Modification of Estimates of Parameters in the Construction of Genetic Selection Indices ('Bending')

J. F. Hayes; W. G. Hill

Biometrics, Vol. 37, No. 3 (Sep., 1981), 483-493.

Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28198109%2937%3A3%3C483%3AMOEOPI%3E2.0.CO%3B2-V

Biometrics is currently published by International Biometric Society.

BIOMETRICS 37, 483–493
September 1981

Modification of Estimates of Parameters in the Construction of Genetic Selection Indices ('Bending')

J. F. Hayes[1] and W. G. Hill
Institute of Animal Genetics, West Mains Road, Edinburgh EH9 3JN, Scotland

SUMMARY

A method, termed 'bending', is proposed for the modification of the estimates of genetic ($\hat{\mathbf{G}}$) and phenotypic ($\hat{\mathbf{P}}$) covariance matrices, that are used in the construction of genetic selection indices for two or more traits. If **P** and **G** are estimated from the between- and within-class covariance matrices, **B** and **W** respectively, in a one-way multivariate analysis of variance, then the method consists of contracting all the eigenvalues of the matrix product $\mathbf{W}^{-1}\mathbf{B}$ towards their mean but with the corresponding eigenvectors unchanged.

The usefulness of the modification procedure is investigated by Monte Carlo simulation and Taylor series approximation methods. In general, the modification procedure improves the achieved response to selection, with the amount of improvement dependent on the (unknown) population parameters and the size of sample for estimation.

In practice, there is some difficulty in selecting the value of the 'bending' factor; some very simple methods of doing this are proposed. If some of the parameter estimates are known to be defective (outside their valid limits), a simple and effective method for the selection of the 'bending' factor is to contract the eigenvalues so that they are all nonnegative.

1. Introduction

The theory of genetic selection indices, as developed by Smith (1936) and Hazel (1943) and subsequently elaborated by Cochran (1951) and Henderson (1963) among others, is based on the assumption that the population parameters such as heritabilities and correlations are known exactly. In practice, however, only estimates of these parameters are available for constructing the index, and such an index is therefore likely to be less efficient than one computed from the parameters themselves. The effects of errors in the parameter estimates and the loss in efficiency in terms of the size of the sample used for estimation have been considered by Williams (1962a, b), Harris (1964), and Sales and Hill (1976a, b).

Williams (1962a) suggested the use of a base index, in which the economic weights are used directly as the index weights, instead of the index computed from the estimated parameters. For two variables he showed that, unless progress from the optimal index is substantially greater than that from the base index, there is a high probability that the estimated index will yield a poorer response than the base index. Williams (1962b) concluded that the base index should be used unless a large amount of data is available for parameter estimation. In a model where only one of two traits was assumed to be of economic importance, but the second trait might be correlated with it, Sales and Hill (1976b) showed that inclusion of the second trait (i.e. use of the estimated rather than the base index) was likely to be worthwhile only when reliable estimates of parameters are

[1] Present address: Animal Science Department, MacDonald Campus of McGill University, Ste Anne de Bellevue, Quebec H9X 1CO, Canada.

Key words: Genetic selection indices; 'Bending'; Errors in parameter estimates.

available. Further, they showed that, if the second trait really contributed nothing useful, the greater the benefit predicted from its inclusion, the greater the real loss in efficiency if it were included.

Instead of using either the index computed from the parameter estimates directly, or simply the base index, it may be possible to modify the parameter estimates that are used, or the index weights themselves. This has some analogies with the procedure of 'ridge regression' (Hoerl and Kennard, 1970) which is sometimes now used in multiple regression analysis. Campbell (1980) has also suggested shrinkage of estimates in discriminant and canonical analysis to improve their stability. Some modification of parameter estimates for selection index construction has always been practised, at least in Monte Carlo simulation studies, when it was obvious that the estimates were out of bounds, for example, heritabilities outside the range 0 to 1, or genetic correlations outside the range -1 to $+1$. In such cases it has been usual practice to set the estimates to the corresponding bound (Harris, 1964; Sales and Hill, 1976b), although it is not clear that this is an optimal procedure. If partial genetic correlations are also considered it is apparent that estimates outside their bounds will frequently be found if there are many variables (Hill and Thompson, 1978); procedures for putting several estimates to their bounds simultaneously are less obvious and not necessarily satisfactory.

In this paper we describe and evaluate a general procedure, termed 'bending', for modifying estimates of parameters in the construction of genetic selection indices. To simplify the presentation we consider solely use of a one-way balanced classification for parameter estimation. Further, we make use of a reparameterization of the selection index suggested by Hayes and Hill (1980) in which the transformed variables have unit phenotypic variance and are uncorrelated genetically or phenotypically. This greatly reduces the parameter space that needs to be analysed, for we can infer from this reduced space to all heritability, genetic and phenotypic correlation combinations.

2. Theory

Let us assume that individual selection is practised on an index $I = \mathbf{b}'\mathbf{x}$, where $\mathbf{b}$ is a vector of index weights, and $\mathbf{x}$ is a vector of observations on p traits, in order to maximize the correlation with overall breeding value $H = \mathbf{a}'\mathbf{g}$, where $\mathbf{a}$ is a vector of economic weights, and $\mathbf{g}$ is a vector of breeding value on the same p traits. If $\mathbf{P} = \text{var}(\mathbf{x})$ and $\mathbf{G} = \text{var}(\mathbf{g})$ are the phenotypic and genetic covariance matrices, respectively, the optimum index is given by

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}\mathbf{a}, \tag{1}$$

and the expected response to selection from its use is

$$R = i(\mathbf{b}'\mathbf{P}\mathbf{b})^{\frac{1}{2}}, \tag{2}$$

where i is the selection intensity.

If only estimates of the parameters $\mathbf{P}$ and $\mathbf{G}$, namely $\hat{\mathbf{P}}$ and $\hat{\mathbf{G}}$, are available then the estimated index weights are usually taken from (1) as

$$\hat{\mathbf{b}} = \hat{\mathbf{P}}^{-1}\hat{\mathbf{G}}\mathbf{a}, \tag{3}$$

and from (2) the predicted progress is

$$R = i(\hat{\mathbf{b}}'\hat{\mathbf{P}}\hat{\mathbf{b}})^{\frac{1}{2}}. \tag{4}$$

The expectation of the response that is actually achieved when $\hat{\mathbf{b}}$ is used subsequently

for making selection decisions in the population is

$$R^a = i\hat{\mathbf{b}}'\mathbf{G}\mathbf{a}(\hat{\mathbf{b}}'\mathbf{P}\hat{\mathbf{b}})^{-\frac{1}{2}}, \tag{5}$$

where the expectations in (2) and (5) are conditional on the index weights that are used. Later we shall consider expectations of quantities, such as R^a, over the distribution of the estimates of the index weights from samples of data. We shall also consider index weights other than the $\hat{\mathbf{b}}$ given by (3); the base index of Williams (1962) is simply $\mathbf{b} = \mathbf{a}$.

In multiple regression analysis the expected sum of squares of the distances of the regression coefficient estimates, $\hat{\boldsymbol{\beta}}$, from their parameter values, $\boldsymbol{\beta}$, has a simple expression: $\mathrm{E}\{(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\,(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\} = \mathrm{tr}(\mathbf{X}'\mathbf{X})^{-1}\sigma^2 = \sigma^2\sum(1/\lambda_i)$, where tr denotes trace, $\mathbf{X}$ is the design matrix, σ^2 the error variance, and λ_i the ith root of $\mathbf{X}'\mathbf{X}$. Many authors, for example Hoerl and Kennard (1970), have noted that if at least one of the roots of $\mathbf{X}'\mathbf{X}$ is small, the quantity $\mathrm{E}\{(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\}$ can be very large, so the technique of ridge regression was suggested to improve the estimates of the regression coefficients. The new regression coefficients are given by $(\mathbf{X}'\mathbf{X}+\mathbf{K})^{-1}\mathbf{X}'\mathbf{y}$ for some diagonal matrix $\mathbf{K}$, and with appropriate choice of $\mathbf{K}$ better predictors can be obtained. In selection indices the corresponding expressions for $\mathrm{E}\{(\hat{\mathbf{b}}-\mathbf{b})'(\hat{\mathbf{b}}-\mathbf{b})\}$ are more complicated and do not readily suggest a technique for improving the estimates. Nevertheless, some other procedures of modifying estimates can be proposed.

Consider the balanced one-way multivariate analysis of variance with half-sib families as follows:

Source	df	MS	E(MS)
Between-sires	$s-1$	$\mathbf{B}$	$\mathbf{P}-\frac{1}{4}\mathbf{G}+\frac{1}{4}n\mathbf{G}$
Within-sires	$s(n-1)$	$\mathbf{W}$	$\mathbf{P}-\frac{1}{4}\mathbf{G}$

From this analysis, $\hat{\mathbf{P}} = \{\mathbf{B}+(n-1)\mathbf{W}\}/n$, $\hat{\mathbf{G}} = 4(\mathbf{B}-\mathbf{W})/n$. In most practical situations the number of degrees of freedom (df) within groups, $s(n-1)$, will be much larger than the df between groups.

It is well-known from multivariate theory that the roots of the sample matrices $\mathbf{B}$ and $\mathbf{W}$ are biased relative to their expectations, a problem discussed further by Hill and Thompson (1978). In particular, the larger roots are biased upwards, the smaller roots downwards, and pairs of equal roots are spread apart. The magnitude of the bias rises as the sample sizes (i.e. degrees of freedom) decrease, the number of variables increases, and the population roots become closer in value to each other. However, the mean of the sample roots, equal to the trace of the corresponding matrix, is unbiased. This suggests that estimates of parameters (and thus the rate of response to selection) might be improved by reduction of the spread of the sample roots, especially when there are rather few degrees of freedom for the between-group (sire) mean square. Again, this could be done in several ways.

For the selection index computations (3) we are concerned with the properties of $\hat{\mathbf{P}}^{-1}\hat{\mathbf{G}}$, and thus with the roots of the determinantal equation $|\hat{\mathbf{P}}^{-1}\hat{\mathbf{G}}-\lambda\mathbf{I}| = 0$ or, equivalently, $|\hat{\mathbf{G}}-\lambda\hat{\mathbf{P}}| = 0$. Alternatively, we can modify the roots of $\mathbf{W}^{-1}\mathbf{B}$, since if v_i is a root of $|\hat{\mathbf{B}}-v\hat{\mathbf{W}}| = 0$, $\lambda_i = 4(v_i-1)/(v_i-1+n)$ is a root of $|\hat{\mathbf{G}}-\lambda\hat{\mathbf{P}}| = 0$. If it is assumed that there are many more degrees of freedom within groups, the mean of the roots of $\mathbf{W}^{-1}\mathbf{B}$ is little biased (just as the usual estimate of intraclass correlation in the analysis of variance is little biased), but the roots are spread excessively about their mean. This suggests that the roots of $\mathbf{W}^{-1}\mathbf{B}$ should be compressed together without altering the average root. If the

modified matrix is defined as $(\mathbf{W}^{-1}\mathbf{B})^*$, it can be computed as follows:

$$(\mathbf{W}^{-1}\mathbf{B})^* = (1-\gamma)\mathbf{W}^{-1}\mathbf{B} + \gamma\bar{v}\mathbf{I}, \qquad 0<\gamma<1, \tag{6}$$

where γ is the bending factor and $\bar{v} = \sum_{i=1}^{p} v_i/p$ is the mean root, with v_i defined above. Alternatively, on the assumption that $\mathbf{W}$, based on many more degrees of freedom, is known accurately relative to $\mathbf{B}$, it is appropriate not to modify $\mathbf{W}$ and thus to take $(\mathbf{W}^{-1}\mathbf{B})^* = \mathbf{W}^{-1}\mathbf{B}^*$, where $\mathbf{B}^*$ is the modified between-group covariance matrix

$$\mathbf{B}^* = (1-\gamma)\mathbf{B} + \gamma\bar{v}\mathbf{W}. \tag{7}$$

The modified genetic and phenotypic covariance matrices, $\hat{\mathbf{G}}^*$ and $\hat{\mathbf{P}}^*$, are then

$$\hat{\mathbf{G}}^* = 4(\mathbf{B}^* - \mathbf{W})/n = 4\{(1-\gamma)\mathbf{B} - (1-\gamma\bar{v})\mathbf{W}\}/n,$$
$$\hat{\mathbf{P}}^* = \{\mathbf{B}^* + (n-1)\mathbf{W}\}/n = \{(1-\gamma)\mathbf{B} + (n-1+\gamma\bar{v})\mathbf{W}\}/n.$$

For $\gamma = 0$, $\hat{\mathbf{P}}^{*-1}\hat{\mathbf{G}}^* = \hat{\mathbf{P}}^{-1}\hat{\mathbf{G}}$; and for $\gamma = 1$, $\hat{\mathbf{P}}^{*-1}\hat{\mathbf{G}}^* = \mathbf{I}$, so from (3) the index weights are $\hat{\mathbf{b}}^* = \mathbf{a}$. This is the base index since the information from the analysis of variance is totally discounted.

It is not claimed that the above is necessarily the best way of modifying the roots, but, as we shall see from the subsequent analysis and Monte Carlo simulation, it seems to be a useful method.

3. Methods

For any positive-definite phenotypic and genetic covariance matrices, $\mathbf{P}$ and $\mathbf{G}$, there exists a transformation $\mathbf{Q}$ such that $\mathbf{QPQ}' = \mathbf{I}$ (the identity matrix) and $\mathbf{QGQ}' = \mathbf{\Lambda}$ is diagonal. The elements of $\mathbf{\Lambda}$ are the roots of $|\mathbf{G} - \lambda\mathbf{P}| = 0$. This transformation is a straightforward application of multivariate theory and is discussed in detail by Hayes and Hill (1980). The existence of this transformation means that we need only consider the sampling and other properties of indices when the phenotypic covariance matrix is the identity matrix and the genetic covariance matrix is diagonal, i.e. unit phenotypic variances and all genetic and phenotypic correlations equal to zero. The only parameters we need to consider are the heritabilities, $\lambda_i = h_i^2$, and economic weights, α_i, of the transformed variables, then any other parameter sets can be analyzed simply by transformation.

Thus Monte Carlo simulation was carried out solely for the case of uncorrelated variables. Mean-square matrices for normally distributed observations were sampled in the manner described by Hill and Thompson (1978). In all simulations, 500 replicates were used and results for $E(R^a)$, for example, were computed as averages over these replicates. To evaluate expected responses $E(R^a)$ analytically, a Taylor series expansion was used, again for uncorrelated variables. The variances and covariances of mean squares are listed by Hayes and Hill (1980), and more details of the calculations for indices after bending of parameter estimates are given in the Appendix.

4. Results

Simulations were conducted for a wide range of parameter values; values of $E(R^a)$ for some samples are given in Table 1. The results include replicates where the sample estimates represented impossible parameter values, i.e. roots of $\hat{\mathbf{P}}^{-1}\hat{\mathbf{G}}$ outside the range 0 to 1, including cases of negative heritability estimates. As discussed by Hayes and Hill (1980), the loss in efficiency expected without any modification of parameter estimates,

Table 1

Mean and standard deviation of R^a/R for different γ values (bending factors), and proportion (Pr %) of 500 replicates in which R^a/R was improved relative to $\gamma = 0$

	γ:	0		.1			.2			.4			.8		
		R^a/R		R^a/R			R^a/R			R^a/R			R^a/R		
a	h^2	Mean	SD	Mean	SD	Pr	Mean	SD	Pr	Mean	SD	Pr	Mean	SD	Pr
							$s = 25$ $n = 8$								
1 1	.1 .2	.465	.648	.490	.641	75	.514	.637	75	.573	.623	72	.662	.631	64
2 1 1	.1 .2 .3	.412	.561	.458	.546	90	.507	.529	90	.616	.477	87	.809	.345	81
1 1 1	.4 .5 .7	.838	.222	.873	.173	94	.903	.129	94	.946	.064	92	.978	.015	86
1 2 1 1	.4 .3 .25 .7	.709	.306	.759	.265	97	.805	.226	97	.881	.148	95	.941	.053	88
							$s = 100$ $n = 16$								
1 1	.1 .25	.973	.047	.978	.034	53	.981	.026	49	.980	.020	40	.950	.014	16
1 3	.1 .25	.983	.028	.986	.024	62	.988	.020	60	.991	.013	55	.989	.054	41
1 1 .5	.3 .6	.993	.010	.994	.008	49	.995	.007	44	.993	.007	32	.979	.005	8
1 1 1	.1 .15 .2	.922	.091	.937	.070	88	.950	.054	85	.968	.033	80	.975	.011	11
1 1 2	.1 .1 .6	.990	.009	.991	.008	57	.991	.008	49	.987	.010	29	.941	.010	1
1 2 2	.3 .5 .6	.984	.015	.987	.012	83	.989	.010	80	.992	.008	75	.990	.003	55
1 2 2 1	.4 .5 .3 .6	.970	.024	.976	.019	89	.980	.017	86	.984	.013	78	.978	.005	53
							$s = 400$ $n = 16$								
1 1	.1 .2	.992	.011	.994	.008	49	.994	.008	42	.990	.009	30	.968	.006	5
1 2	.1 .2	.994	.009	.995	.008	51	.996	.006	47	.995	.006	34	.986	.004	13
1 1	.4 .5	.998	.003	.998	.002	53	.998	.002	49	.998	.002	40	.996	.001	20
1 1 1	.1 .2 .3	.989	.011	.991	.009	56	.990	.009	46	.985	.010	27	.953	.006	2
1 1 1	.4 .5 .6	.996	.004	.996	.004	65	.997	.003	58	.996	.003	46	.991	.002	15
1 1 1 1	.1 .2 .2 .1	.977	.019	.981	.016	73	.983	.013	70	.984	.011	60	.966	.006	22
1 1 1 1	.3 .4 .5 .6	.993	.005	.994	.004	64	.995	.004	57	.993	.004	37	.981	.002	4

$E(1-R^a/R)$, was greater the closer the heritabilities were to zero and to each other, the greater the number of variables included and the smaller the sample size for estimation.

For all parameter sets, it was possible to select a value of the bending factor, γ, that improved $E(R^a/R)$ compared with its unmodified value ($\gamma = 0$). The optimal value of γ depended on the heritabilities, size of experiment, and number of traits. The variance of R^a/R was reduced as γ increased, perhaps a desirable feature since, in practice, selection programmes are conducted only once and the response is, therefore, not accumulated over many replicates. For $\gamma = 1$, all R^a are those given by the base index and $\text{var}(R^a) = 0$. Examination of individual replicates indicated that when R^a/R was very poor before modification, it was usually vastly improved by taking $\gamma > 0$; if R^a/R was very good before modification, i.e. close to unity, it was further slightly improved or was only slightly reduced.

It is clear from Table 1 that the value of γ which maximizes $E(R^a)$ depends on the heritabilities, number of traits, and size of experiment. The optimal value of γ is investigated further in Table 2. The optimal γ appears to be greater the more alike are the heritabilities and the smaller the experiment, but it is rather insensitive to the number of traits and the actual values of the heritabilities although the improvement in $E(R^a/R)$ is greater with many traits and low heritabilities. If all the heritabilities (population roots) are equal, the optimum value of γ is unity, i.e., the base index in which the index weights equal the economic weights is best. The problem of choosing γ in any particular case is a critical one which we consider further.

One situation where it is obvious that the estimates are defective is when the sample covariance matrix, $\hat{\mathbf{G}}$, or alternatively $\hat{\mathbf{P}}^{-1}\hat{\mathbf{G}}$, has one or more negative roots, which is not unlikely if there are many traits (Hill and Thompson, 1978). A reasonable procedure in

Table 2

Mean values over 500 *replicates of* R^a/R *for* $\gamma = 0$ *and for* $\gamma =$ *optimal value, where* γ *values tried were* 0.0, 0.1, 0.2, ... , 1.0. *All traits are of equal economic importance.* $n = 16$

Heritabilities	Number of sires														
	25			50			100			200			400		
	R^a/R		γ opt	R^a/R		γ opt	R^a/R		γ opt	R^a/R		γ opt	R^a/R		γ opt
	$\gamma = 0$	γ opt		$\gamma = 0$	γ opt		$\gamma = 0$	γ opt		$\gamma = 0$	γ opt		$\gamma = 0$	γ opt	
.1 .2	.746	.918	.8	.894	.969	.6	.961	.981	.4	.984	.989	.3	.992	.994	.1
.1 .4	.927	.958	.4	.972	.979	.2	.988	.989	.1	.994	.995	.1	.997	.997	0
.1 .6	.968	.976	.2	.986	.988	.1	.994	.994	.1	.997	.997	0	.998	.998	0
.3 .4	.929	.992	.8	.973	.993	.7	.988	.995	.5	.994	.996	.3	.997	.998	.2
.3 .6	.959	.980	.4	.983	.988	.3	.992	.994	.2	.996	.997	.1	.998	.998	0 or .1
.3 .8	.974	.982	.2	.989	.991	.2 or .1	.995	.995	.1	.997	.997	0	.999	.999	0
.1 .2 .3	.767	.945	.8	.897	.960	.6	.955	.973	.4	.979	.984	.2 or .3	.990	.991	.1
.1 .4 .7	.929	.954	.3	.969	.975	.2	.985	.987	.1	.993	.993	.1	.996	.996	0
.1 .5 .9	.950	.963	.2	.978	.981	.1	.989	.990	.1	.995	.995	0	.997	.997	.2
.3 .4 .5	.894	.984	.8	.953	.986	.7	.978	.990	.5	.989	.993	.3 or .4	.995	.996	.2
.3 .6 .9	.944	.969	.4	.975	.981	.3	.988	.990	.2	.994	.995	.1	.997	.997	0
.3 .7 .9	.946	.970	.4	.975	.982	.3	.988	.990	.1	.994	.995	.1	.997	.997	0 or .1
.1 .2 .3 .4	.800	.937	.8	.907	.955	.5	.956	.970	.4	.978	.983	.2	.989	.990	.1
.1 .4 .7 .9	.927	.951	.3	.966	.972	.2	.984	.985	.1	.992	.992	.1	.996	.996	0
.3 .4 .5 .6	.883	.977	.8	.945	.981	.6	.973	.986	.5	.987	.991	.3	.993	.995	.2
.2 .5 .8 .9	.927	.959	.4	.966	.975	.3	.984	.986	.1	.992	.993	.1	.996	.996	0

such a case would be to eliminate the offending variable(s) by setting the negative roots of $\hat{\mathbf{G}}$ or $\hat{\mathbf{P}}^{-1}\hat{\mathbf{G}}$ to zero, in effect reducing the number of independent variables. This would be standard practice in canonical analysis. An alternative would be to 'bend' the estimates by reducing the range of the roots as proposed, because it is not just the negative roots which are in error. Since the optimum value of γ cannot be predetermined for any single data set (replicate), one possible procedure would be to choose γ just sufficiently large that the smallest root of $\hat{\mathbf{P}}^{*-1}\hat{\mathbf{G}}^*$ equals zero. Examples using the first replicates sampled which had at least one negative root of $\hat{\mathbf{P}}^{-1}\hat{\mathbf{G}}$ for particular sets of parameters are given in Table 3, where the procedure of setting negative roots to zero, or 'bending' until the smallest root was zero are compared. Both procedures led to improved values of achieved response, R^a, but the bending procedure was almost always better. In the examples with two variables, where initially some R^a values were negative (i.e. selecting in the wrong direction), the R^a values obtained by bending were always positive, even though this was not always so if only the negative root was set to zero. In general, the improvements were smaller with more traits, and the bending did not always then lead to a higher R^a value, but the reduction, if present, was always small and in replicates in which R^a was always high.

Whilst the above procedure for choice of γ when a negative root is found seems quite effective, it does not tell us how to operate more generally when all roots are positive and parameters are, of course, unknown. One simple method would be to choose γ solely as a function of the size of the sample used. For $n = 16$ individuals per family, on the basis of the results shown in Table 2, reasonable choices of γ if $s = 25$ would be 0.5 when the number of traits is two or three, and 0.6 when the number of traits is four. If $s = 50$, reasonable choices of γ would be 0.3 when the number of traits is two or three, and 0.4 when the number of traits is four. If $s = 100$, $\gamma = 0.2$ would be a reasonable choice. Examples using this procedure are given in Table 4 for a range of population parameters. This procedure was effective, especially, for the smaller numbers of families. The procedure certainly has the merit of simplicity.

Table 3

Effect on R^a/R in individual replicates, where $\hat{\mathbf{G}}$ has at least one negative eigenvalue before modification (initial), of either setting the negative roots to zero (A) *or 'bending' till all roots are zero or positive* (B). *All traits are of equal economic importance.* $s = 50$, $n = 16$

Heritabilities:	.05 .1			.05 .1 .1		
	R^a/R			R^a/R		
	Initial	Modified		Initial	Modified	
		A	B		A	B
	−.300	−.175	.052	.942	.999	.996
	−.027	.755	.946	−.908	.283	.962
	−.924	−.235	.973	.886	.902	.906
	.545	.702	.799	.015	.756	.881
	.637	.839	.912	.652	.690	.721
	−.999	−.268	.450	.409	.762	.832
	.281	.643	.900	.808	.840	.909
	.184	.717	.945	−.116	.511	.747
Mean:	−.075	.372	.747	.336	.718	.869

Heritabilities:	.05 .1 .1 .2			.05 .05 .1 .1 .15		
	R^a/R			R^a/R		
	Initial	Modified		Initial	Modified	
		A	B		A	B
	.822	.895	.899	.869	.884	.938
	.828	.871	.882	.458	.506	.778
	.944	.944	.934	.980	.980	.965
	.924	.935	.933	.528	.772	.824
	.844	.933	.939	.787	.821	.885
	.723	.884	.954	.493	.689	.846
	.815	.864	.917	.651	.749	.866
	.077	.545	.799	.795	.852	.884
Mean:	.747	.859	.907	.695	.782	.873

A more complicated procedure for obtaining γ would be first to assume that the sample roots were the true roots and find the value of γ which would maximize R^a/R for this set of parameters, given s, n and the economic weights. This value of γ would then be used to bend the estimates. Such a procedure cannot be used iteratively since the sample estimates are biased, and a larger and larger value of γ is chosen as the roots converge, eventually giving $\gamma = 1$. A check on this procedure is not feasible by Monte Carlo methods, for individual replicate sample matrices have to be taken as a starting point for further replicates to find the optimum value of γ. Thus the Taylor series approximation referred to previously was adopted.

Before the Taylor series approximation was used to evaluate the above procedure it was checked for a range of parameters against the Monte Carlo simulation. Results are shown in Table 5 only for $s > 100$; with smaller numbers of families the agreement was poorer. With this restriction, however, there is seen to be good agreement between the approximate and simulated values of $E(R^a/R)$ before bending, and reasonably good agreement between the predicted optimal value of γ and the corresponding mean bent values $E(R^a/R)$.

Table 4

Values of R^a/R obtained when γ is chosen solely as a function of family size, for different sets of heritabilities and family sizes. All traits are of equal economic importance. $n = 16$

Heritabilities		R^a/R					
	s:	25		50		100	
	γ:	0.0	0.5	0.0	0.3	0.0	0.2
.1 .2		.746	.882	.894	.947	.961	.975
.1 .4		.927	.954	.972	.978	.986	.989
.1 .6		.968	.955	.986	.982	.994	.991
.3 .4		.929	.983	.973	.987	.988	.992
.3 .6		.959	.980	.983	.988	.992	.994
.3 .8		.974	.974	.989	.988	.995	.994
Mean:		.917	.955	.966	.978	.986	.989
	γ:	0.0	0.5	0.0	0.3	0.0	0.2
.1 .2 .3		.767	.923	.897	.945	.955	.969
.1 .4 .7		.929	.948	.969	.973	.985	.986
.1 .5 .9		.950	.948	.978	.976	.989	.988
.3 .4 .5		.894	.972	.953	.976	.978	.985
.3 .6 .9		.944	.968	.975	.981	.988	.990
.3 .7 .9		.946	.975	.975	.982	.988	.990
Mean:		.905	.955	.941	.972	.980	.985
	γ:	0.0	0.6	0.0	0.4	0.0	0.2
.1 .2 .3 .4		.800	.935	.907	.952	.956	.968
.1 .4 .7 .9		.927	.939	.966	.966	.984	.984
.3 .4 .5 .6		.883	.972	.945	.976	.973	.982
.2 .5 .8 .9		.927	.954	.966	.973	.984	.986
Mean:		.884	.950	.946	.967	.974	.980

The Taylor series approximation was, therefore, thought justified, particularly since R^a is fairly constant over a range of γ values near the optimum and the simulation results are subject to sampling error. In Table 6, use of a fixed value of γ (0.2) is compared with the use of the γ value determined to be optimal from the sample roots. There is no benefit in using the latter.

5. Discussion

A selection index computed from estimates of parameters based on limited data can be far short of the efficiency of an optimum index computed from the parameters themselves. We have shown that, rather than discard such an index altogether, it can be improved by bending the estimates in a simple way. We do not rule out the possibility that better methods of modifying the estimates exist; we have concentrated solely on changing the sample roots and have left the corresponding eigenvectors unchanged. We suspect, however, that other procedures which modify the roots of $\hat{\mathbf{G}}$ or similar matrices in order to reduce their spread will have similar properties to the procedure used here. We need to

Table 5

Comparison of Monte Carlo simulation (MCS) *and Taylor series approximation* (TS) *in the prediction of R^a/R before modification (initial), optimal γ value and R^a/R at the optimal γ value*

Heritabilities		R^a/R					
		Initial, $\gamma = 0$		γ opt		γ opt	
		MCS*	TS	MCS	TS	MCS	TS
.1 .2	100	.961	.970	.981	.981	.4	.4
	200	.984	.985	.989	.988	.3	.3 or .2†
	400	.992	.993	.994	.994	.1 or .2	.1
.3 .4	100	.988	.988	.995	.995	.5	.5
	200	.994	.994	.996	.996	.3	.3
	400	.997	.997	.998	.998	.2	.2
.1 .6	100	.994	.994	.994	.994	.1	.0
	200	.997	.997	.997	.997	.0	.0
	400	.998	.998	.998	.998	.0	.0
.3 .8	100	.995	.995	.995	.995	.1	.1
	200	.997	.997	.997	.997	.0	.0
	400	.999	.999	.999	.999	.0	.0
.1 .2 .3	100	.955	.962	.973	.974	.4	.4
	200	.979	.981	.984	.985	.2 or .3	.2
	400	.990	.990	.991	.991	.1	.1
.3 .4 .5	100	.978	.980	.990	.990	.5	.5
	200	.989	.990	.993	.993	.3 or .4	.3
	400	.995	.995	.996	.996	.2	.2

* Mean of 500 replicates.

† Pair of values equally good. Values of R^a/R were computed for γ increasing in steps of 0.1.

Table 6

Comparison, using Taylor series, of R^a/R for individual replicates, obtained before modification (initial), by setting $\gamma = 0.2$, regardless of sample roots, and $\gamma = x$, the optimal value of γ assuming the sample roots are the population roots. $s = 100$, $n = 16$

Heritabilities:	.1 .2			.1 .2 .3			.1 .2 .3 .4		
	R^a/R			R^a/R			R^a/R		
	Initial	$\gamma = .2$	$\gamma = x$	Initial	$\gamma = .2$	$\gamma = x$	Initial	$\gamma = .2$	$\gamma = x$
	.902	.912	.932	.912	.918	.918	.960	.990	.996
	.993	.989	.982	.865	.934	.902	.977	.989	.989
	.999	.996	.985	.922	.956	.956	.969	.985	.985
	.977	.994	.999	.984	.983	.981	.995	.999	.999
	.953	.987	.983	.997	.995	.987	.978	.994	.994
	.958	.956	.951	.998	.999	.996	.959	.967	.966
	.926	.976	.976	.983	.989	.987	.975	.966	.954
	.965	.962	.954	.996	.997	.997	.991	.993	.993
Mean:	.959	.971	.970	.957	.971	.965	.975	.985	.984

emphasize, lest there by any confusion, that although we have dealt solely with uncorrelated variables, the methodology applies generally.

The main difficulty lies, of course, in choosing the appropriate value for the bending factor, γ, in the absence of prior information. We assume that there is little or no reliable prior knowledge about the parameters, otherwise some kind of Bayesian procedure could be adopted, which could be computationally difficult even if logically sound. We have made two suggestions: (i) if any roots of $\hat{\mathbf{P}}^{-1}\hat{\mathbf{G}}$ are negative, bend until the smallest is zero; (ii) bend on the basis of the sample size alone. The latter procedure has not been fully worked out, however, as our purpose in this paper has not been to delve into computational details, but to put forward a general procedure for discussion.

The methodology can readily be extended to take account of selection incorporating relativest information, since the same basic covariance matrices $\hat{\mathbf{P}}$ and $\hat{\mathbf{G}}$ are still required. In a hierarchical data structure of sire families, dam families and within-dam families, the fewest degrees of freedom are at the sire level, suggesting that the estimates of the sire mean square matrix of the genetic covariance matrix computed from it should be bent towards the dam or within-dam family component. The more realistic case of unbalanced data sets presents more difficulties since that initial parameter estimation procedure is more complicated and open to choice. However, this is not the place for a full analysis.

Acknowledgement

This work was supported in part by a grant from the Meat and Livestock Commission of Great Britain.

Résumé

Une méthode, dénommée 'courbure' est proposée pour modifier les estimations des matrices de covariance génétique ($\hat{\mathbf{G}}$) et phénotypique ($\hat{\mathbf{P}}$) que l'on utilise dans la construction d'index de sélection génétique pour deux (ou d'avantage) caractères. Si $\hat{\mathbf{P}}$ et $\hat{\mathbf{G}}$ sont estimées à partir des matrices de covariance inter- et intra-classes, $\mathbf{B}$ et $\mathbf{W}$ respectivement, dans une analyse de variance multivariate à un seul facteur, alors la méthode consiste dans la contraction de toutes les valeurs propres de la matrice produit $\mathbf{W}^{-1}\mathbf{B}$ vers leur moyenne, mais sans modifier les vecteurs propres correspondants.

L'intérêt de la méthode de modification est étudiée par les méthodes de simulation de Monte-Carlo et d'approximation par les séries de Taylor. En général, la méthode de modification améliore la réponse à la sélection, le degré d'amélioration dépendant des paramètres (inconnus) de la population et de l'effectif de l'échantillon servant à l'estimation. En pratique, on a quelque difficulté à choisir la valeur du facteur de 'courbure'; quelques méthodes très simples pour le faire sont proposées. Si quelques estimations des paramètres semblent défectueuses (en dehors de leurs limites de validité), une méthode simple et effective pour le choix du facteur de courbure consiste à contracter les valeurs propres de sorte qu'elles soient toutes non négatives.

References

Cochran, W. G. (1951). Improvement by means of selection. *Proceedings of Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman (ed.), 449–470. Berkeley: University of California Press.

Campbell, N. A. (1980). Shrunken estimators in discriminant and canonical variate analysis. *Applied Statistics* **29,** 5–14.

Harris, D. L. (1964). Expected and predicted progress from index selection involving estimates of population parameters. *Biometrics* **20,** 46–72.

Hayes, J. F. and Hill, W. G. (1980). A reparameterization of a genetic selection index to locate its sampling properties. *Biometrics* **36,** 237–248.

Hazel, L. N. (1943). The genetic basis for constructing selection indexes. *Genetics* **28,** 476–490.

Henderson, C. R. (1963). Selection index and expected genetic advance. In *Statistical Genetics and Plant Breeding*, W. D. Hanson and H. F. Robinson (eds), 141–163. Publication 982, National Academy of Sciences. Washington, D.C.: National Research Council.

Hill, W. G. and Thompson, R. (1978). Probabilities of non-positive definite between-group or genetic covariance matrices. *Biometrics* **34,** 429–439.

Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: biased estimation for non-orthogonal problems. *Technometrics* **12,** 55–67.

Sales, J. and Hill, W. G. (1976a). Effect of sampling errors on efficiency of selection indices. 1. Use of information from relatives for single trait improvement. *Animal Production* **22,** 1–17.

Sales, J. and Hill, W. G. (1976b). Effect of sampling errors on efficiency of selection indices. 2. Use of information on associated traits for improvement of a single important trait. *Animal Production* **23,** 1–14.

Smith, H. F. (1936). A discriminant function for plant selection. *Annals of Eugenics* **7,** 240–250.

Williams, J. S. (1962a). Some statistical properties of a genetic selection index. *Biometrika* **49,** 325–337.

Williams, J. S. (1962b). The evaluation of a selection index. *Biometrics* **18,** 375–393.

Received February 1980; *revised August* 1980

APPENDIX

The Taylor series approximation to the expected value, $E(R^a)$, of R^a is

$$E(R^a) = R^a\,|_{\hat{\mathbf{y}}=\mathbf{y}} + \tfrac{1}{2}\sum_i\sum_j \mathrm{cov}(\hat{y}_i, \hat{y}_j)\,\frac{\partial^2 R^a}{\partial \hat{y}_i \partial \hat{y}_j}\,\Big|_{\hat{\mathbf{y}}=\mathbf{y}}, \tag{A1}$$

where $\mathbf{y}$ is a vector of the elements b_{ij} of $\mathbf{B}$, and w_{ij} of $\mathbf{W}$. We need only consider the genetic and phenotypic variables in their transformed state (see text). The variances and covariances of $\mathbf{y}$ are given by Hayes and Hill (1980); for example,

$$\mathrm{var}(b_{ii}) = 2\{1+(n-1)t_i\}^2/(s-1).$$

It is convenient to write the achieved response $R^a = (\mathbf{a}'\hat{\mathbf{G}}\hat{\mathbf{P}}\mathbf{G}\mathbf{a})(\mathbf{a}'\hat{\mathbf{G}}\hat{\mathbf{P}}^{-1}\mathbf{P}\hat{\mathbf{P}}^{-1}\hat{\mathbf{G}}\mathbf{a})^{-\frac{1}{2}}$ as $4(\mathbf{a}'\hat{\mathbf{T}}\hat{\mathbf{P}}\mathbf{T}\mathbf{a})(\mathbf{a}'\hat{\mathbf{T}}\hat{\mathbf{P}}^{-1}\mathbf{P}\hat{\mathbf{P}}^{-1}\hat{\mathbf{T}}\mathbf{a})^{-\frac{1}{2}} = 4YZ^{-\frac{1}{2}}$, say, where $\hat{\mathbf{T}} = (\mathbf{B}-\mathbf{W})/n$ is the estimate of the between-sire covariance matrix.

The derivatives $\partial^2 R^a/\partial b_{ij}^2$ in terms of $\partial Y/\partial b_{ij}$, $\partial^2 Y/\partial b_{ij}^2$, $\partial Z/\partial b_{ij}$ and $\partial^2 Z/\partial b_{ij}^2$, and these derivatives in terms of $\partial\hat{\mathbf{T}}/\partial b_{ij}$, together with the corresponding derivatives with respect to w_{ij}, are given in the Appendix to Hayes and Hill (1980). The only change is that the derivatives of $\hat{\mathbf{T}}$ now involve terms in γ and are given below.

Let $\mathbf{D}_{ij}$ be a $p \times p$ matrix with 1 in the ijth and jith positions and zeros elsewhere. Then evaluating at the parameter level, we have

$$\frac{\partial\hat{\mathbf{T}}}{\partial b_{ij}} = \frac{1}{n}(1-\gamma)\mathbf{D}_{ij} = \frac{\partial\hat{\mathbf{P}}}{\partial b_{ij}},$$

$$\frac{\partial\hat{\mathbf{T}}}{\partial b_{ii}} = \frac{1}{n}(1-\gamma)\mathbf{D}_{ii} + \frac{\gamma}{np}\frac{1}{1-t_i}(\mathbf{I}-\mathbf{T}) = \frac{\partial\hat{\mathbf{P}}}{\partial b_{ii}},$$

$$\frac{\partial\hat{\mathbf{T}}}{\partial w_{ij}} = \frac{1}{n}\mathbf{D}_{ij} + \frac{\gamma}{n}\left\{1+\frac{n}{p}\sum\left(\frac{t_i}{1-t_i}\right)\right\}\mathbf{D}_{ij},$$

$$\frac{\partial\hat{\mathbf{T}}}{\partial w_{ii}} = -\frac{1}{n}\mathbf{D}_{ii} - \frac{\gamma}{np}\left\{\frac{1+(n-1)t_i}{(1-t_i)^2}(\mathbf{I}-\mathbf{T}) + \frac{\gamma}{np}\left\{p+n\sum\left(\frac{t_i}{1-t_i}\right)\right\}\mathbf{D}_{ii},\right.$$

$$\frac{\partial\hat{\mathbf{P}}}{\partial w_{ij}} = \frac{n-1}{n}\mathbf{D}_{ij} + \frac{\gamma}{n}\left\{1+\frac{n}{p}\sum\left(\frac{t_i}{1-t_i}\right)\right\}\mathbf{D}_{ij} \quad \text{and}$$

$$\frac{\partial\hat{\mathbf{P}}}{\partial w_{ii}} = \frac{n-1}{n}\mathbf{D}_{ii} - \frac{\gamma}{np}\left\{\frac{1+(n-1)t_i}{(1-t_i)^2}\right\}(\mathbf{I}-\mathbf{T}) + \frac{\gamma}{np}\left\{p+n\sum\left(\frac{t_i}{1-t_i}\right)\right\}\mathbf{D}_{ii}.$$

A1 can now be evaluated.