Inequality constrained estimation of genetic parameters

T. R. Famula

Department of Animal Science, University of California, Davis, CA 95616, USA

Summary. A method is presented for computing estimates of genetic parameters under linear inequality constraints such that solutions are within theoretical limits. The method produces biased estimators, yet a small scale numerical study, also presented, shows that the inequality constrained estimators have a small mean squared error of prediction than the best of unbiased estimators. The increase in efficiency of estimation is particularly useful for traits where heritability is near the boundary values of zero or one.

Key words: Variance components – Heritability – Genetic correlation

Introduction

One of the principal aims of animal breeding is the identification of sources of variation. Associated with this goal is the characterization of the direction and magnitude or relationships between traits. To this end, the estimation of components of variance and covariance is a topic which traditionally interests animal geneticists.

Although, by definition, variance components are non-negative parameters, estimates of heritability have been reported below the theoretical limit of zero, as well as above the theoretical limit of one (1.0) (Allaire and Lin 1980; Jungst et al 1981). Similarly, estimates of genetic correlations have been reported outside the parameter space of $[-1.0, 1.0]$ (Thrift et al. 1981). The objective of this paper is to present computing strategies which insure estimates within the defined theoretical boundaries.

The necessary and sufficient conditions for the existence of non-negative quadratic unbiased estimators of linear functions of variance components were examined by LaMotte (1973), who found that for the general mixed linear model, the only variance component which can be so estimated is σ_e^2 (the residual error variance). Hence, for quadratic estimates of variance components, unbiasedness and non-negativity are

incompatible optimality criteria. An extension of LaMotte's (1973) Lemma 3 illustrates a similar incompatibility between unbiasedness and constrained estimates of heritability below 1.0.

Suggested strategies, which could be adopted when negative heritability estimates arise, have been put forward by Searle (1971, section 9.8b), who, in so doing, noted that none are wholly satisfactory. Another possibility is to choose a method which guarantees positive estimates of variance components, for example Henderson's (1973) algorithm for maximum likelihood estimators. However, none of these suggested strategies are applicable to estimates of heritability which are greater than one (1.0).

Computational alternatives which are conceptually straightforward and which add little additional computing effort will be presented. The method permits linear inequality constraints on the solutions for variance and covariance components such that parameter estimates fall within the prescribed by quantitative genetic theory.

Statistical preliminaries

The general mixed linear model assumed throughout this paper is

$$y = X b + Z_1 u_1 + Z_2 u_2 + \ldots + Z_k u_k \tag{1}$$

where y is a vector of N observations, X of order $N \times p$ and Z of order $N \times c_i$ $(i = 1, \ldots, k)$ are known incidence matrices, b is an unknown vector of p fixed effects and the u_i of order c_i are nonobservable vectors of random effects (customarily, $u_k = e$, a vector of error effects and $Z_k = I$ of order N) such that

(i) $E[u_i] = 0$, $i = 1, \ldots, k$
(ii) the elements of u_i are independent with common variance σ_i^2, and
(iii) u_i and u_j are independent for $i \neq j$.

Accordingly y has mean vector

$$E[y] = X b \tag{2}$$

and variance-covariance matrix

$$Var[y] = \sum_{l=1}^{k} Z_i \, Z_i' \, \sigma_i^2 = V. \tag{3}$$

Methods for estimating the variance components, σ_i^2, typically involve the construction and solution of a set of linear equations of the form

$$A \sigma = g \tag{4}$$

where the coefficient matrix A and the vector g are known and σ is a vector of the k unknown variance components. For most methods of estimation g is a vector of k quadratic forms and, for unbiased estimators, $E[g] = A \sigma$. Estimation of genetic parameters with known upper and lower bounds suggests that a solution, σ, to (4) be stated as

$$\text{minimize } \| A \sigma - g \| \text{ subject to } F \sigma \geqq h \tag{5}$$

for F a known $m \times k$ matrix and h a known vector of order m. The elements of F and h are determined by the desired inequality constraints.

The characterize the algorithm by which (5) is solved it is convenient to define

$$\begin{aligned} \Phi(\sigma) &= \| A \sigma - g \| \\ S &= \{ \sigma \,|\, F \sigma \geqq h \}. \end{aligned} \tag{6}$$

The basic iterative algorithm is

Step 1: Find $\sigma \in S$ (if you can't, stop),
Step 2: Find w such that $\sigma + w \in S$ and $\Phi(\sigma + w)$
   $< \Phi(\sigma)$ (if you can't, stop),
Step 3: Set new solution for σ to $\sigma + w$,
Step 4: Go to step 2.

The vector w (or the negative of this vector) is sometimes called the dual vector for problem (5). Complete characterization of the solution of (5) is provided by the Kuhn-Tucker Theorem (Fiacco and McCormick 1968, p. 20 and p. 90). In addition, the numerical stability of the solution as well as the convergence properties of the algorithm are developed by Stoer (1971). The discussion of these properties, though important to numerical analysts concerned with the general application of these techniques, is not appropriate here because the particular form of inequality constraints imposed in the stimation of heritability (i.e., that the constraints are continuously differentiable functions of the variance components) do guarantee a unique solution to (5). This is true provided A is of full column rank which, for variance component estimation, it is. For the case where $F \sigma$ is

much greater than h (i.e., $F \sigma \gg h$) the inequality constrained least squares solutions are identical to the solutions of ordinary least squares. Conversely, for the case where $F \sigma$ approach h the inequality constrained estimators behave as if they were linear equality constraints (Liew 1972).

Inequality constrained estimators

Having established that solving (4) subject to a set of linear inequality constraints is a plausible method of obtaining estimates of genetic parameters within theoretical limits, a small example is presented to illustrate the application of the algorithm to the estimation of heritability. Of interest is the determination of the elements of F and h of (5). The example, though unrealistically small, considers the problem of negative estimates of heritability as well as estimates which exceed one (1.0).

Consider Table 1 where six progeny records are classified by sire and dam. An appropriate linear model might be the two-way crossed classification (without interaction) with algebraic representation

$$y_{ijk} = \mu + d_i + s_j + e_{ijk}, \tag{7}$$

where y_{ijk} is the k-th progeny record of the j-th sire mated to the i-th dam, μ is an unknown constant, d_i is the random effect of the i-th dam with $E[d_i] = 0$ and $V[d_i] = \sigma_D^2$, s_j is the random effect of the j-th sire with $E[s_j] = 0$ and $V[s_j] = \sigma_S^2$, and e_{ijk} is a random residual effect with $E[e_{ijk}] = 0$ and $V[e_{ijk}] = \sigma_E^2$. Dam, sire and residual effects are mutually uncorrelated. Under this model it is possible to compute two estimates of narrow sense heritability,

$$h_D^2 = \frac{4\sigma_D^2}{\sigma_D^2 + \sigma_S^2 + \sigma_E^2}$$

and

$$h_S^2 = \frac{4\sigma_S^2}{\sigma_D^2 + \sigma_S^2 + \sigma_E^2}.$$

Method 1 estimates of the variance components (Henderson 1953) can be computed from the following

Table 1. Example data set

	Sires	
	1	2
1	10	7
	12	6
Dams		
2	8	–
	9	–

linear equations

$$1/6 \begin{bmatrix} 16 & 4 & 6 \\ 4 & 16 & 6 \\ -4 & -4 & 18 \end{bmatrix} \begin{bmatrix} \hat{\sigma}_D^2 \\ \hat{\sigma}_S^2 \\ \hat{\sigma}_E^2 \end{bmatrix}$$

$$= \begin{bmatrix} R(d|\mu) \\ R(s|\mu) \\ T - R(\mu, d) - R(\mu, s) + R(\mu) \end{bmatrix} = \begin{bmatrix} .0833 \\ 14.0833 \\ 9.1667 \end{bmatrix} \quad (8)$$

where the $R(\ )$-notation is the reduction in sums of squares for fitting various submodels of (7) (Searle 1971) and T is the total sum of squares. Solving (8) in the usual manner yields estimates

$$\hat{\sigma}_D^2 = -2.4338$$

$$\hat{\sigma}_S^2 = 4.5662$$

$$\hat{\sigma}_E^2 = 3.5294.$$

Clearly both estimates of heritability are outside the theoretical range. Estimation of these variance components under inequality constraints (after (5)) such that heritability estimates are within the interval (0, 1) implies the following linear inequality constraints

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 1 & 1 \\ 1 & -3 & 1 \end{bmatrix} \begin{bmatrix} \hat{\sigma}_D^2 \\ \hat{\sigma}_S^2 \\ \hat{\sigma}_E^2 \end{bmatrix} \cong \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (9)$$

Solving (8) subject to these constraints under the algorithm presented in the previous section yields estimates of the variance components as

$$\hat{\sigma}_D^2 = 0.0000$$

$$\hat{\sigma}_S^2 = 1.3609$$

$$\hat{\sigma}_E^2 = 4.0827$$

with dual vector $\mathbf{w}' = [0.0 \ 0.0193 \ 0.0773]$. Note that heritability estimates have now been placed on the boundary values of zero (0.0) and one (1.0). Moreover, if one assumes that the sire and dam variance are both estimating one-fourth of the additive genetic variance, an equality constraint of $\hat{\sigma}_S^2 = \hat{\sigma}_D^2$ may be employed.

In addition, inequality constrained estimation can be used to compute estimates of the environmental and genetic correlations between two traits within the theoretical range $[-1, 1]$. In the example presented in Van Vleck (1973, p. 51) we first use inequality constraints to estimate the sire and error variances of traits A and B. This is to insure that heritability estimates of traits A and B are within the range $[0, 1]$. Having obtained estimates of $\sigma_{S_A}^2$, $\sigma_{S_B}^2$, $\sigma_{E_A}^2$ and $\sigma_{E_B}^2$ (the sire and error variances for traits A and B) these values are used in establishing the linear inequality constraints for the estimators of the genetic and environmental covariances so the correlations are in $[-1, 1]$. Thus the form of $\mathbf{F}\,\sigma > \mathbf{h}$ for the estimation of the sire and environmental covariances of traits A and B is

$$\begin{bmatrix} 1 & 0 \\ -1 & 0 \\ -3 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} \sigma_{S_A S_B} \\ \sigma_{E_A E_B} \end{bmatrix} \cong \begin{bmatrix} -\hat{\sigma}_{S_A} & \hat{\sigma}_{S_B} \\ -\hat{\sigma}_{S_A} & \hat{\sigma}_{S_B} \\ -C \\ -C \end{bmatrix}$$

where $C = \{(\sigma_{E_A}^2 - 3\,\sigma_{S_A}^2)(\sigma_{E_B}^2 - 3\,\sigma_{S_B}^2)\}^{1/2}$.

Extending this algorithm to multiple trait problems of more than two traits may impose significant computing problems. For example, in a problem with four traits there are six genetic correlations to be estimated. Constraining the estimates of variance and covariance components will greatly increase the order of $\mathbf{F}$ and $\mathbf{h}$. The effect of these simultaneous constraints, along with constraints to satisfy parameter restrictions on all the intermediate partial correlations, may adversely effect the stability of the solutions.[1] Certainly, such multiple trait problems may impose computing difficulties that make this, or perhaps any other computing algorithm, impractical.

An alternative formulation of inequality constraints, in multiple trait models, could be used to force the variance-covariance metrices to be positive definite (suggested by Van Vleck, personal communication). This is equivalent to constraints on the variance and covariance components such that the latent roots of the genetic and environmental covariance matrices are all positive. Constraints on the latent roots are equivalent to those above provided constraints include coefficients used in the estimation of genetic parameters.

Small sample properties of inequality constrained estimation of variance components

A Monte Carlo study is presented to examine the sampling properties of the inequality constrained (IC) estimators of variance components. A small scale numerical study was used to assess the accuracy of IC estimators relative to other methods, because the sampling properties cannot be explicitly derived. The question of primary importance is whether the increase in bias of IC estimation is offset by a smaller variance of prediction errors. That is, is it advantageous to accept some bias in our estimators while decreasing the variance of prediction errors?

Sample data sets were generated from two small designs, described in Table 2, after Quaas and Bolgiano

1 The author wishes to acknowledge the reviewer responsible for offering this comment

Table 2. Description of data sets for small sample numerical study

	Design I	Design II
No. of observations	100	100
No. of herds	5	5
No. of sires	10	10
Percent filled cells	40.0	54.0
No. observations/filled cell	5.0	3.7

(1977). An appropriate model is a two-way crossed classification mixed model without interaction. Observations generated as multivariate normal (International Mathematical and Statistical Libraries 1979) were classified by herd (a fixed effect) and sire (a random effect) and thus, after model (1) $k = 2$, $\mathbf{b}$ is vector of fixed herd effects, $\mathbf{u}_1$, is a nonobservable vector of random sire effects and $\mathbf{u}_2 = \mathbf{e}$, a vector of residuals. Heritability is estimated as $4\,\sigma_1^2/(\sigma_1^2 + \sigma_2^2)$ and the inequality constraints, $\mathbf{F}\,\boldsymbol{\sigma} > \mathbf{h}$, are written as

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} \hat{\sigma}_1^2 \\ \hat{\sigma}_2^2 \end{bmatrix} \geqq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Observations were generated at seven (7) parameter values of σ_1^2 and σ_2^2 corresponding to heritabilities of 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, and 0.95.

The method of variance component estimation used in this sampling experiment was MIVQUE (Rao 1971), although any method which can be formulated after (4) may have been used. The matrix of coefficients and quadratic forms used for MIVQUE is

$$\begin{bmatrix} \operatorname{tr}(\mathbf{Z}'\,\mathbf{P}\mathbf{Z})^2 & \operatorname{tr}(\mathbf{Z}'\,\mathbf{P}^2\mathbf{Z}) \\ \operatorname{tr}(\mathbf{Z}'\,\mathbf{P}^2\mathbf{Z}) & \operatorname{tr}(\mathbf{P}^2) \end{bmatrix} \begin{bmatrix} \hat{\sigma}_1^2 \\ \hat{\sigma}_2^2 \end{bmatrix} = \begin{bmatrix} \mathbf{y}'\,\mathbf{P}\mathbf{Z}\mathbf{Z}'\,\mathbf{P}\mathbf{y} \\ \mathbf{y}'\,\mathbf{P}^2\,\mathbf{y} \end{bmatrix} \quad (10)$$

where for

$$\mathbf{V} = \mathbf{I} + \tilde{\gamma}\,\mathbf{Z}\mathbf{Z}',$$

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\,\mathbf{X}(\mathbf{X}'\,\mathbf{V}^{-1}\,\mathbf{X})^{-}\,\mathbf{X}'\,\mathbf{V}^{-1} \quad \text{and}$$

$$\gamma = \sigma_1^2/\sigma_2^2 = \mathrm{h}^2/(4 - \mathrm{h}^2).$$

The symbol "~" is used to denote a prior value of γ one which does not necessarily agree with the value of γ. The notation MIVQUE $(\tilde{\gamma})$ is used to distinguish between MIVQUE estimators computed at different prior values of γ. Computations were performed at seven (7) possible priors, corresponding to heritabilities of 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, and 0.95.

The sampling experiment proceed with the generation of 500 data sets for each of the 49 possible combinations of true values (of σ_1^2 and σ_2^2) and priors. This permits an examination as to the optimal choice

of a prior when using IC estimators. In addition, the efficiency of estimation was examined to determine if the increased accuracy of IC estimators was offset by the increase in bias.

The percentage of estimates of heritability that were outside the range [0, 1] was a function of the true value of heritability and not the prior. The percentage of negative estimates of heritability range from approximately 45% of the 500 data sets for a true heritability of 0.05 to 2% at a true heritability of 0.95 (independent of the choice of prior). Similarly, the percentage of heritability estimates greater than one (1.0) ranged from less than 1% (2 about of 500 data sets) at a true heritability of 0.05 to approximately 41% of the data sets for a true value of 0.95.

Results are presented for Design II only, although similar findings are established under Design I. Figure 1 presents the estimated bias of σ_1^2, standardized by dividing the difference between estimated and true value by the true standard deviation of sire effects. As expected the bias is minimized for intermediate true values of σ_1^2 (i.e., values corresponding to heritabilities of 0.25 to 0.75). As shown earlier, if $\mathbf{F}\,\boldsymbol{\sigma}$ is much greater than $\mathbf{h}$, IC estimation is equivalent to ordinary least squares, which yields unbiased estimates. On the other hand, at the boundaries, the IC algorithm behaves so as to overestimate σ_1^2 at low heritability and underestimate σ_1^2 at high heritability to insure solutions within theoretical limits. Precisely the opposite is true for inequality constrained estimation of σ_2^2. Figure 2 presents the estimated bias (standardized by dividing by the error standard deviation) of constrained estimates of σ_2^2. As is obvious from the figure, underestimates of σ_2^2 occur at low heritability while overestimation is a problem at high values of heritability. However, with a
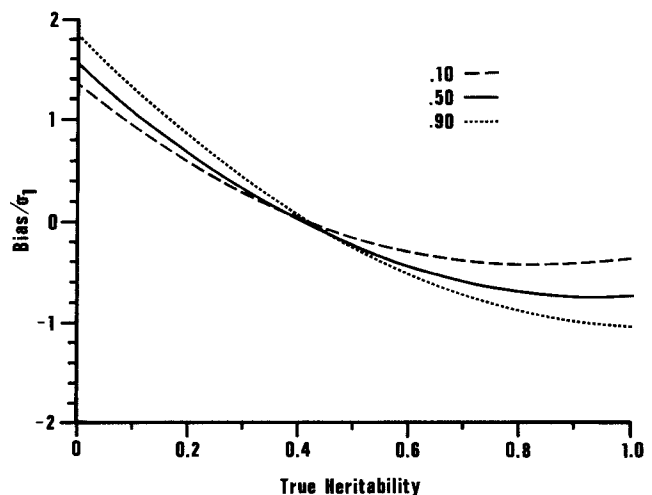

Fig. 1. Estimated bias of inequality constrained estima5ion of σ_1^2 in Design II for three prior heritabilities
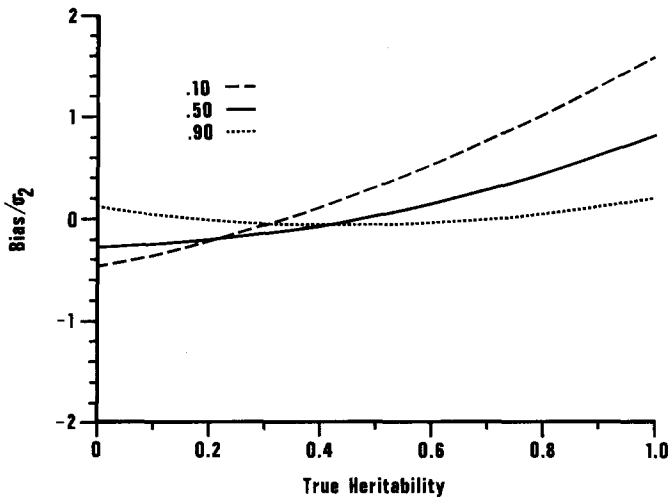
Fig. 2. Estimated bias of inequality constrained estimation of σ_2^2 in Design II for three prior heritabilities
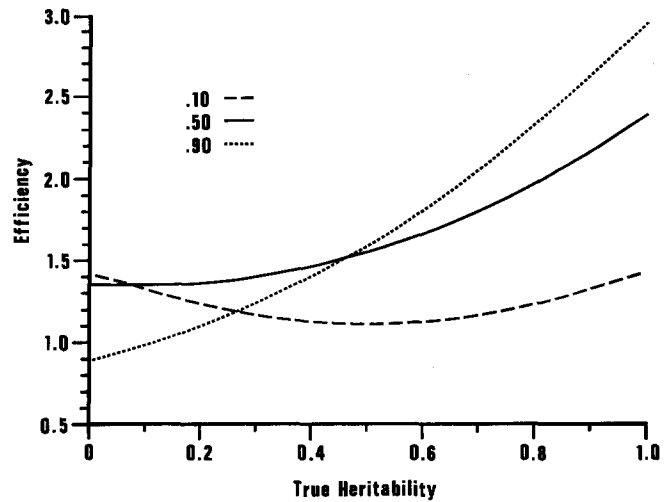

Fig. 4. Estimated efficiency of inequality constrained estimation of σ_1^2 in Design II for three prior heritabilities
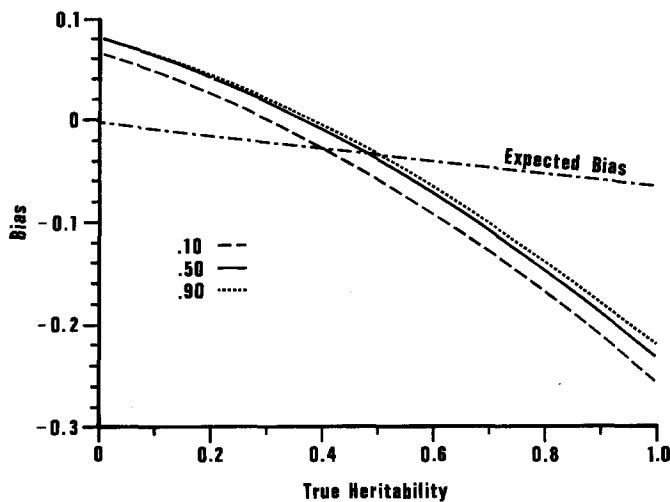

Fig. 3. Estimated bias of inequality constrained estimation of heritability in Design II for three prior heritabilities

prior corresponding to as high a heritability as 0.90, the bias remains close to zero over the range of true parameter values.

Bias in the estimation of heritability is presented in Fig. 3. As in the inequality constrained estimation of σ_1^2, heritability is overestimated for low true parameter values and underestimated for true values above 0.5. Interestingly, the prior assumption for heritability has little influence on the magnitude of the bias as all plots follow a nearly identical path. Also included in Fig. 3 is a plot of the expected bias for standard MIVQUE, which although unbiased for the individual variance components, the expected value of the ratio is not the value of the true parameter (Mood et al. 1974; p. 181).

As the figure indicates, standard MIVQUE is biased downwards across the entire range of true parameter values.

Figure 4 presents the efficiency of IC estimation of σ_1^2 as compared to MIVQUE estimators computed with priors that are identical to true value. Efficiency of estimation is computed as the ratio of the theoretical mean squared error (MSE) of MIVQUE (γ) to the estimated MSE of ICMIVQUE at various prior values. Recall that the MSE of MIVQUE (γ) may be derived analytically given γ, $\mathbf{X}$ and $\mathbf{Z}$ of model (1) whereas the MSE of ICMIVQUE must be estimated from the sampling experiment because there is no explicit expression for the variance of inequality constrained estimators — owing to the iterative nature of the algorithm. Note also that the MSE is the sum of the prediction error variance and the squared bias so that the estimated MSE of ICMIVQUE contains the bias and thus allows an examination of the increased accuracy of constrained estimation relative to the increase in the bias of the estimators. The plots presented in Fig. 4 (which are similar to plots available for Design I) illustrate the ICMIVQUE estimators of σ_1^2 are, over repeated sampling, more efficient than standard MIVQUE. Conversely, Fig. 5 demonstrates a dramatic decline in the efficiency of estimation of σ_2^2 under inequality constraints, particularly in the range of true heritabilities over 0.75. This may have been expected since the only constraint on σ_2^2 is that the estimate be positive, which is satisfied for unbiased estimators (LaMotte 1973). Thus, the inequality constraints imposed on σ_1^2, which simultaneously alter the estimation of σ_2^2, can alter the efficiency of estimation of σ_2^2.
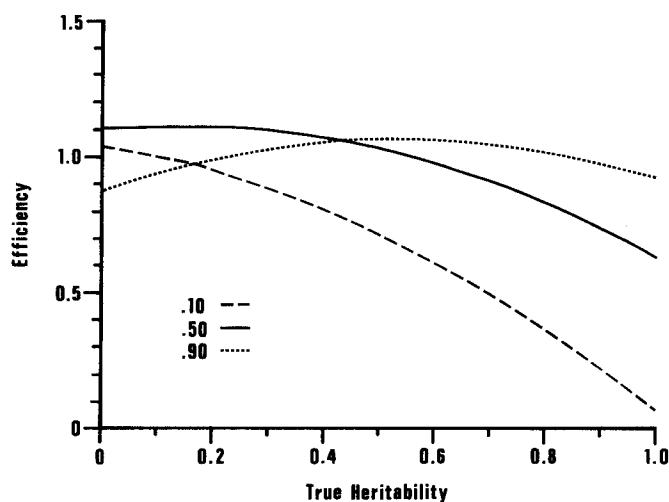
Fig. 5. Estimated efficiency of inequality constrained estimation of σ_2^2 in Design II for three prior heritabilities
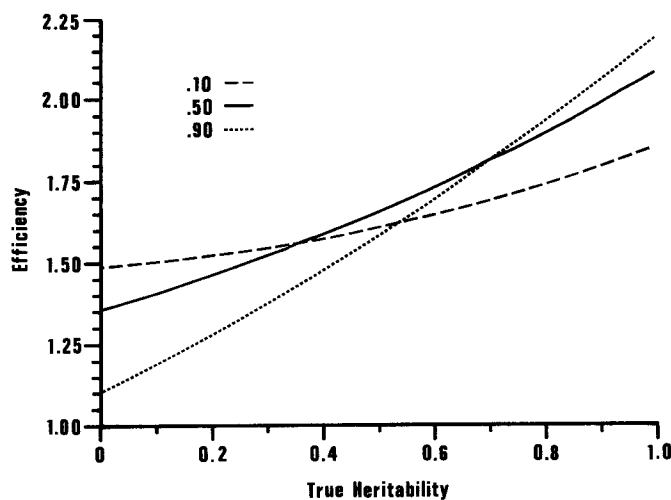

Fig. 6. Estimated efficiency of inequality constrained estimation of heritability in Desgn II for three prior heritabilities

Finally, Fig. 6 presents the increase in efficiency of inequality constrained estimation of heritability over that of unbiased estimation. The increase is particularly evident for high values of heritability. In addition, it is obvious from the several plots in Fig. 6 to deduce the choice of prior most appropriate for a given true value of heritability. Not surprisingly, the assumed prior should agree as closely as possible to the true heritability.

Conclusion

The sampling experiment, presented in the previous section, illustrates the increase in the efficiency of

estimation of heritability (in terms of minimizing the ratio of MSE's of MIVQUE to ICMIVQUE) when inequality constraints are used to insure estimates within theoretical limits. The performance of other variance estimation methods (e.g., Henderson methods) can be examined in a similar fashion. The poor efficiency of estimation for the residual variance suggests the following algorithm for inequality constrained estimation of heritability:

1) Estimate the residual variance with any "local" unbiased method (e.g., the within smallest subclass mean square);

2) Using equality constraints, force the solution for the residual variance to this value, allowing only σ_1^2 to vary so as to insure an estimate of heritability within the interval $[0, 1]$.

This algorithm may increase the efficiency of estimation above that shown in Fig. 6.

References

Allaire, FR, Lin CY (1980) Heritability of age at first calving. J Dairy Sci 63:171–173

Fiacco AV, McCormick QP (1968) Nonlinear programming: sequential unconstrained minimization techniques. Wiley New York

Henderson CR (1953) Estimation of variance and covariance components. Biometrics 9:226–252

Henderson CR (1953) Sire evaluation and genetic trands. In: Proc Animal Breed Genet Symp in honor of Dr. J. L. Lush. Am Soc Animal Sci Am Dairy Sci Assoc, pp 10–41

International Mathematical and Statistical Libraries (1979) IMSL Library Reference Manual, 7th edn IMSL, Houston

Jungst, SB, Christian, LL, Kuhlers DL (1981) Response to selection for feed efficiency in individually fed Yorkshire boars. J Anim Sci 53:323–331

LaMotte LR (1973) On non-negative quadratic unbiased estimation of variance components. J Am Stat Assoc 68:728–730

Liew CK (1972) Inequality constrained least-squares estimation. J Am Stat Assoc 71:746–751

Mood AM, Graybill FA, Boes DC (1974) Introduction to the theory of statistics. McGraw-Hill, New York

Quaas RL, Bolgiano DC (1977) Notes on invariant quadratic unbiased estimators and their sampling variances in the two way mixed model with no interaction. Mimeo, Dept Anim Sci. Cornell University, Ithaca, NY

Rao CR (1971) Estimation of variance and covariance components-MIVQUE theory. J Multivariate Anal 1:257–275

Searle SR (1971) Linear models. Wiley, New York

Stoer J (1971) On the numerical solution of constrained least-squares problems. SIAM J Numerical Anal 8:382–411

Thrift FA, Dillard EV, Shrode RR, Butts WT (1981) Genetic parameter estimates based on selected and control beef cattle populations. J Anim Sci 3:57–61

Van Vleck LD (1973) Summary of methods for estimating genetic parameters using simple statistical models. Dept Anim Sci, Cornell University, Ithaca, NY