University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Faculty Papers and Publications in Animal Science

Animal Science Department

January 1985

Successive Overrelaxation, Block Iteration, and Method of Conjugate Gradients for Solving Equations for Multiple Trait Evaluation of Sires

L. Dale Van Vleck
University of Nebraska-Lincoln, dvan-vleck1@unl.edu

D. J. Dwyer
Cornell University

Follow this and additional works at: https://digitalcommons.unl.edu/animalscifacpub

 Part of the Animal Sciences Commons

Van Vleck, L. Dale and Dwyer, D. J., "Successive Overrelaxation, Block Iteration, and Method of Conjugate Gradients for Solving Equations for Multiple Trait Evaluation of Sires" (1985). *Faculty Papers and Publications in Animal Science*. 292.
https://digitalcommons.unl.edu/animalscifacpub/292

This Article is brought to you for free and open access by the Animal Science Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Papers and Publications in Animal Science by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Successive Overrelaxation, Block Iteration, and Method of Conjugate Gradients for Solving Equations for Multiple Trait Evaluation of Sires

L. D. VAN VLECK and D. J. DWYER
Department of Animal Science
Cornell University
Ithaca, NY 14853

ABSTRACT

A potential difficulty with mixed model equations for multiple trait evaluation of sires is solving the equations as the number of equations increases proportionally to the number of traits. Time required to obtain inverse solutions increases by the number cubed. Thus, iterative procedures often are used. Three iterative procedures, successive overrelaxation, block iteration with relaxation, and the method of conjugate gradients, were compared for four sets of multiple trait equations for sire evaluation. Equations were solved after absorption of equations for random herd-year-season effects. Equations for two and four traits each with test and complete data sets made up the four sets of equations. The two-trait system featured high heritabilities and large negative correlations among effects whereas the four-trait system had low heritabilities and smaller negative correlations. Rate of convergence for block iteration was faster than for successive overrelaxation, especially for the four-trait system and especially for more exacting convergence criteria. The method of conjugate gradients was efficient only for test data sets (30 and 60 equations) and was not competitive with the other methods for complete data sets (1426 and 2852 equations). Test data sets accurately predicted optimum relaxation factors for successive overrelaxation for complete data sets. Optimum relaxation factor for the two-trait system was 1.5 to 1.7 and for the four-trait system was 1.3 to 1.5. Gauss-Seidel iteration took 33 to 400% more rounds than successive overrelaxation with the optimum relaxation factor depending on stopping criteria and data set.

INTRODUCTION

One potential problem with multiple trait mixed model evaluation is difficulty of solving equations. Number of equations to be solved increases by number of traits, number of coefficients increases by number squared, and number of computations required for direct solution increases by the number cubed even though symmetry reduces the number of computations by about half. Efficient computing strategies are needed to facilitate multiple trait evaluation. Prediction of category frequencies for traits such as calving difficulty and type traits by best linear unbiased prediction is a special form of multiple trait evaluation of sires. Such sets of equations were available from analysis of Brown Swiss type data (12). Equations for a test set of data and for the complete data set were available for multiple subtraits of two traits, the first having three categories (equivalent to two traits) and the second having five categories (equivalent to four traits). The two pseudo-multiple trait systems also had different variance-covariance matrices, corresponding to relatively high and relatively low heritabilities.

Canonical transformation (1, 2, 6, 11) can be used to reduce multiple trait, mixed model equations to independent sets of single trait mixed model equations if observations on all traits have the same design matrices and if there is only one random classification other than residual effects. The model underlying available type equations, however, contains two random classifications so that canonical transformation cannot be applied.

The goals of the study were: 1) to find optimum relaxation parameters for successive

Received April 30, 1984.

overrelaxation (SOR) modification of Gauss-Seidel (G-S) iteration, 2) to compare SOR with block iteration (BLOCK), 3) to find optimum relaxation parameters for BLOCK, 4) to compare SOR and BLOCK with the conjugate-gradient (CON GRAD) method of solving equations, and 5) to determine if optimum relaxation parameters can be estimated from a small set of data, eliminating the need to use a large set of data for that purpose.

METHODS

The model for sire evaluation was:

$$\mathbf{y} = \mathbf{Wh} + \mathbf{Xm} + \mathbf{Zs} + \mathbf{e}$$

where:

- **y** is a vector of observations for all traits ordered by traits within animals (each animal has a measurement on each trait that will be zero or one),
- **h** is a vector of random herd-year effects ordered by traits within herd-years,
- **m** is a vector of fixed effects (one for each trait corresponding to the mean frequency for that trait),
- **s** is a vector of random effects of sires of animals ordered by traits within sires,
- **e** is a vector of random residual effects associated with the vector of observations,
- **W** is a matrix of zeros and ones that associates the herd-year effects with the observations,
- **X** is a matrix of zeros and ones associating effects in **m** with observations, and
- **Z** is a matrix of zeros and ones associating sire effects with observations.

The expectation of the observation vector is:

$$E(\mathbf{y}) = \mathbf{Xm}$$

The variance-covariance matrix of the random effects is:

$$V \begin{bmatrix} \mathbf{h} \\ \mathbf{s} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{H} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix}$$

where **H**, **G**, and **R** are block diagonals with the right direct product forms:

$$\mathbf{H} = \mathbf{I}_H * \mathbf{V}_H$$
$$\mathbf{G} = \mathbf{I}_S * \mathbf{V}_S$$
$$\mathbf{R} = \mathbf{I}_N * \mathbf{V}_E$$

where:

- $\mathbf{V}_H$ is the variance-covariance matrix of herd-year effects for multiple traits observed in the same herd-year,
- $\mathbf{V}_S$ is the variance-covariance matrix of sire effects for multiple traits for animals having the same sire,
- $\mathbf{V}_E$ is the variance-covariance matrix of residual effects for traits observed on the same animal, and
- $\mathbf{I}_H$, $\mathbf{I}_S$, and $\mathbf{I}_N$ are identity matrices of order the number of herd-years, sires, and animals.

Mixed model equations for predicting sire values are:

$$\begin{bmatrix} \mathbf{W'R^{-1}W+H^{-1}} & \mathbf{W'R^{-1}X} & \mathbf{W'R^{-1}Z} \\ & \mathbf{X'R^{-1}X} & \mathbf{X'R^{-1}Z} \\ \text{Symmetric} & & \mathbf{Z'R^{-1}Z+G^{-1}} \end{bmatrix} \begin{bmatrix} \mathbf{h} \\ \mathbf{m} \\ \mathbf{s} \end{bmatrix} = \begin{bmatrix} \mathbf{W'R^{-1}y} \\ \mathbf{X'R^{-1}y} \\ \mathbf{Z'R^{-1}y} \end{bmatrix}$$

Before solving for **s**, herd-year equations were absorbed leaving symmetric equations of the form as illustrated for two sires and two traits:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{bmatrix} \begin{bmatrix} s_{11} \\ s_{12} \\ s_{21} \\ s_{22} \\ s_{31} \\ s_{32} \end{bmatrix} = \begin{bmatrix} r_{11} \\ r_{12} \\ r_{21} \\ r_{22} \\ r_{31} \\ r_{32} \end{bmatrix} \quad [1]$$

To simplify description of the computing procedure, notation has been changed slightly. Now the first two equations are for the **m** vector (i.e., $s_{11} = m_1$ and $s_{12} = m_2$), the next two equations are for the first sire, and the last two equations are for the second sire (i.e., first subscript in solution and right-hand side vectors refers to sequentially ordered sire number plus 1 and second subscript to trait).

The matrix of coefficients can be partitioned into blocks of order corresponding to number of traits.

The G-S method of iteration is well known (5, 8) but will be described here for completeness. Let elements of the solution and right-hand side vectors be identified by single subscripts (1, . . ., 6), for the example. Then the jth solution in the nth round of iteration is:

$$s_j^n = \frac{1}{a_{jj}} (r_j - \sum_{i=1}^{j-1} a_{ji}s_i^n - \sum_{i=j+1}^{6} a_{ji}s_i^{n-1})$$

which can be rewritten when w = 1 as:

$$s_j^n = s_j^{n-1} + \frac{w}{a_{jj}} (r_j - \sum_{i=1}^{j-1} a_{ji}s_i^n - a_{jj}s_j^{n-1} - \sum_{i=j+1}^{6} a_{ji}s_i^{n-1}) \quad [2]$$

When 2 > w > 1 this modification of G-S is known as SOR, and when 0 < w < 1 the modification is known as successive under-relaxation (SUR), where w is known as the relaxation factor (e.g., 5, 7, 8, 10, 13).

Thus, when written in this form, G-S is modified easily for SOR or SUR. In addition, calculation yields a difference between each right-hand side and the right-hand side regenerated from solutions calculated to that moment. These residuals were used as an approximation to a standard method of measuring rate of convergence.

Block iteration was suggested by R. L. Quaas (personal communication, 1980) as a method of obtaining more rapid convergence. Subsequently, a textbook reference was found to such a method as early as 1843 (3, 13). Multiple trait equations are blocked naturally when data and solution vectors are ordered by traits within animal and traits within herd-year or sire.

Let partitioned equations [1] be rewritten in block form as:

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix}$$

where each $\mathbf{A}_{ij}$ corresponds to a 2 × 2 submatrix in [1] and each $\mathbf{s}_i$ and $\mathbf{r}_i$ corresponds to a vector of length 2.

Then the block method of iteration corresponding to G-S, SOR, and SUR where **p** = the number of solution vectors each of length 2 can be written as:

$$\mathbf{s}_k^n = \mathbf{s}_k^{n-1} + w\mathbf{A}_{kk}^{-1}(\mathbf{r}_k - \sum_{i=1}^{k-1} \mathbf{A}_{ki}\mathbf{s}_i^n - \mathbf{A}_{kk}\mathbf{s}_k^{n-1} - \sum_{i=k+1}^{p} \mathbf{A}_{ki}\mathbf{s}_i^{n-1}) \quad [3]$$

Computation of each vector of solutions need not involve the inverse of $\mathbf{A}_{kk}$. When the order of each block is small, an inverse procedure may be as efficient as an indirect procedure such as a forward and a backward procedure based on a Cholesky decomposition of $\mathbf{A}_{kk}$,$\mathbf{L}_{kk}$ where $\mathbf{L}_{kk}$ is a lower triangular matrix and $\mathbf{L}_{kk}\mathbf{L}'_{kk} = \mathbf{A}_{kk}$. In either case, $\mathbf{A}_{kk}^{-1}$ or $\mathbf{L}_{kk}$ needs to be calculated only once and not for each round.

Approximate differences of right-hand sides from regenerated right-hand sides can be calculated a block at a time as indicated in [3].

The other method of iteration that was compared is the method of conjugate gradients (CONGRAD) (4, 8, 9). Technically the method is not iterative but is an exact method that gives the exact solution in as many steps as there are equations. The approach to the exact solution appears iterative in nature, and thus, solutions at earlier steps may be sufficiently accurate for most purposes. The method, however, accumulates rounding errors from round to round in contrast to other methods in which the solution vector at the end of any round is simply a new starting point for the next round.

The CONGRAD method has some computational advantages such as requiring less storage space (advantage can be taken of symmetry) and the vector of residuals of right-hand sides and regenerated right-hand sides is calculated directly from round to round without having to save the original right-hand sides. A disadvantage is all of the usual convergence criteria may increase or may decrease from round to round, whereas with other iterative methods for equations guaranteed to converge, convergence criteria generally decrease from round to round. For completeness the CONGRAD method is listed here for equations **As** = **r**.

For a starting point, choose some s_0, e.g., $\mathbf{s}_0 = \mathbf{D}^{-1}\mathbf{r}$, where **D** is the diagonal of **A**.

Let:

$\mathbf{P}_0 \Leftarrow \mathbf{r}$

$\mathbf{P}_0 \Leftarrow \mathbf{P}_0 - \mathbf{A}\mathbf{s}_0$

$\mathbf{r}_0 \Leftarrow \mathbf{P}_0$

where $\Leftarrow$ indicates replace the term on the left of the arrow with the term on the right of the arrow as in Fortran programming. Then for each round i (i = 0, . . .) repeat the following steps until convergence criterion is satisfied.

1) $\mathbf{x}_i \Leftarrow \mathbf{A}\mathbf{p}_i$ (vector)
2) $\lambda_i \Leftarrow \mathbf{p}_i'\mathbf{x}_i$ (scalar)
3) $\alpha_i \Leftarrow (\mathbf{p}_i'\mathbf{r}_i)/\lambda_i$ (scalar)
4) $\mathbf{s}_{i+1} \Leftarrow \mathbf{s}_i + \alpha_i\mathbf{p}_i$ (vector)
5) $\mathbf{r}_{i+1} \Leftarrow \mathbf{r}_i - \alpha_i\mathbf{x}_i$ (will be algebraically equal to vector of residuals from original right-hand sides)
6) $\beta_i \Leftarrow (-\mathbf{r}_{i+1}'\mathbf{x}_i)/\lambda_i$ (scalar)
7) $\mathbf{p}_{i+1} \Leftarrow \mathbf{r}_{i+1} + \beta_i\mathbf{p}_i$

Stopping Point

A difficulty with iterative methods is determining when to stop and accept solutions (5). Convergence criteria based on maximum change from round to round, on sum of absolute changes from round to round, or on sum of squares of changes from round to round are relatively easy to calculate but depend more on a feeling for the data than on an absolute criterion because such criteria depend basically on the trait or traits measured rather than on the coefficient matrix. Therefore, after considerable reflection as well as searching through numerical analyses texts (5, 7, 8, 10), the following criterion was adopted.

The square root of the sum of squares of the right-hand sides, $(\Sigma r_i^2)^{.5} = (\mathbf{r}'\mathbf{r})^{.5}$, the Euclidean norm, if divided by the number of equations is much like the standard deviation of right-hand sides and reflects the scale of traits if a single trait is being used or an average of the scales of traits for a multiple trait system. A common procedure for checking whether a set of solutions satisfies the original equations is to regenerate the right-hand sides as a product of the coefficient matrix and the solution vector. The difference or residual from the original right-hand sides reflects the scale of the trait. Thus, the square root of the sum of squares of residuals divided by the number of equations is similar to the standard deviation of residuals and reflects the scale of the traits and the frequency each trait occurs in the solution vector.

In symbols, for the nth round of iteration:

$$\mathbf{e} = \mathbf{r} - \mathbf{A}\mathbf{s}^n \text{ and } (\Sigma e_i^2)^{.5} = (\mathbf{e}'\mathbf{e})^{.5}$$

The standardized stopping procedure would be to examine the ratio:

$$(\mathbf{e}'\mathbf{e})^{.5}/(\mathbf{r}'\mathbf{r})^{.5}$$

and stop if this ratio is less than some constant, C. One textbook (8) suggested $C \leq .001$. Experience with several sets of equations suggested C between .001 and .0001. The use projected for the solutions may determine the necessary convergence. For ranking animals, a less precise stopping point may be required than when

TABLE 1. Description of data sets for studying optimum solution strategies.

Data set	Number				
	Categories[1]	Animals	Herd-years	Sires	Equations after absorption of herd-years
Test 1	3	21	10	14	30
Complete 1	3	12,838	2,295	712	1,426
Test 2	5	21	10	14	60
Complete 2	5	12,838	2,295	712	2,852

[1] The last category is ignored in the analysis. The solution for that category is obtained as a difference from the other categories.

genetic evaluations are used to predict genetic trend.

The approximation used for e in this study was to calculate the regenerated right-hand side for each equation (see term in square brackets in [2] and [3]) just prior to solving that equation in a particular round of iteration and then use that residual rather than the residual which could be calculated at the end of a round. Approximation to the residual is conservative in the sense that it averages half a round behind the residual calculated at the end of the round. Actual sum of squares of residuals would be smaller at the end of the round.

Starting Point

A problem in comparing iterative procedures is that rate of convergence often depends on the initial solution. A number of arbitrary procedures were tried including 1) dividing right-hand sides by diagonal coefficients, 2) inserting raw means for fixed effects and zeroes for s, 3) solving $\mathbf{A}_{11}\mathbf{s}_1 = \mathbf{r}_1$ for the fixed effects and setting sire effects to zero, and 4) using the same solutions to $\mathbf{A}_{11}\mathbf{s}_1 = \mathbf{r}_1$ for up to three rounds. The conclusion was to start with solutions that approximate the expected values of the solutions. Thus, starting method 3 was employed for all iterative procedures, which is also the first step in a block iteration procedure when the fixed effect equations come first.

Relaxation Parameters

The original intent was to vary the relaxation parameter over the range of .7 to 1.9 by increments of .1 for SOR and BLOCK SOR. This plan was followed for all except the largest set of equations with SOR. Only four relaxers were tried for that set of equations (.7, 1.0, 1.3, and 1.6) because of the time required to obtain convergence for each relaxer and because by that time BLOCK SOR appeared decidely superior to SOR with an optimum relaxation parameter.

Data Sets

Two sets of equations for sire evaluation (12) were chosen for categorically scored traits of front end (5 categories = 4 subtraits) and stature (3 categories = 2 subtraits) corresponding to traits with low and moderately high herit-

TABLE 2. Variance-covariance matrices for test and complete data sets for stature (2 subtraits) and for front end (4 subtraits).

Category	Residual				Herd-year				Sire			
	1	2	3	4	1	2	3	4	1	2	3	4
					Trait 1 (stature)							
1	.2021	−.1895			.0261	−.0216			.0136	−.0105		
2	−.1895	.2195			−.0216	.0203			−.0105	.0085		
					Trait 2 (front end)							
1	.1859	−.1573	−.0130	−.0086	.0276	−.0208	−.0040	−.0023	.0049	−.0033	−.0004	−.0008
2		.2228	−.0311	−.0208		.0201	.0009	−.0001		.0031	−.0003	.0004
3			.0473	−.0016			.0031	−.0002			.0008	−.0002
4	Sym.[1]			.0321	Sym.			.0023	Sym.			.0007

[1] Sym = Symmetry.

TABLE 3. Number of rounds to reach convergence criteria by successive overrelaxation (SOR), by block iteration with relaxation (BLOCK), and by the method of conjugate gradients (CONGRAD) for a trait with three categories (two subtraits).

Stopping point, C[1]	Relaxation parameter .7	.8	.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	CON GRAD
						Test data set, trait 1								
	SOR													
.001	*	88	74	62	52	44	37	30	24	18	15	32	51	8
.0001	*	*	*	*	90	74	61	49	37	26	21	77	97	9
.00001	*	*	*	*	*	*	85	67	51	35	28	*	*	12
						Complete data set, trait 1								
	SOR													
.001	10	10	8	8	8	7	7	7	6	7	7	8	11	15
.0001	72	61	52	45	38	33	28	24	19	15	13	19	35	42
.00001	*	*	*	*	96	80	65	53	40	25	52	76	*	62
						Test data set, trait 1								
	BLOCK													
.001	6	5	4	3	3	4	5	5	7	9	11	14	14	
.0001	8	7	6	4	4	5	6	8	10	14	19	15	15	
.00001	11	9	7	5	5	7	8	10	14	18	24	15	16	
						Complete data set, trait 1								
	BLOCK													
.001	17	14	12	10	9	8	6	6	8	8	8	8	9	
.0001	27	23	19	16	13	11	8	9	11	12	14	13	11	
.00001	38	31	26	21	17	14	9	12	15	13	19	19	16	

[1] $(e'e)^{.5}/(r'r)^{.5} < C$.

*Indicates > 100 rounds.

ability. A small set of data had been used to test the sequence of programs that set up and solved the equations. Descriptions of the four sets of equations are in Table 1. Two basic comparisons were intended: 1) to determine whether the same iteration procedure would be optimum for test and complete data sets and 2) to determine whether the same iteration procedure would be optimum for both traits.

The same variance-covariance matrices for herd-year, sire, and residual effects were used for the pair of test and complete data sets for the same trait. Matrices are in Table 2.

RESULTS AND DISCUSSION

Trait 1. Number of rounds of iteration to reach convergence criteria are in Table 3 for trait 1. The method of conjugate gradients was most efficient for the test data set as might be expected for a system with only 30 equations because exact convergence (unless disrupted by rounding errors) is guaranteed in 30 rounds or less. For the complete data set, CONGRAD, although more efficient than G-S for more strict convergence criteria (C = .0001 and .00001), was not as efficient as SOR with an optimum relaxation parameter. Optimum block iteration was more efficient than optimum SOR for both the test and complete data sets.

Optimum relaxation parameter for SOR for the test data set was a good indicator of the relaxation factor for the complete data set. Although 1.7 was optimum for the test data for all three convergence criteria (C = .001, .0001, .00001), the optimum parameter for the complete data set depended on the convergence criterion but ranged from 1.5 to 1.7 with 1.6 good for all three C's.

Optimum relaxation parameter for block iteration was smaller than for SOR. The test data set also underestimated the optimum

TABLE 4. Number of rounds to reach convergence criteria by successive overrelaxation (SOR), by block iteration with relaxation (BLOCK), and by the method of conjugate gradients (CONGRAD) for a trait with five categories (four substraits).

Stopping point, C[1]	Relaxation parameter													
	.7	.8	.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	CON GRAD
						Test data set, trait 2								
	SOR													
.001	41	35	30	27	24	21	20	19	20	22	22	27	30	16
.0001	*	85	71	59	48	40	32	27	35	45	52	58	62	18
.00001	*	*	*	90	73	58	42	43	52	65	84	90	94	22
						Complete data set, trait 2								
	SOR													
.001	2	m	m	1	m	m	2	m	m	3	m	m	m	9
.0001	24			18			14			19				46
.00001	*			60			35			62				86
						Test data set, trait 2								
	BLOCK													
.001	3	2	2	2	2	2	2	3	3	4	4	6	10	
.0001	5	4	3	3	3	4	4	5	6	8	10	15	21	
.00001	7	6	5	3	4	5	6	8	9	12	17	29	32	
						Complete data set, trait 2								
	BLOCK													
.001	5	5	4	4	4	4	4	3	5	5	5	7	8	
.0001	11	10	8	7	6	5	5	6	7	8	9	10	11	
.00001	18	15	13	10	9	7	7	8	10	11	12	13	14	

[1] $(e'e)^{.5}/(r'r)^{.5} < C$.

*Indicates > 100 rounds.

[m]Indicates these relaxers were not tried.

relaxation parameter for the complete data set (1.0 or 1.1 *vs.* 1.3).

Trait 2. Numbers of rounds to reach convergence criteria for trait 2 are in Table 4. Because categories of trait 2 correspond to four subtraits, the number of equations is twice the number for trait 1. The pattern for trait 2 parallels that for trait 1. The main difference is that optimum relaxation parameter for SOR was smaller for trait 2 than for trait 1. In fact, although not all relaxation parameters were tried for the complete set for trait 2, it appears that the test data set did a good job of estimating the relaxation parameter for the complete data set. Efficiency of convergence for BLOCK for data set 2 did not depend as much on choice of relaxation parameter as compared to data set 1; i.e., a wider range of relaxation parameters was nearly optimum for data set 2 than for data set 1.

Increase in efficiency due to optimum relaxation for SOR or for block iteration was greater for more severe convergence criteria (Table 5). BLOCK was more efficient than other methods. The method of conjugate gradients generally was not competitive.

CONCLUSIONS

Extrapolating from results for one or two kinds of data to other sets of data seems extraordinarily hazardous for iterative procedures. What may work well for one set may not work at all for similar sets.

Results presented here do suggest, somewhat surprisingly, that a test set of data might be used to indicate a reasonably optimum relaxation parameter for a larger set of data having the same variance-covariance structure.

It is likely that block iteration will be more efficient than G-S iteration, and efficiency of

TABLE 5. Comparison of number of rounds of iteration to convergence for Gauss-Seidel (G-S), optimum successive overrelaxation (SOR), optimum block (BLOCK), and conjugate-gradient (CONGRAD) iteration for the complete data sets for traits 1 and 2.

Stopping point, C[1]	Trait 1				Trait 2			
	G-S	SOR (1.6)	BLOCK (1.3)	CONGRAD	G-S	SOR (1.3)	BLOCK (1.3)	CONGRAD
.001	8	7	6	15	1	2	4	9
.0001	45	15	8	42	18	14	5	46
.00001	>100	25	9	62	60	35	7	86

[1] $(e'e)^{.5}/(r'r)^{.5} < C$.

convergence will be greater for more strict convergence criteria. Relaxation parameters greater than 1 and up to 1.6 or 1.7 are likely to provide as good or better efficiency of convergence than G-S iteration, although for these equations G-S iteration at a less severe convergence criterion (.001) was surprisingly efficient, possibly due in part to having starting values relatively close to the final solutions.

A step-wise procedure for obtaining an optimum iteration procedure for an evaluation program which will be run a number of times would seem to be the following. If the equations have a natural block structure as do multiple trait equations: 1) choose the desired convergence, 2) for a small set of data, find the optimum relaxation parameter for block iteration, 3) use that relaxation parameter on a representative complete set of data and increase and decrease by some interval (e.g., .1) until the optimum relaxation parameter is bracketed, 4) occasionally repeat step 3 for larger sets of data.

ACKNOWLEDGMENTS

This research was supported in part by a grant from Eastern AI Cooperative, Inc., Ithaca, NY.

REFERENCES

1 Anderson, T. W. 1958. An introduction to multivariate statistical analysis. John Wiley and Sons, Inc., New York, NY.
2 Arnason, T. 1982. Prediction of breeding values for multiple traits in small non-random mating (horse) populations. Acta Agric. Scand. 32:171.
3 Gerling, C. L. 1843. Page 96 *in* Die Ausgleichs-Rechungen der practischen Geometrie oder die Methode der Kleinsten Quadrate mit ihren Anwendungen fur geodatische Aufgarben, Gothe, Hamburg: 96. [As cited by R. S. Varga, (13, p. 96)].
4 Hestenes, M. R., and E. Stiefel. 1952. Method of conjugate gradients for solving linear systems. J. Res. Natl. Bur. Stand. 49:409.
5 Hornbeck, R. W. 1975. Numerical methods. Quantum Publishers, Inc., New York, NY.
6 Hotelling, H. 1936. Relations between two sets of variates. Biometrika 28:321.
7 Isaacson, E. 1966. Analysis of numerical methods. John Wiley and Sons, Inc., New York, NY.
8 Jennings, A. 1977. Matrix computation for engineers and scientists. John Wiley and Sons, Inc., New York, NY.
9 Reid, J. K. 1971. On the method of conjugate gradients for the solution of large sparse systems of linear equations. Page 231 *in* Large sparse sets of linear equations. J. K. Reid, ed. Academic Press, New York, NY.
10 Stewart, G. W. 1973. Introduction to matrix computations. Academic Press, New York, NY.
11 Thompson, R. 1977. Estimation of quantitative genetic parameters. Page 639 *in* Proc. Int. Conf. Quant. Genet. E. Pollack, O. Kempthorne, and T. B. Bailey, ed. Iowa State Univ. Press, Ames.
12 Van Vleck, L. D., and P. J. Karner. 1980. Sire evaluation by best linear unbiased prediction for categorically scored type traits. J. Dairy Sci. 63:1328.
13 Varga, R. S. 1962. Matrix iterative analysis. Prentice-Hall, Englewood, NJ.