}^{-1} = -P^{-1} K$$

and

$$P^{-1} V_{21} K V_{12} = \bar{k} P^{-1} G^2 \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

Thus $\bar{k} h^4 P/2$ is added to the phenotypic and the genetic variances of the parental generation and $\bar{k} h^4 P/4$ is added to the variances of the offspring generation.

Let $P(i, j)$, $G(i, j)$ be the phenotypic and genetic variances in generation j after i generations of selection, $i \leqq j$, then

$$V_{22} = \begin{bmatrix} P(1,1) & 0 & G(1,2)/2 \\ 0 & P(1,1) & G(1,2)/2 \\ G(1,2)/2 & G(1,2)/2 & P(1,2) \end{bmatrix}.$$

Apply the analysis of Pedigree 1 again, and it is clear that

$$\begin{aligned} P(2,2) &= P(1,2) + P(1,1) h_1^4 \bar{k}/2, \quad h_1^2 = G(1,1)/P(1,1) \\ &= P + P h^4 \bar{k}/4 + P(1,1) h_1^4 \bar{k}/2. \end{aligned} \tag{10}$$

This generalises in an obvious way and writing P as P_0; $P(i, i)$, $G(i, i)$ as P_i, G_i and putting $h_i^2 = G_i/P_i$

$$P_n = P_0 + \bar{k} \sum_{i=0}^{n-1} P_i h_i^4 / 2^{n-i} \tag{11}$$

and clearly $P_n - P_0$ is the D_n of Bulmer.

Thus, equation (5) can be derived from first principles by elementary statistical argument. It is clear that selection need not be the same for males and females, nor is it essential that selection be constant across generations. In fact, to accomodate changing selection pressure, separate $\bar{k}$'s for each generation can be calculated, $\bar{k}_i$ for generation i, and (11) becomes

$$P_n = P_0 + \sum_{i=0}^{n-1} \bar{k}_i P_i h_i^4 / 2^{n-i}. \tag{12}$$

Each selection process alters the variance in existing and later generations, and the alterations are cumulative.

The validity of the repeated application of (8) is open to question. The first reduction can be justified; however subsequent use of (8) requires, if necessary, that at each stage the existing distribution be approximated by one with the same moment characteristics and for which (8) is valid.

The errors incurred by assuming, for example, normality at each stage can be investigated in simple cases such as Pedigree 1. It certainly seems that sufficiently precise results are obtained by assuming normality stage by stage to capture the essential numerical characteristics of the selection process. Refined comparisons may require more exacting calculations, but these are obviously difficult to do for more than two generations.

Ancestral regression

Consider now a two generation model appropriate for selection in parental and grand-parental generations;

$$g_2 = b_1 p_1 + b_2 p_0 + \varepsilon_2. \tag{13}$$

Note that $p_1 = p_{11} + p_{12}$, $p_0 = p_{01} + p_{02} + p_{03} + p_{04}$ and that b_1, b_2 and ε_2 are different to the corresponding expressions of (1) and (3). Moreover, although the b_i of (13) are derived by Bulmer, they are re-derived here as a preparation for extensions. We note that

$$C[g_2, p_1] = C[g_2, p_0] = G, \quad V[p_1] = 2P, \quad V[p_0] = 4P,$$
$$C[p_1, p_0] = 2G$$

and the appropriate equations for determining b_1 and b_2 are

$$\begin{bmatrix} 2P & 2G \\ 2G & 4P \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} G \\ G \end{bmatrix}$$

and

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \frac{h^2}{2} \begin{bmatrix} 1 & h^2 \\ h^2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

whence

$$b_1 = \tfrac{1}{2} h^2 (2 - h^2)/(2 - h^4), \quad b_2 = \tfrac{1}{2} h^2 (1 - h^2)/(2 - h^4).$$

These are interesting formulae showing that when $h^2 = 0$ there is no regression on ancestors, and when $h^2 = 1$ only the parents provide predictive information.

Moreover, the amount of $V[g_2]$ contributed by regression is

$$V_2 = G \frac{h^2}{2} [1 \;\; 1] \begin{bmatrix} 1 & h^2 \\ h^2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = G \frac{h^2 (3 - 2h^2)}{2(2 - h^4)}. \tag{14}$$

For instance, in Bulmer's numerical example, $G = 50$, $h^2 = 0.5$ and $V_2 = 14.29$.

It is now necessary, as well as instructive, to establish the degree to which parents, grandparents,

Table 1. Values of $\gamma_n(h^2)$ for various n and h^2

h^2 / n	1	2	3	4	5	6	∞
0.1	0.05000	0.07035	0.07869	0.08212	0.08354	0.08412	0.08452
0.2	0.10000	0.13265	0.14361	0.14731	0.14857	0.14900	0.14922
0.3	0.15000	0.18848	0.19895	0.20184	0.20264	0.20286	0.20295
0.4	0.20000	0.23913	0.24760	0.24947	0.24988	0.24997	0.25000
0.5	0.25000	0.28571	0.29167	0.29268	0.29286	0.29289	0.29289
0.6	0.30000	0.32927	0.33283	0.33327	0.33333	0.33333	0.33333
0.7	0.35000	0.37086	0.37254	0.37268	0.37269	0.37269	0.37269
0.8	0.40000	0.41176	0.41228	0.41230	0.41230	0.41230	0.41230
0.9	0.45000	0.45378	0.45384	0.45384	0.45384	0.45384	0.45384
1	0.50000	0.50000	0.50000	0.50000	0.50000	0.50000	0.50000

great-grandparents and so on can contribute to V[g]. This will be labelled V_∞ and it will be found by first calculating V_n, and then letting n tend to infinity.

The appropriate model is

$$g_n = b_1 p_{n-1} + b_2 p_{n-2} + \ldots + b_n p_0 + \varepsilon_n \qquad (15)$$

where $p_i = \sum_{j=1}^{2^{n-i}} p_{ij}$. Writing

$$\Sigma_n = \begin{bmatrix} 1 & h^2 & \ldots h^2 \\ h^2 & 2 & \ldots 2h^2 \\ \vdots & \vdots & \vdots \\ h^2 & 2h^2 \ldots & 2^{n-1} \end{bmatrix}$$

the generalisation of V_2 is

$$V_n = G\frac{h^2}{2} 1_n' \Sigma_n^{-1} 1_n = G\,\gamma_n(h^2), \quad \text{say.} \qquad (16)$$

where 1_n is a vector of n 1's.

Values of $\gamma_n(h^2)$ are given in Table 1 for various values of n and h^2. The limiting values of γ are found in the last column, and it is clear that convergence to these limits is rapid, especially for the larger values of h^2. Convergence is guaranteed since, from genetic considerations, $\gamma_n(h^2) \leqq \frac{1}{2}$, and from regression theory, $\gamma_n(h^2)$ is monotone non-decreasing in n.

For the numerical example cited, $h^2 = 0.5$, and $V_\infty \cong 14.64$. This is only a minor increase over V_2.

Now, $G_n = V_n + V[\varepsilon_n]$, and if selection is practiced, V_n will be changed to V_{sn}, say. Then, $D_n = V_{sn} - V_n$ and the calculation of V_{sn} will now be outlined for n = 2.

Under assumptions of normality, p_{01}, p_{02} and p_{11} have a trivariate normal distribution with mean vector 0 and covariance matrix

$$P = P\begin{bmatrix} 1 & 0 & h^2/2 \\ 0 & 1 & h^2/2 \\ h^2/2 & h^2/2 & 1 \end{bmatrix}.$$

Using the formulae of Tallis (1961) for the first two moments of a rectangularly truncated normal distribution of three dimensions and a numerical integration program, it is possible to compute the effect of truncation on P, P_s. Then, using the explicit formulae for b_1 and b_2 it is easey to find V_{s2}, and hence $D_2 = V_{s2} - V_2$.

Table 2. Bulmer's numerical example

n	V_n	V_{sn}	$-D_n$	G_n	P_n	h_n^2
1	12.50	2.73	9.77	40.23	90.23	0.446
2	14.29	2.40	11.89	38.11	88.11	0.433
3	14.58	2.20	12.38	37.62	87.62	0.429
4	14.63	2.13	12.50	37.50	87.50	0.429
∞	14.64	2.11	12.53	37.47	87.47	0.428

One point that needs to be made is that the truncation levels must be carefully set in the calculation of P_s. For example, if a fixed proportion α of the population is selected at each stage, the cut-off points for p_{01} and p_{02}, a say, can be set in the usual way for a one dimensional normal distribution. However, the truncation point for p_{21}, b, must be found, by numerical methods, from the trivariate normal integration process and must satisfy

$$P(p_{21} > b \mid p_{01} > a, p_{02} > a) = \alpha \quad \text{or}$$

$$P(p_{21} > b, p_{01} > a, p_{02} > a) = \alpha^3.$$

Further details are omitted.

It turns out that this approach gives results marginally different to those of (11). This is probably because the assumption that the reduction formula (8) is repeatably applicable requires a distribution other than the multinormal. In fact D_1 values for both approaches agree exactly while D_2 calculated by the two differ, by at most, a small fraction of 1% for the cases investigated.

Because of the close agreement of the two methods up to n = 2, and D_n converges rapidly, it is safe to write $V_{sn} = V_n + D_n$ and to calculate D_n by (5) and V_n by (16). In this way Table 2 was compiled for Bulmer's numerical example.

One observation comes from (15). The vector of regression coefficients, b, satisfies the equation

$$\Sigma_n b = \frac{h^2}{2} 1_n, \tag{17}$$

and, when $h^2 = 1$ it is easy to see that the vector $b' = (\frac{1}{2}, 0, \ldots, 0)$ is the unique solution. This implies that if the genotypes of the two parents, $g_{(n-1)1}$ and $g_{(n-1)2}$, are known, the more remote ancestral genotypes are of no predictive value, and the regression equation becomes

$$g_n = \tfrac{1}{2}(g_{(n-1)1} + g_{(n-1)2}) + \varepsilon_n. \tag{18}$$

Model (18) has important applications in assortative mating when it is required to find the effect of n generations of assortation on genetic variance. Under the standard distributional assumptions, in (18) the variance of g_n depends only on the variance of $\frac{1}{2}(g_{(n-1)1} + g_{(n-1)2})$ and $V[\varepsilon_n]$, the latter being independent of assortation. Since (18) holds under panmixia, $V[\varepsilon_n] = G_0/2$, see Bulmer (1980). This leads immediately to the recurrence equation $G_n = G_{n-1}(1 + m_g)/2 + G_0/2$, where m_g is the correlation between parental genotypes due to assortation. The limiting result $\hat{G} = G_0/(1 - m_g)$ is found immediately from this formula, and the whole analysis is justified on the basis of the ancestral regression model.

Multiple characters

Selection on one character

In order to discuss the effect of selection on more than a single character, the notation must be expanded. Let there be q charachters; $K_1, K_2, \ldots, K_q$; then the first problem to be considered is the effect of selection of one character, K_1 say, on the genetic variances and covariances of others.

With respect to K_s, $p_{ij}(s)$ is the phenotype of P_{ij} and $g_n(s)$ is the additive genotype n generations after selection commences. Further, with an unavoidable conflict of notation with the previous section, let

$$C[p_{ij}(s), p_{ij}(t)] = P_{st}, \quad V[g_n(s)] = G_{ss},$$
$$C[g_n(s), p_{ij}(t)] = 2^{-(n-i)} G_{st}$$

and without loss of generality set $s = 1$, $t = 2$. The appropriate model is

$$g_n(2) = c_1 p_{n-1}(1) + \ldots + c_n p_0(1) + \varepsilon_n(1) \tag{19}$$

where obviously $p_i(1) = \sum_{j=1}^{2^{n-1}} p_{ij}(1)$, and so on. The variance due to regression is

$$V_n[2|1] = \frac{G_{12}^2}{2P_{11}} 1_n' \Sigma_n^{-1}(1) 1_n = \beta_{12}^2 V_n(1) \tag{20}$$

where $\beta_{ij} = G_{ij}/G_{ii}$, $\Sigma_n(1)$ and $V_n(1)$ are Σ_n and V_n for K_1.

Using (15) and (19) it is easy to deduce that the covariance due to regression between K_1 and K_2 is

$$C_n[1, 2|1] = \frac{G_{12}}{G_{11}} V_n(1) = \beta_{12} V_n(1). \tag{21}$$

Similarly, considering the effect of selection for K_1 on the covariance of K_2 and K_3

$$C_n[2, 3|1] = \frac{G_{12} G_{13}}{G_{11}^2} V_n(1) = \beta_{12}\beta_{13} V_n(1). \tag{22}$$

From these results it is easy to see that the changes to the genetic variance and covariance structure as a result of selection for K_1 is

$$\Delta_n(G_{11}) = D_n, \quad \Delta_n(G_{12}) = \beta_{12} D_n, \quad \Delta_n(G_{22}) = \beta_{12}^2 D_n,$$
$$\Delta_n(G_{23}) = \beta_{12}\beta_{13} D_n. \tag{23}$$

Since there is no loss of generality by restricting the study to K_1, K_2 and K_3, the effect of selection of a single character on the genetic covariance structure of an arbitrary set of characters is fully specified by D_n and (23).

Selection for more than one character

Suppose there are q characters with phenotypic and genetic covariance matrices P and G respectively. If selection is practiced on these characters, then it is assumed that P is altered to $(I + K)P$, where K is a suitable $(q \times q)$ matrix.

If, for example, truncation selection is practiced on the characters, then the effect of this can be assessed by using appropriate formulae for the truncated normal distribution, Tallis (1961). Thus P would become P_s, say, and $K = (P_s - P)P^{-1}$.

The matrices G and P can now replace their scalar counterparts in the formulation leading to (12). If males and females are selected differently in generation i, matrices K_{0i} and K_{1i} will define this and these two matrices play the role of k_{0i} and k_{1i} in the single character case. It is easy to infer that, if P_n and G_n are the phenotypic and genetic covariance matrices after n generations of selection,

$$G_n = G_0 + D_n, \quad P_n = P_0 + D_n$$

where

$$D_n = \sum_{i=0}^{n-1} 2^{-(n-i)} H_i \bar{K}_i H_i P_i \tag{24}$$

where $\bar{K}_i = 2^{-1}(K_{0i} + K_{1i})$ and $H_i = G_i P_i^{-1}$. If

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} 2^{-(n-1)} \bar{K}_i = \bar{K},$$

equilibrium is defined by

$$\hat{D} = \hat{H}\bar{K}\hat{G}, \quad \hat{G} = G_0 + \hat{D}, \quad P_0 = P_0 + \hat{D} \tag{25}$$

which is best solved by iteration.

The use of ancestral covariance structure to obtain these results is efficient. Moreover, this approach should permit further generalisations to non-panmictic populations, see discussion.

This section concludes with a brief discussion of ancestral regression when there are multiple characters.

Let $p_i' = (p_i(1), \ldots, p_i(q))$, where $p_i(s) = \sum_{j=1}^{2^{n-1}} p_{ij}(s)$,

then from $p = (p_1, p_2, \ldots, p_n)$ it is required to predict additive genotype, $g_n' = (g_n(1), \ldots, g_n(q))$. Hence, under the previously stated assumptions,

$$g_n = B\,p + \varepsilon_n \tag{26}$$

applies for suitable $q \times nq$ matrix B.

Notice that

$$C[g_n, p] = [G, G, \ldots, G]$$

where G is the $q \times q$ variance matrix $V[g_n]$. Moreover, if P is the phenotypic variance matrix for $K_1, \ldots, K_q$,

$$V[p] = 2\begin{bmatrix} P & G & \ldots & G \\ G & 2P & \ldots & 2G \\ \vdots & \vdots & & \vdots \\ G & 2G & \ldots & 2^{n-1}P \end{bmatrix}.$$

It is immediate that

$$B = C[g_n, p]\,[V[p]]^{-1} \tag{27}$$

and that the covariance due to regression is

$$C_n = C[g_n, p]\,[V[p]]^{-1}\,C[g_n, p]'. \tag{28}$$

Under selection V[p] is altered to $V_s[p]$, say, and C_{ns} is defined as for C_n with $V_s[p]$ replacing V[p]. Thus $D_n = C_{ns} - C_n$, although in practice $V_s[p]$ would be difficult to calculate. Fortunately, again, D_n can be determined directly from (24) and C_{ns} can be found indirectly.

Of more importance, when P = G, the unique solution to the equation $BV[p] = C[g_n, p]$ is $[\frac{1}{2}I_q, 0, \ldots, 0]$. This shows that, in this case, the model reduces to

$$g_n = 2^{-1}(g_{(n-1)1} + g_{(n-1)2}) + \varepsilon_n, \quad V[\varepsilon_n] = G/2 \tag{29}$$

where, of course, the g's are q-dimensional vectors of the additive genotypes of the q characters for the parents. This result allows the analysis of assortative mating outlined at the end of the Section concerning ancestral regression to be extended to multiple characters, Tallis (1985).

Discussion

The work of Bulmer (1971) resolved a basic paradox concerning the effects of selection on additive genetic variance. In addition, he showed how to use multivariate normal models to quantify these effects.

The present paper re-examines the problem using ancestral covariances, and Bulmer's fundamental recursion formula for the disequilibrium variance is obtained by a general reduction formula for variables undergoing selection. The results easily extend to cover selection for more than a single character.

Ancestral regressions are also introduced to provide an alternative viewpoint. This approach has practical limitations, but it yields important formulae used in assortative mating and also gives results close to those of Bulmer.

The reduction procedure used to derive the Bulmer formulae is of some generality. It seems that situations more complex than that of panmixia should be negotiable, and work along these lines is proceeding. In particular, efforts will be made to analyse the joint influence of selection and assortation on genetic covariance structure.

These population level analyses are of limited direct relevance to breeders working with small populations of unspecified genetic structure. Nevertheless, numerical calculations using the various formulae and pertinent values for the parameters to provide some insight as to the level of effect of some breeding practices. Even a rough mathematical analysis is better than guesses based on intuition, because the processes are sufficiently complicated as to make excessive demands on even the highest order intuition.

Acknowledgements. The author wishes to thank the referees for constructive and useful comments on an early version of this paper, and to Mr. P. Leppard, for tireles assistance with computing problems.

References

Bulmer MG (1971) The effect of selection on genetic variability. Am Nat 105:201–211

Bulmer MG (1980) The mathematical theory of quantitative genetics. Clarendon Press, Oxford

Crow JF, Kimura M (1970) An introduction to population genetic theory. Burgess, Minneapolis, Minn

Johnson NL, Kotz S (1972) Distributions in statistics: continuous multivariate distributions. Wiley & Sons, New York

Tallis GM (1961) The moment generating function of the truncated multi-normal distribution. J R Stat Soc B 23: 223–229

Tallis GM (1985) Transfer systems and covariance under assortative mating. Theor Appl Genet 70:497–504