Prediction of future records

C. R. Henderson
DEPARTMENT OF ANIMAL SCIENCE
CORNELL UNIVERSITY, ITHACA, NY 14853

1. INTRODUCTION

The problem discussed in this paper is the following. The breeder has available a set of records, y, from which he wishes to predict the relative magnitude of future records. These predictions are then used to make selection decisions. For example, y might be milk records of the progeny of a set of dairy bulls from which the production of future progeny of these bulls is to be predicted. As a second example, y could represent milk records on all cows in a herd from the time that milk recording began until the present. The dairyman wishes to predict future milk yields on those cows still surviving that have records as well as the yields on heifers that have not yet reached producing age. As a third example, y represents a record on each of three traits on a set of animals. The breeder wishes to predict some function (not necessarily linear) of the records on future progeny of these animals.

Now, of course, the method of prediction should be dictated by what we want to do with the predictions, by computational costs, and by how well the results are understood by potential users. From the viewpoint of what we want to do with the predictions, two alternative and not necessarily incompatible goals are:

1. Maximizing the expected value of the functions of future records when individuals are selected upon the basis of their predictions.
2. Finding the most accurate "estimates" of functions of future records.

I tend to consider the second of these the more important since my responsibility as a researcher is not to make selections but rather to assist breeders in their selections by providing them with the best possible tools for selection. One must realize that the best selection for one breeder may not necessarily be the best for some other breeder, for their goals may differ, particularly with respect to relative emphasis to be placed on different traits. Consequently, we should provide "best" evaluations for each trait and then tell the breeder how to utilize this information to meet his goals. Further, the question of cost complicates the selection problem. The best selection from the standpoint of genetic improvement is not necessarily the best from an economic viewpoint. For example, the cost of an insemination from one dairy bull may be very different from that for another one. With the above considerations in mind, my efforts have been and will continue to be to emphasize most accurate evaluation possible within the limitations of computing costs and the costs of obtaining records suitable for evaluation. Accordingly, this paper deals with evaluation methods, which I choose to call prediction.

The problem of predicting future records is a special case of prediction of a nonobservable random vector, say $\underline{g}$, that is distributed jointly with $\underline{y}$. In the case of prediction of records, $\underline{g}$ is some function of future records. Also, $\underline{g}$ could be some variable, the prediction of which leads to a prediction of future records. Examples are given later. (An underscored lower case letter denotes a column vector, an underscored upper case letter a matrix, and an italicized lower case letter a scalar. Any of these can have one or more subscripts.)

2. BEST PREDICTION

We wish to predict g_i from a function of $\underline{y}$, say $f_i(\underline{y}) = \hat{g}_i$. How should this be done? It seems logical to try to minimize the prediction error, $\hat{g}_i - g_i$, or some function of error. The most commonly used notion is to minimize, if possible, $E(\hat{g}_i - g_i)^2$. Such a prediction method is called "best prediction" (BP), or the predictor itself is called a "best predictor" (also BP). It turns out that

$$\hat{g}_i = E(g_i|\underline{y}) \ ;$$

that is, the minimizing value is the conditional mean of g_i given $\underline{y}$. Some additional properties of BP are:

1. $E(\hat{\underline{g}}) = E(\underline{g})$. Thus, the predictor is unbiased.
2. $E[(\hat{\underline{g}}-\underline{g}),(\hat{\underline{g}}-\underline{g})'] =$ variance-covariance matrix of the prediction errors $= E_y[\mathrm{Var}(\underline{g}|\underline{y})]$, that is, the mean in repeated sampling of the conditional variance-covariance matrix of $\underline{g}$ given $\underline{y}$.
3. $\mathrm{Cov}(\hat{\underline{g}},\underline{g}') = \mathrm{Var}(\hat{\underline{g}})$.
4. $\mathrm{Cov}(\hat{\underline{g}},\underline{y}') = \mathrm{Cov}(\underline{g},\underline{y}')$.
5. Correlation between $\hat{g}_i$ and g_i is maximum and is $\sigma_{\hat{g}_i}/\sigma_{g_i}$.
6. If $\underline{y}'$ can be partitioned into subvectors $[\underline{y}_1' \ \cdots \ \underline{y}_q']$ each of the same length and with the corresponding elements of $\underline{g}$ as $[g_1 \ \cdots \ g_q]$ such that the density function $h(\underline{y}_i, g_i)$ is the same for all $i = 1, \ldots, q$ and such that $(\underline{y}_i, g_i)$ is distributed independently of $(\underline{y}_j, g_j)$ for all $i \neq j$, the following property proved by Cochran (1951) is true. If truncation is carried out on $\hat{g}_i$, the resulting expected mean of the selected g_i is maximum for *any* function of $\underline{y}$.

Proofs of properties 1-5 above can be found in Searle (1973).

Now, it would appear to be impossible to find a more desirable prediction method than the one described above. Unfortunately, however, best predictors can seldom be found. Note that the form of distribution must be known, we must have the mathematical tools required for finding the conditional mean, and finally the

numerical values of the parameters of the distribution must be known. Is the last requirement ever fulfilled? Further, in order for property 6 to be true, minimum requirements would be that the candidates for selection are unrelated (so that g_i are independent) and every candidate must have the same type of information and number of records (so that every $\underline{y}_i$ has the same length and the joint distribution of every $[\underline{y}_i, g_i]$ is the same).

Fortunately, if the distribution is multivariate normal, best prediction simplifies greatly, for then the conditional mean of $\underline{g}$ is linear in $\underline{y}$ and the only parameter values needed are the first and second moments. Best prediction in this case is identical computationally to selection index discussed in the next section.

3. BEST LINEAR PREDICTION (SELECTION INDEX)

Suppose we are willing to restrict ourselves to predictions that are linear in $\underline{y}$. That is,

$$\hat{g}_i = a_i + \underline{b}_i\underline{y}$$

where a_i is some scalar and $\underline{b}_i$ is a vector of weights chosen so that $E(\hat{g}_i - g_i)^2$ is minimum. It is found that

$$\hat{\underline{g}} = E(\underline{g}) + \underline{C}'\underline{V}^{-1}[\underline{y} - E(\underline{y})]$$

where $\underline{C}' = \mathrm{Cov}(\underline{g},\underline{y})$ and $\underline{V} = \mathrm{Var}(\underline{y})$. Some useful properties of best linear prediction (BLP) are:

1. $E(\hat{\underline{g}}) = E(\underline{g})$. Thus, the predictor is unbiased.
2. $\mathrm{Var}(\hat{\underline{g}}-\underline{g}) = \mathrm{Var}(\underline{g}) - \underline{C}'\underline{V}^{-1}\underline{C}$.
3. $\mathrm{Var}(\hat{\underline{g}}) = \mathrm{Cov}(\hat{\underline{g}},\underline{g}') = \underline{C}'\underline{V}^{-1}\underline{C}$.
4. The correlation between $\hat{g}_i$ and g_i is maximum in the class of linear functions of the form $a_i + \underline{b}_i'\underline{y}$.
5. When the distribution is multivariate normal, $\hat{\underline{g}}$ is the Bayesian estimate of $\underline{g}$ under a quadratic loss function (Solomon, 1971).
6. When the distribution is multivariate normal, the probability of a correct choice between g_i and g_j is maximized for any linear function of $\underline{y}$ (see Henderson, 1963).
7. BP of $\underline{m}'\underline{g}$, a linear function of $\underline{g}$, is $\underline{m}'\hat{\underline{g}}$, that is, the same linear function of the BP of $\underline{g}$.

Best linear prediction is easy to compute when the number of records is small and is well known by animal breeders, but it does have the serious deficiency that $E(\underline{y})$ is usually if not always unknown. In contrast, $\mathrm{Var}(\underline{y})$ and $\mathrm{Cov}(\underline{g},\underline{y}')$ are often assumed to have been estimated quite accurately, and consequently most animal breeders seem willing to substitute these estimates for the parameter values and then to proceed as though these are the true values. With regard to the mean of $\underline{y}$, the assumption is usually made that it has the form, $\underline{X}\underline{\beta}$, where $\underline{X}$ is a known matrix and $\underline{\beta}$ is a fixed, unknown vector. Often, however, certain of the elements of $\underline{\beta}$ have been estimated quite accurately from prior data and can be used accordingly, as for example effects of age on milk production. Let $\underline{\beta}' = [\underline{\beta}_1' \vdots \underline{\beta}_2']$ and assume that $\underline{\beta}_1$ is known. Then, the record vector can be defined as $\underline{y}-\underline{X}_1\underline{\beta}_1$.

It is most unlikely, however, that all elements of $\underline{\beta}$ have been well estimated. In fact, in some cases, the sample at hand provides the only information. Consequently, the usual procedure used by breeders is to estimate $\underline{X}\underline{\beta}$ from the data, compute $\underline{y}-\underline{X}\hat{\underline{\beta}}$, and then proceed to apply selection index (BLP) as though $\hat{\underline{\beta}} = \underline{\beta}$. The next section deals with the question of the optimum choice of an estimate of $\underline{\beta}$.

4. BEST LINEAR UNBIASED PREDICTION

It was pointed out in the preceding section that the best linear prediction contains the numerical values of the means of $\underline{g}$ and of the records. It has also been assumed, naively I think, that the animals to be evaluated are all members of the same population. Consequently, their means have been assumed to be zero. Of course, any other constant would do equally well since its

choice would not affect the ranking of predictions. Having made the assumption of equal means, the practice has then been to estimate $\underline{X\beta}$ by some method, sometimes by arithmetic means of levels of factors but often by regular least squares. The crucial question is what is the best estimate to use to substitute in selection index. Henderson (1963) solved a more general problem; namely, he found the linear function of $\underline{y}$ with mean $\underline{0}$ which minimizes the prediction error variance. That is, he minimized $E(\hat{g}_i - g_i)^2$ subject to $E(\hat{g}_i) = 0$. The result was $\hat{\underline{g}} = \underline{C}'\underline{V}^{-1}(\underline{y} - \underline{X}\hat{\underline{\beta}})$ where $\hat{\underline{\beta}}$ is a solution to the generalized least squares equations, $\underline{X}'\underline{V}^{-1}\underline{X}\hat{\underline{\beta}} = \underline{X}'\underline{V}^{-1}\underline{y}$. This method has been called best linear unbiased prediction (BLUP). Note that this is the selection index result with $\underline{X}\hat{\underline{\beta}}$ substituted for $\underline{X\beta}$.

It is probably not surprising to find that of *all* estimates of $\underline{X\beta}$, the GLS (also best linear unbiased) estimate is best to use in selection index. What is not so intuitively obvious is that of *all* linear functions of $\underline{y}$ the best one is in fact selection index using a GLS estimate of $\underline{X\beta}$ for the parameter value. Some useful properties of BLUP are:

1. $E(\hat{\underline{g}}) = E(\underline{g})$. This, of course, was required in the method.
2. $\mathrm{Var}(\hat{\underline{g}} - \underline{g}) = \mathrm{Var}(\underline{g}) - \underline{C}'\underline{V}^{-1}\underline{C} + \underline{C}'\underline{V}^{-1}\underline{X}(\underline{X}'\underline{V}^{-1}\underline{X})^{-}\underline{X}'\underline{V}^{-1}\underline{C}$ where $(\underline{X}'\underline{V}^{-1}\underline{X})^{-}$ denotes any generalized inverse of $\underline{X}'\underline{V}^{-1}\underline{X}$.
3. $\mathrm{Var}(\hat{\underline{g}}) = \mathrm{Cov}(\hat{\underline{g}}, \underline{g}') = \underline{C}'\underline{V}^{-1}\underline{C} - \underline{C}'\underline{V}^{-1}\underline{X}(\underline{X}'\underline{V}^{-1}\underline{X})^{-}\underline{X}'\underline{V}^{-1}\underline{C}$.
4. The correlation between $\hat{g}_i$ and g_i is maximum in the class of linear functions of $\underline{y}$ that have zero expectations.
5. When the distribution is multivariate normal, $\hat{\underline{g}}$ is BLUE and MLE of $E(\underline{g}|\underline{y})$, the conditional mean of $\underline{g}$ for fixed $\underline{y}$.
6. When the distribution is multivariate normal, the linear function in the class of such functions with mean zero that maximizes the probability of selecting the better of g_i and g_j is $\hat{g}_i - \hat{g}_j$.
7. BLUP of $\underline{m}'\underline{g}$ is $\underline{m}'\hat{\underline{g}}$.

4.1. Different and Unknown Means of Variables to Be Predicted

Suppose now that we cannot assume validly that all animals come from the same population. Rather, we assume that

$$g_i = \underline{k}_i'\underline{\beta} + \underline{m}_i'\underline{u}$$

where $\underline{u}$ is assumed to have null means. It seems logical in this case to predict g_i by the sum of the GLS estimate of $\underline{k}_i'\underline{\beta}$ plus BLUP of $\underline{m}_i'\underline{u}$. Henderson (1975c) proved that this minimizes $E(\hat{g}_i - g_i)^2$ in the class of linear functions of $\underline{y}$ that are unbiased, that is, for which $E(\hat{g}_i) = E(g_i) = \underline{k}_i'\underline{\beta}$.

When choices must be made between animals from different populations with different means and these means are unknown, the use of the method of this section has not been proved to have optimum properties with regard to ranking, as is the case with ranking normally distributed variables with known or with equal means. This is one of the unsolved problems in selection.

4.2. An Efficient Method for Computing BLP and BLUP

The foregoing methods described for BLP (selection index) and BLUP are difficult computationally if there are many elements in $\underline{y}$, that is, if there are many records. The difficulty arises in that the large nondiagonal matrix $\underline{V}$ must be inverted. A much easier method was described by Henderson (1950) for certain prediction problems.

Suppose now that a linear model for $\underline{y}$ can be written as

$$\underline{y} = \underline{X\beta} + \underline{Zu} + \underline{e}$$

where $\underline{X}$ and $\underline{Z}$ are known matrices; $\underline{\beta}$ is a fixed vector, usually unknown; $\underline{u}$ and $\underline{e}$ are nonobservable random vectors with $\mathrm{Var}(\underline{u}) = \underline{G}\sigma^2$, $\mathrm{Var}(\underline{e}) = \underline{R}\sigma^2$, and $\mathrm{Cov}(\underline{u}, \underline{e}') = \underline{0}$. $\underline{G}$ and $\underline{R}$ are known, nonsingular matrices; σ^2 is a scalar, the value of which is not needed in prediction. Now write the following equations:

$$\begin{pmatrix} \underline{X}'\underline{R}^{-1}\underline{X} & \underline{X}'\underline{R}^{-1}\underline{Z} \\ \underline{Z}'\underline{R}^{-1}\underline{X} & \underline{Z}'\underline{R}^{-1}\underline{Z} + \underline{G}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\underline{\beta}} \\ \hat{\underline{u}} \end{pmatrix} = \begin{pmatrix} \underline{X}'\underline{R}^{-1}\underline{y} \\ \underline{Z}'\underline{R}^{-1}\underline{y} \end{pmatrix}.$$

Henderson et al. (1959) proved that $\hat{\underline{\beta}}$ is a GLS solution for $\underline{\beta}$, and Henderson (1963) proved that $\hat{\underline{u}}$ is BLUP of $\underline{u}$. Then, of course, because of the invariance property of BLUP, if $g_i = \underline{m}_i'\underline{u}$, BLUP of g_i is $\underline{m}_i'\hat{\underline{u}}$.

Also, if $\underline{\beta}$ is known, $\hat{\underline{u}}$, the solution to the following equations, is BLP (selection index);

$$(\underline{Z}'\underline{R}^{-1}\underline{Z}+\underline{G}^{-1})\hat{\underline{u}} = \underline{Z}'\underline{R}^{-1}(\underline{y}-\underline{X}\underline{\beta}) .$$

In most applications, $\underline{R} = \underline{I}$, and $\underline{G}$ has a form that is easy to invert. This method has been applied to some very large problems, for example, the evaluation of approximately 3000 sires (the $\underline{u}$ vector) with progeny in at least 240,000 herd-year-seasons comprising most of the elements of the $\underline{\beta}$ vector.

The method has the further advantage in that the sampling variances derive directly from a g-inverse of the coefficient matrix. Let a g-inverse be

$$\begin{pmatrix} \underline{C}_{11} & \underline{C}_{12} \\ \underline{C}_{12}' & \underline{C}_{22} \end{pmatrix} .$$

Then, if $\underline{K}'\underline{\beta}$ is estimable

$$\mathrm{Var}(\underline{K}'\underline{\beta}) = \underline{K}'\underline{C}_{11}\underline{K}\sigma^2 ;$$

$$\mathrm{Cov}(\underline{K}'\underline{\beta},\hat{\underline{u}}'-\underline{u}') = \underline{K}'\underline{C}_{12}\sigma^2 ;$$

$$\mathrm{Var}(\hat{\underline{u}}-\underline{u}) = \underline{C}_{22}\sigma^2 .$$

5. PREDICTION WHEN VARIANCES AND COVARIANCES ARE UNKNOWN

It has been assumed in the preceding sections that the needed variances and covariances, $\underline{V}$ and $\underline{C}$ (or $\underline{G}$ and $\underline{R}$ in the mixed linear model) are known.

How do we proceed if these parameters are unknown? Geneticists commonly use a combination of genetic theory and estimates of variances from whatever data have been published to derive estimates of these parameters. If these prior estimates have small sampling variances and if the data to be used for prediction can legitimately be regarded as a random sample from the same population as the data from which the prior estimates were obtained, the

predictions should be about as good as if we knew the variances. Some work has been done on the sensitivity of the predictions to errors in these variances. In general, it appears that accuracy of ranking of single traits is quite insensitive to errors in estimates of variances, but multiple trait selection is much more sensitive to errors.

A particularly difficult problem is that of prediction from a body of data that provides the only information available on the variances. Econometricians have been concerned with this problem with respect to estimation of $\underline{\beta}$. If I understand the literature correctly, they commonly estimate the variance-covariance matrix of the error vector by some method that yields consistent estimators, regard these estimates as the parameter values, and then apply generalized least squares. This seems somewhat strange to me since consistency is a large sample property and many of their samples are not large. In contrast, animal breeders, often with very large samples, seem to prefer to use unbiased estimates of variances in their selection index methods and more recently in their BLUP equations. The question as to which is the best procedure to follow if the variances are unknown indeed may be an unsolvable problem.

One possibility is to estimate variances by maximum likelihood. Then, if these estimates are substituted for the corresponding parameter values in the mixed model equations, the resulting solution to $\underline{\beta}$ is maximum likelihood, and the solution to $\underline{u}$ is the maximum likelihood estimate of the conditional mean of $\underline{u}$ given $\underline{y}$. The mixed model equations solution can be used conveniently to estimate the variances as pointed out by Henderson (1973). Whether or not the resulting predictions are better or worse than those resulting from substitution of variances estimated by easier methods such as Henderson's (1953) methods is not known. Of course, it is known that maximum likelihood does have certain optimum properties in large samples.

The variance components model is

$$\underline{y} = \underline{X}\underline{\beta} + \underline{Z}_1\underline{u}_1 + \ldots + \underline{Z}_c\underline{u}_c + \underline{e}$$

where $\mathrm{Var}(\underline{u}_i) = \underline{I}\sigma_i^2$, $\mathrm{Var}(\underline{e}) = \underline{I}\sigma_e^2$, $\mathrm{Cov}(\underline{u}_i, \underline{u}_j') = \underline{0}$ for all $i \neq j$, and $\mathrm{Cov}(\underline{u}_i, \underline{e}') = \underline{0}$ for all i. Then, $\underline{Z}'\underline{R}^{-1}\underline{Z}+\underline{G}^{-1}$ of the mixed model equations becomes

$$\begin{bmatrix} \underline{Z}_1'\underline{Z}_1 + \frac{\underline{I}\sigma_e^2}{\sigma_1^2} & \underline{Z}_1'\underline{Z}_2 & \cdots & \underline{Z}_1'\underline{Z}_c \\ \underline{Z}_1'\underline{Z}_2 & \underline{Z}_2'\underline{Z}_2 + \frac{\underline{I}\sigma_e^2}{\sigma_2^2} & \cdots & \underline{Z}_2'\underline{Z}_c \\ \vdots & \vdots & & \vdots \\ \underline{Z}_c'\underline{Z}_1 & \underline{Z}_c'\underline{Z}_2 & \cdots & \underline{Z}_c'\underline{Z}_c + \frac{\underline{I}\sigma_e^2}{\sigma_c^2} \end{bmatrix}.$$

Let the inverse of this matrix be

$$\begin{bmatrix} \underline{T}_{11} & \underline{T}_{12} & \cdots & \underline{T}_{1c} \\ \underline{T}_{12}' & \underline{T}_{22} & \cdots & \underline{T}_{2c} \\ \vdots & \vdots & & \vdots \\ \underline{T}_{1c}' & \underline{T}_{2c}' & \cdots & \underline{T}_{cc} \end{bmatrix}.$$

The estimate of σ_e^2 is $(\underline{y}'\underline{y} - \underline{\beta}'\underline{X}'\underline{y} - \underline{u}_1'\underline{Z}_1'\underline{y} - \ldots - \underline{u}_c'\underline{Z}_c'\underline{y})/n$ where $\underline{\beta}$ and $\underline{u}$ are solutions to mixed model equations and n is the number of elements in $\underline{y}$. The estimate of σ_i^2 is $(\underline{u}_i'\underline{u}_i + \sigma_e^2 \mathrm{tr}\underline{T}_{ii})/q_i$ where q_i is the number of elements in $\underline{u}_i$. The solution to $\underline{\beta}$ is MLE of $\underline{\beta}$, and the solution to $\underline{u}$ is the MLE of the conditional mean of $\underline{u}$. These equations require, of course, an iterative solution. The proof of these results is given in a paper submitted to *Biometrics* by Henderson, Ufford, and Schaeffer.

6. APPLICATIONS

This section and the remainder of this paper will deal with some applications of the principles presented above and in particular BLUP.

6.1. Prediction of Future Records in an Additive Genetic Model Using Single Records

Suppose that we wish to predict future records of a set of animals, some of which have one record and others none. An additive genetic model is assumed. Let the model for the records be

$$\underline{y} = \underline{X}\underline{\beta} + \underline{Z}\underline{a} + \underline{e}$$

where $\underline{a}$ represents the vector of additive genetic merits of the animals to be evaluated, and $\underline{Z}$ is a matrix with as many rows as there are records and as many columns as animals to be evaluated. The columns pertaining to animals with no records have all null elements. For example, if there are 5 animals, the 2nd, 3rd, and 4th only having records,

$$\underline{Z} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

$$\mathrm{Var}(\underline{a}) = \underline{A}h^2\sigma_y^2 ;$$

$$\mathrm{Var}(\underline{e}) = \underline{I}(1-h^2)\sigma_y^2 ;$$

$$\mathrm{Cov}(\underline{a}, \underline{e}') = \underline{0}$$

where $\underline{A}$ is the numerator relationship matrix of the animals to be evaluated, h^2 = heritability of the trait in an unselected, noninbred population, and σ_y^2 is the variance of records in this population. Then, the variance of the record on the ith animal is $(1+f_i h^2)\sigma_y^2$, and the covariance between records on the ith and jth animals is $a_{ij}h^2\sigma_y^2$ where f_i is the inbreeding coefficient of the ith animal and a_{ij} is the numerator relationship between the ith and jth animal. Then, the BLUP equations for predicting $\underline{a}$ are

$$\begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z}+\underline{A}^{-1}\frac{1-h^2}{h^2} \end{bmatrix} \begin{bmatrix} \hat{\underline{\beta}} \\ \hat{\underline{a}} \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{y} \\ \underline{Z}'\underline{y} \end{bmatrix}.$$

The prediction of future records on the animals evaluated in these equations is $\hat{\underline{a}}$.

The additive genetic merit (and predicted record) of an animal not appearing in the $\underline{a}$ vector above is

$$\underline{c}'\underline{A}^{-1}\hat{\underline{a}}$$

where $\underline{c}'$ is the row vector of numerator relationships between this animal and those in the equations above. In the special case of an animal with no records and no progeny with records but with evaluated sire and dam, $\underline{c}'\underline{A}^{-1}\hat{\underline{a}}$ simplifies to the mean of the $\hat{a}_i$ of its parents.

If the number of animals in $\underline{A}$ is large, the rapid methods described by Henderson (1975a, 1975b, 1976a, 1976b) with a modification of one of them by Quaas (1976) can be used to compute $\underline{A}^{-1}$.

In this application using $\underline{A}$ and in subsequent ones, it is essential that all records be included that were used in making selection decisions.

6.2. Prediction of Future Records in an Additive Model Using Multiple Records

In this section, we deal with traits on which more than one record can be obtained, lactation yields in dairy cows, for example. Let the model for all records be

$$\underline{y} = \underline{X}\beta + \underline{Z}\underline{a} + \underline{Z}\underline{p} + \underline{e}$$

where $\underline{a}$ is the vector of additive genetic merits of all animals, $\underline{p}$ is the vector of permanent environmental effects, and $\underline{Z}$ is a matrix of 1's and 0's relating a particular record to a particular individual. If the individual has no record, the column of $\underline{Z}$ relating to that animal is null. For example, if there are 3 animals to be evaluated, the first with 2 records, the second with 1 record, and the third with none,

$$\underline{Z} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} .$$

$$\mathrm{Var}(\underline{a}) = \underline{A}h^2\sigma_y^2 ;$$
$$\mathrm{Var}(\underline{p}) = \underline{I}(r-h^2)\sigma_y^2 ;$$
$$\mathrm{Var}(\underline{e}) = \underline{I}(1-r)\sigma_y^2$$

where r = repeatability and σ_y^2 = variance of single records in a noninbred population. $\mathrm{Cov}(\underline{a},\underline{p}')$, $\mathrm{Cov}(\underline{a},\underline{e}')$, and $\mathrm{Cov}(\underline{p},\underline{e}')$ are null. The BLUP equations are

$$\begin{pmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z}+\underline{A}^{-1}\frac{1-r}{h^2} & \underline{Z}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} & \underline{Z}'\underline{Z}+\underline{I}\frac{1-r}{r-h^2} \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{\underline{a}} \\ \hat{\underline{p}} \end{pmatrix} = \begin{pmatrix} \underline{X}'\underline{y} \\ \underline{Z}'\underline{y} \\ \underline{Z}'\underline{y} \end{pmatrix} .$$

Note that all elements of $\hat{\underline{p}}$ pertaining to animals with no records all equal 0. Consequently, the equations and coefficients corresponding to $\underline{p}$ of these animals can be deleted.

The prediction of the future production of the ith animal is $\hat{a}_i+\hat{p}_i$. The prediction of production of animals without records is the same as in the preceding section.

6.3. Prediction of Future Production Under an Additive Plus Dominance Model

The model is the same as in Section 6.2 except now a dominance term has been added and the method is restricted to a noninbred population.

The model is

$$\underline{y} = \underline{X}\beta + \underline{Z}\underline{a} + \underline{Z}\underline{d} + \underline{Z}\underline{p} + \underline{e}$$

where $\underline{d}$ refers to dominance value with partitioning as described by Cockerham (1963). The covariances between $\underline{d}$ and $\underline{a}$, $\underline{p}$, and $\underline{e}$ are null.

$$\mathrm{Var}(\underline{a}) = \underline{A}\sigma_a^2 ;$$
$$\mathrm{Var}(\underline{d}) = \underline{D}\sigma_d^2$$

where $\underline{D}$ is the dominance relationship matrix;

$$\mathrm{Var}(\underline{p}) = \underline{I}\sigma_p^2 ;$$
$$\mathrm{Var}(\underline{e}) = \underline{I}\sigma_e^2 .$$

Then the variance of a record is $\sigma_a^2 + \sigma_d^2 + \sigma_p^2 + \sigma_e^2$. Covariance between two records on the same individual is $\sigma_a^2 + \sigma_d^2 + \sigma_p^2$. Covariance between records on the ith and jth animal is $\alpha_{ij}\sigma_a^2 + d_{ij}\sigma_d^2$. The BLUP equations are

$$\begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} & \underline{X}'\underline{Z} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z}+\underline{A}^{-1}\frac{\sigma_e^2}{\sigma_a^2} & \underline{Z}'\underline{Z} & \underline{Z}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} & \underline{Z}'\underline{Z}+\underline{D}^{-1}\frac{\sigma_e^2}{\sigma_d^2} & \underline{Z}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} & \underline{Z}'\underline{Z} & \underline{Z}'\underline{Z}+\underline{I}\frac{\sigma_e^2}{\sigma_p^2} \end{bmatrix} \begin{bmatrix} \hat{\underline{\beta}} \\ \hat{\underline{a}} \\ \hat{\underline{d}} \\ \hat{\underline{p}} \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{y} \\ \underline{Z}'\underline{y} \\ \underline{Z}'\underline{y} \\ \underline{Z}'\underline{y} \end{bmatrix} .$$

The prediction of a future record (expressed as a deviation from 0) on the ith animal included in the equations above is

$$\hat{a}_i + \hat{d}_i + \hat{p}_i .$$

The predicted production of an animal not appearing in these equations is

$$\underline{c}'\underline{A}^{-1}\hat{\underline{a}} + \underline{d}'\underline{D}^{-1}\hat{\underline{d}}$$

where $\underline{c}'$ is the row vector of the numerator relationship between the animal and those above and $\underline{d}'$ is a similar vector for the dominance relationship.

The methods of this section can be extended to other components of genetic variation, e.g., additive × additive.

6.4. Evaluation of Multiple Traits in an Additive Genetic Model

This section deals with the problem of evaluating several correlated traits from records on the individual and its relatives. Suppose that the additive genetic variance matrix of traits in the same animal is $\underline{G}_0$, the corresponding environmental variance-covariance matrix is $\underline{R}_0$, the genetic and environmental variables are uncorrelated, and consequently the phenotypic variance-covariance matrix is $\underline{G}_0 + \underline{R}_0 = \underline{V}_0$, say. Suppose that the ith animal has a record on each of the traits and the model for the vector of records is denoted by $\underline{y}_i$, the model for which is

$$\underline{y}_i = \underline{X}_i\underline{\beta} + \underline{a}_i + \underline{e}_i ;$$
$$\mathrm{Var}(\underline{a}_i) = \underline{G}_0 ;$$
$$\mathrm{Var}(\underline{e}_i) = \underline{R}_0 ;$$
$$\mathrm{Cov}(\underline{a}_i, \underline{e}_i') = \underline{0} .$$

Now, if $\underline{\beta}$ is known and we do not wish to use records on relatives of the ith animal, the regular selection index evaluation of

$$\underline{a}_i = \underline{G}_0\underline{V}_0^{-1}(\underline{y}_i - \underline{X}_i\underline{\beta}) .$$

Then, if the net merit is defined as $\underline{m}'\underline{a}_i$, the solution for this animal is $\underline{m}'\hat{\underline{a}}_i$.

If certain records are missing, we substitute for $\underline{G}_0$ and $\underline{V}_0$, these matrices reduced according to records missing.

Now suppose we want to use records on relatives and we cannot assume that $\underline{\beta}$ is known. We first consider the case in which we have n animals each with a record on all q of the traits. Then, if we order the records traits within animals, we have for our general BLUP equations $\underline{X}' = [\underline{X}_1' \ \underline{X}_2' \ \cdots \ \underline{X}_n']$, $\underline{Z} = \mathrm{I}$, and

$$\underline{G} = \begin{bmatrix} a_{11}\underline{G}_0 & a_{12}\underline{G}_0 & \cdots & a_{1n}\underline{G}_0 \\ a_{12}\underline{G}_0 & a_{22}\underline{G}_0 & \cdots & a_{2n}\underline{G}_0 \\ \vdots & \vdots & & \vdots \\ a_{1n}\underline{G}_0 & a_{2n}\underline{G}_0 & & a_{nn}\underline{G}_0 \end{bmatrix}$$

where a_{ij} are elements of $\underline{A}$, the numerator relationship matrix. $\underline{G}$ can be written as $\underline{A}*\underline{G}_0$ where * denotes the direct product operation (Searle, 1966). Similarly,

$$\underline{R} = \begin{bmatrix} \underline{R}_0 & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{R}_0 & \cdots & \underline{0} \\ \vdots & \vdots & & \vdots \\ \underline{0} & \underline{0} & \cdots & \underline{R}_0 \end{bmatrix} = \underline{I}*\underline{R}_0 .$$

Now, because of the form of $\underline{G}$ and $\underline{R}$, these can be inverted easily;

$$\underline{G}^{-1} = \underline{A}^{-1}*\underline{G}_0^{-1}$$

where $\underline{A}^{-1}$ is easy to compute by the rapid methods cited above and $\underline{G}_0^{-1}$ is a small matrix. Similarly,

$$\underline{R}^{-1} = \underline{I}*\underline{R}_0^{-1} = \begin{bmatrix} \underline{R}_0^{-1} & & 0 \\ & \ddots & \\ 0 & & \underline{R}_0^{-1} \end{bmatrix} .$$

Then, the BLUP equations can be written, realizing that because $\underline{Z} = \underline{I}$,

$$\underline{X}'\underline{Z} = \underline{X}' ;$$
$$\underline{Z}'\underline{Z} = \underline{I} ;$$
$$\underline{Z}'\underline{y} = \underline{y} .$$

A more realistic situation is one in which certain of the animals to be evaluated have one or more missing records. In fact, some may have no records. In that case, $\underline{Z} \neq \underline{I}$ but rather is written by deleting from an $nq \times nq$ identity matrix all rows pertaining to missing records. For example, if there are three animals and two traits of interest and the first animal has records on both traits, the second has a record on only the first trait, and the third animal has no records,

$$\underline{Z} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} .$$

Also, $\underline{R}$ is modified by deleting both rows and columns pertaining to missing records. In the above example,

$$\underline{R} = \begin{pmatrix} r_{11} & r_{12} & 0 \\ r_{12} & r_{22} & 0 \\ 0 & 0 & r_{11} \end{pmatrix}$$

where r_{ij} are elements of $\underline{R}_0$. $\underline{G}$ and $\underline{a}$ are the same as in the full information case, and consequently all traits on all animals can be predicted.

6.5. Evaluation from Progeny with One Record Each

In this section, we are concerned with the evaluation of animals using only progeny records. Each progeny has only one record used in the analysis. The vector of progeny records, $\underline{y}$, has the model

$$\underline{y} = \underline{X\beta} + \underline{Zs} + \underline{e} .$$

The most commonly assumed elements in $\underline{\beta}$ by those evaluating dairy sires in artificial insemination programs are herd-year-season effects and group effects (generation and stud, usually). $\underline{s}$ refers to a vector of sire values, with possibly some of the sires having no progeny. $\underline{Z}$ relates records to sires. For example, suppose there are three sires, the first with 4 progeny, the second with 2, and the third with none. Then, if records are ordered within sire,

$$\underline{Z} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} .$$

The assumptions about $\underline{s}$ and $\underline{e}$ are as follows in addition to the assumption of null means:

$$\mathrm{Var}(\underline{s}) = \tfrac{1}{4}h^2\underline{A}\sigma_y^2 ;$$

$$\mathrm{Var}(\underline{e}) = (1-\tfrac{1}{4}h^2)\underline{I}\sigma_y^2 ;$$

$$\mathrm{Cov}(\underline{s},\underline{e}') = \underline{0} .$$

Then the BLUP equations are

$$\begin{pmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z}+\underline{A}^{-1}\frac{4-h^2}{h^2} \end{pmatrix} \begin{pmatrix} \hat{\underline{\beta}} \\ \hat{\underline{s}} \end{pmatrix} = \begin{pmatrix} \underline{X}'\underline{y} \\ \underline{Z}'\underline{y} \end{pmatrix} .$$

If all sires are assumed to be members of the same population, the relative merit of their future progeny is predicted by differences in the elements of $\hat{\underline{s}}$. If groups are in the model, the prediction of the ith sire is $\hat{s}_i$ + an element of $\hat{\underline{\beta}}$ corresponding to the group in which the sire falls.

6.6. Evaluation from Progeny with Multiple Records

The problem in this section is the same as in 6.5 except that some or all of the progeny have more than one record. Now we write the model for the records as

$$\underline{y} = \underline{X\beta} + \underline{Z}_s\underline{s} + \underline{Z}_p\underline{p} + \underline{e}$$

where $\underline{p}$ refers to individual progeny. To illustrate $\underline{Z}_s$ and $\underline{Z}_p$, suppose there are three sires. The first has 3 progeny with 3, 2, and 1 records, respectively. The second sire has 2 progeny with 2 and 1 records, respectively, and the third sire has no progeny. Then, if records are ordered progeny within sires,

$$\underline{Z}_s = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix};$$

$$\underline{Z}_p = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

$\underline{s}$, $\underline{p}$, and $\underline{e}$ are random vectors with null means.

$$\mathrm{Var}(\underline{s}) = \tfrac{1}{4}h^2\underline{A}\sigma_y^2 ;$$

$$\mathrm{Var}(\underline{p}) = (r-\tfrac{1}{4}h^2)\underline{I}\sigma_y^2 ;$$

$$\mathrm{Var}(\underline{e}) = (1-r)\underline{I}\sigma_y^2 .$$

Cov(s,p'), Cov(s,e'), and Cov(p,e') are all null. This, of course, is an additive genetic model, the mates of sires are assumed unrelated, and the usual simple repeatability model is assumed.

Then, in a noninbred population, it is seen that the variance of a record is $(\frac{1}{4}h^2 + r - \frac{1}{4}h^2 + 1 - r)\sigma_y^2 = \sigma_y^2$. The covariance between two records on the same progeny is $(\frac{1}{4}h^2 + r - \frac{1}{4}h^2)\sigma_y^2 = r\sigma_y^2$. The covariance between half-sibs is $\frac{1}{4}h^2\sigma_y^2$. The covariance between a progeny of the ith sire and the progeny of the jth sire is $\frac{1}{4}h^2 a_{ij}\sigma_y^2$.

Then the BLUP equations are

$$\begin{pmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z}_s & \underline{X}'\underline{Z}_p \\ \underline{Z}_s'\underline{X} & \underline{Z}_s'\underline{Z}_s + \frac{4(1-r)}{h^2}\underline{A}^{-1} & \underline{Z}_s'\underline{Z}_p \\ \underline{Z}_p'\underline{X} & \underline{Z}_p'\underline{Z}_s & \underline{Z}_p'\underline{Z}_p + \frac{4(1-r)}{4r-h^2}\underline{I} \end{pmatrix} \begin{pmatrix} \hat{\underline{\beta}} \\ \hat{\underline{s}} \\ \hat{\underline{p}} \end{pmatrix} = \begin{pmatrix} \underline{X}'\underline{y} \\ \underline{Z}_s'\underline{y} \\ \underline{Z}_p'\underline{y} \end{pmatrix}.$$

$\underline{Z}_s'\underline{Z}_s$ is simply a diagonal matrix with a total number of progeny records for a sire in the diagonal. Similarly, $\underline{Z}_p'\underline{Z}_p$ is a diagonal matrix with the diagonal elements being the number of records on an individual progeny. $\underline{Z}_s'\underline{Z}_p$ contains in each row the number of records on the individual progeny of the sire and zeros. $\underline{Z}_s'\underline{y}$ and $\underline{Z}_p'\underline{y}$ are totals of records by sire and progeny, respectively.

6.7. Multiple Trait Evaluation by Progeny Tests

If we wish to evaluate sires with respect to several correlated traits from records on their progeny, a combination of the methods of Sections 6.4 and 6.5 can be used. Details are presented by Henderson (1976c).

6.8. Nonlinear Merit Function

There are some practical situations in which the net merit of an individual is a nonlinear function of certain traits (see, for example, Wilton et al., 1968). As stated in Section 2, best prediction would be the logical method to use in this case, but it is unlikely that the required distribution and its parameters are known. A feasible method is available if $(\underline{y},\underline{u})$ are jointly normally distributed, if the first and second moments of the distribution are known, if the merit function is a polynomial in $\underline{u}$, and if we are willing to accept a prediction that is a polynomial in $\underline{y}$. In this case one can describe the problem as one of best linear prediction where the data vector available for prediction contains elements that are products and powers of the individual elements of $\underline{y}$.

For example, suppose that the merit function is

$$g = u_1 + 2u_2 + .5u_1^2 + .6u_1u_2 + .8u_2^2 .$$

Suppose that $\underline{y} = [y_1 \quad y_2]$ and we choose $\hat{g}$ of the form $\hat{g} = E(g) + b_1[y_1 - E(y_1)] + b_2[y_2 - E(y_2)] + b_3[y_1^2 - E(y_1^2)] + b_4[y_1y_2 - E(y_1y_2)] + b_5[y_2^2 - E(y_2^2)]$. Then the prediction of this form that minimizes prediction error variance requires for its computation the parameters

$$E[u_1 \quad u_2 \quad u_1^2 \quad u_1u_2 \quad u_2^2 \quad y_1 \quad y_2 \quad y_1^2 \quad y_1y_2 \quad y_2^2] \;;$$
$$E[y_1 \quad y_2 \quad y_1^2 \quad y_1y_2 \quad y_2^2]\,[y_1 \quad y_2 \quad y_1^2 \quad y_1y_2 \quad y_2^2]' \;;$$
$$E[y_1 \quad y_2 \quad y_1^2 \quad y_1y_2 \quad y_2^2]\,[u_1 \quad u_2 \quad u_1^2 \quad u_1u_2 \quad u_2^2]' .$$

Fortunately, the expectations of products of normally distributed random variables can be written as functions of the means and of the variance-covariance matrix of variables comprising the product. These can be derived from the moment generating function. Let the vector $\underline{w}$ have a multivariate normal distribution with mean $\underline{\mu}$ and variance-covariance matrix $\underline{V}$. Then the moment generating function from Mood (1950) is

$$\exp(\underline{t}'\underline{\mu} + \underline{t}'\underline{V}\underline{t}) .$$

Using this function, some results are

$$E(w_1w_2) = \mu_1\mu_2 + v_{12} \;;$$
$$E(w_1w_2w_3) = \mu_1\mu_2\mu_3 + \mu_1v_{23} + \mu_2v_{13} + \mu_3v_{12} \;;$$
$$\begin{aligned} E(w_1w_2w_3w_4) = {} & \mu_1\mu_2\mu_3\mu_4 \\ & + \mu_1\mu_2v_{34} + \text{all other terms of this type} \\ & + v_{12}v_{34} + \text{all other terms of this type} \;; \end{aligned}$$
$$\begin{aligned} E(w_1w_2w_3w_4w_5) = {} & \mu_1\mu_2\mu_3\mu_4\mu_5 \\ & + \mu_1\mu_2\mu_3v_{45} + \text{all other terms of this type} \\ & + \mu_1v_{23}v_{45} + \text{all other terms of this type} \;; \end{aligned}$$
$$\begin{aligned} E(w_1w_2w_3w_4w_5w_6) = {} & \mu_1\mu_2\mu_3\mu_4\mu_5\mu_6 \\ & + \mu_1\mu_2\mu_3\mu_4v_{56} + \text{other terms of this type} \\ & + \mu_1\mu_2v_{34}v_{56} + \text{other terms of this type} \\ & + v_{12}v_{34}v_{56} + \text{other terms of this type} . \end{aligned}$$

6.9. Restricted Best Linear Unbiased Prediction

Kempthorne and Nordskog (1959) solved the problem of constructing an index to select individuals such that the expected value of a set of linear functions of additive merits of several traits is null in the individuals selected for some defined merit function. Quaas and Henderson (1976) have generalized this method to solve the same problem but with the means of records unknown and using records on relatives. They call this method "restricted best linear unbiased prediction."

Let $\underline{u}_i$ be the subvector of additive genetic values on several traits on the ith animal. We want to predict $\underline{m}'\underline{u}_i$ so that $\underline{C}'\underline{u}_i$ is uncorrelated with the predictor $\underline{m}'\hat{\underline{u}}_i$, so that $\underline{m}'\hat{\underline{u}}_i$ has expectation zero, and with these 2 restrictions has minimum prediction error variance. The method for doing this is found to be identical to the method of Section 6.4 except that in place of $\underline{R}_0^{-1}$ a block diagonal matrix is used that is computed from $\underline{R}_0$, $\underline{G}_0$, $\underline{C}$, and $\underline{Z}$.

6.10. Prediction of the Merit of Single Crosses

We assume that a random set of lines has been derived from the same original population, that these lines all have the same inbreeding coefficients, and a random set of progeny is obtained from single crosses among these lines. The problem is to predict the merit of a single cross (including those on which there are no data). It is assumed that the "pure" line data will not be used. Let the model for an observation on the cross between the ith line used as the male line and the jth line used as the female line be

$$y_{ijk} = \underline{x}'_{ijk}\underline{\beta} + s_i + d_j + r_{ij} + e_{ijk}$$

where $\underline{x}'_{ijk}\underline{\beta}$ defines the fixed effects in the model.

$$\begin{aligned} \mathrm{Var}(s_i) &= \sigma_s^2 \;; \\ \mathrm{Var}(d_j) &= \sigma_d^2 \;; \\ \mathrm{Cov}(s_i, d_i) &= \sigma_{sd} \;; \\ \mathrm{Var}(r_{ij}) &= \sigma_r^2 \;; \\ \mathrm{Cov}(r_{ij}, r_{ji}) &= \sigma_{rr'} \text{ for } i \neq j \;; \\ \mathrm{Var}(e_{ijk}) &= \sigma_e^2 . \end{aligned}$$

All other covariances = 0.

Then the BLUP equations for predicting $\underline{s}$, $\underline{d}$, and $\underline{r}$ are written by first writing least squares equations as though $\underline{s}$, $\underline{d}$, and $\underline{r}$ are fixed. These equations have the form

$$\begin{pmatrix} \underline{C}_{11} & \underline{C}_{12} & \underline{C}_{13} & \underline{C}_{14} \\ \underline{C}'_{12} & \underline{C}_{22} & \underline{C}_{23} & \underline{C}_{24} \\ \underline{C}'_{13} & \underline{C}'_{23} & \underline{C}_{33} & \underline{C}_{34} \\ \underline{C}'_{14} & \underline{C}'_{24} & \underline{C}'_{34} & \underline{C}_{44} \end{pmatrix} \begin{pmatrix} \hat{\underline{g}} \\ \hat{\underline{s}} \\ \hat{\underline{d}} \\ \hat{\underline{r}} \end{pmatrix} = \begin{pmatrix} \underline{t}_1 \\ \underline{t}_2 \\ \underline{t}_3 \\ \underline{t}_4 \end{pmatrix}.$$

Now modify these equations as follows:

to $\begin{pmatrix} \underline{C}_{22} & \underline{C}_{23} \\ \underline{C}'_{23} & \underline{C}_{33} \end{pmatrix}$, add $\begin{pmatrix} \underline{I}\sigma_s^2 & \underline{I}\sigma_{sd} \\ \underline{I}\sigma_{sd} & \underline{I}\sigma_d^2 \end{pmatrix}^{-1} \sigma_e^2$;

to $\underline{C}_{44}$, add $\underline{V}^{-1}$ where $\underline{V}$ is a matrix with σ_r^2 in the diagonals and $\sigma_{rr'}$ in the positions corresponding to r_{ij} with r_{ji}; all other elements of $\underline{V}$ are zero.

From the solution to these equations, the prediction of the merit of the cross of line of sire i by line of dam j is

$$\hat{s}_i + \hat{d}_j + \hat{r}_{ij} .$$

The prediction of the merit of the ith line of sire in crosses with a random sample of dams from other random lines of the population is $\hat{s}_i$ and similarly for the prediction of the merit of the jth line of dam.

BIBLIOGRAPHY

[1] Cochran, W.G. (1951). Improvement by selection. *Proc. 2nd Berkeley Symp. Math. Statist. and Prob.*, 449.

[2] Cockerham, C.C. (1963). Estimation of genetic variance. *Statistical Genetics and Plant Breeding* (W.D. Hanson & H.F. Robinson, Eds.). Washington: National Academy of Sciences--National Research Council, 53-93.

[3] Henderson, C.R. (1950). Estimation of genetic parameters. *Biom. 6*, 186.

[4] Henderson, C.R. (1953). Estimation of variance and covariance components. *Biom. 9*, 226.

[5] Henderson, C.R. (1963). Selection index and expected genetic advance. *Statistical Genetics and Plant Breeding* (W.D. Hanson & H.F. Robinson, Eds.). Washington: National Academy of Sciences--National Research Council, 141-63.

[6] Henderson, C.R. (1973). Sire evaluation and genetic trends. *Proc. of the Anim. Breeding and Genet. Symp. in Honor of Dr. Jay L. Lush.* July 29, 1972, Blacksburg, Va.

[7] Henderson, C.R. (1975a). Rapid method for computing the inverse of a relationship matrix. *J. Dairy Sci. 58*, 1727.

[8] Henderson, C.R. (1975b). Inverse of a matrix of relationships due to sires and maternal grandsires. *J. Dairy Sci. 58*, 1917.

[9] Henderson, C.R. (1975c). Best linear unbiased estimation and prediction under a selection model. *Biom. 31*, 423.

[10] Henderson, C.R. (1976a). A simple method for computing the inverse of a numerator relationship matrix used for prediction of breeding values. *Biom. 32*, 69.

[11] Henderson, C.R. (1976b). Inverse of a matrix of relationships due to sire and maternal grandsires in an inbred population. *J. Dairy Sci.*, in press.

[12] Henderson, C.R. (1976c). Multiple trait sire evaluation using the relationship matrix. *J. Dairy Sci. 59*, 769.

[13] Henderson, C.R., Kempthorne, O., Searle, S.R., & von Krosigk, C.M. (1959). The estimation of environmental and genetic trends from records subject to culling. *Biom. 15*, 192.

[14] Kempthorne, O. & Nordskog, A.W. (1959). Restriced selection indices. *Biometrics 15*, 10.

[15] Mood, A.M. (1950). *Introduction to the Theory of Statistics*. New York: McGraw-Hill.

[16] Quaas, R.L. (1976). Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics*, in press.

[17] Quaas, R.L. & Henderson, C.R. (1976). Restricted best linear unbiased prediction of breeding values. *Biometrics*, submitted.

[18] Searle, S.R. (1966). *Matrix Algebra for the Biological Sciences (Including Applications in Statistics)*. New York: John Wiley and Sons, Inc.

[19] Searle, S.R. (1973). Derivation of prediction formulae. BU-482-M, Biometrics Unit, Cornell Univ., Ithaca, New York.

[20] Solomon, D.L. (1971). A Bayesian interpretation of the genetic selection index. BU-362-M, Biometrics Unit, Cornell Univ., Ithaca, New York.

[21] Wilton, J.W., Evans, D.A., & VanVleck, L.D. (1968). Selection indices for quadratic models of total merit. *Biometrics 24*, 937.