Computing numerator relationships between any pair of animals

L.D. Van Vleck

Roman L. Hruska U.S. Meat Animal Research Center, ARS-USDA, Lincoln, NE, USA

Corresponding author: L.D. Van Vleck
E-mail: lvanvleck@unlnotes.unl.edu

ABSTRACT. We describe a simple method to compute the numerator relationship between any or all pairs of animals in the numerator relationship matrix. The method depends on output of the MTDFNRM program from the MTDFREML set of programs. An option of the MTDFNRM program creates a file that includes the inbreeding coefficient for each animal. The method also makes use of how the inbreeding coefficient is traditionally calculated: one-half of the relationship between the animal's parents. To obtain the numerator relationship between any pair of animals, the original pedigree file is augmented with a dummy animal with an identification number (ID) greater than for any animal in the original pedigree file. The ID of the pair of animals for which the relationship is wanted is included as parents. MTDFNRM is then run with the option to create a file of ordered and original IDs for animals and their parents along with the inbreeding coefficients. We then multiply the inbreeding coefficient for a dummy animal by two to obtain the numerator relationship between the two animals designated as parents.

Key words: Inbreeding coefficient, Pairs of animals, Numerator relationships

INTRODUCTION

The history of computing relationships between pairs of animals traces to Wright (1918, 1922, 1934) through the method of path coefficients, which was used for 30 to 40 years.

Cruden (1949) and Emik and Terrill (1949) developed recursive methods for calculating numerator relationships with a tabular method where the entry for column j of row i is one-half the sum of entries for the sire and dam of animal j in the same row (i). The diagonal element for animal i is 1 + F(i), where F(i) is the inbreeding coefficient which is equal to one-half the additive relationship between parents of animal i, which appears earlier in the table.

Both the path coefficients of Wright and the tabular method can create the matrix of numerator relationships, A. However, use of the mixed model equations (Henderson et al., 1959; Henderson, 1949, 1963, 1984) to predict genetic values requires the inverse of A. Henderson (1975, 1976) developed simple rules for calculating the elements of the inverse of A from a list of animals with corresponding sires and dams. Quaas (1976) expanded the rules of Henderson to include inbreeding coefficients. These rules are now used for nearly all genetic evaluations for most classes of livestock.

Sometimes the numerator relationship of a pair of animals is desired, rather than the corresponding inverse element of the inverse of A (Ferraz B and Hodges J, personal communication, 2007).

We will describe a method that uses the output of MTDFNRM, one of the three programs in the MTDFREML package (Boldman et al., 1995), which is based on the algorithm of Quaas (1976) along with traditional calculation of the inbreeding coefficient as noted in the last sentence of his epic paper.

METHODS

The MTDFNRM program calculates inverse elements of A, for use with the mixed model equations to account for partial genetic replication. The method to calculate the relationship between a pair of animals depends on the file MTDF13, which is an option for MTDFNRM. The MTDF13 file includes recoded and original animal, sire, and dam identification numbers and also the inbreeding coefficients of each animal and its sire and dam.

The method to calculate the numerator relationship matrix from the inbreeding coefficient (or for any pair of animals) depends on the correspondence of F of an animal and one-half the relationship between its sire and dam. The F calculated by MTDFNRM is used to determine the additive relationship between any pair of animals in the original A matrix. Dummy animals are added to the pedigree file with their sires and dams being the pairs of animals for which the numerator relationship is wanted. Once the inbreeding coefficient is known, the numerator relationship between the parents is automatically known:

F(i) = 0.5 (relationship of sire of i with dam of i). Multiplying by 2 gives the relationship between the sire of i and the dam of i. This process is reversed to calculate the relationship between the "parents" of animal (i). The "sire" and "dam" can both be males or both be females, or male and female, or female and male. In other words, animal (i) can be a dummy animal, but if any pair of animals in the original pedigree file is given as its parents, the inbreeding coefficient will be calculated for animal (i). In that way, the numerator relationship between any pair of animals can be found. One adds a dummy animal with an identification number (ID) to

the pedigree file and specifies the pair of animals as parents of the dummy animal. The dummy animal should not be in the correct pedigree file but would be entered as a dummy animal ID, with the ID for the pair of animals for which the relationship is wanted. The dummy animal ID must be greater than those of animals in the original pedigree file, which can be recoded to 1, 2, 3,…, N by MTDFNRM from the original pedigree file and written to MTDF13. For example, the relationship between animals 2 and 10 is wanted. Specify an ID for a dummy animal not in the original pedigree file, such as 5001. Then add to the pedigree file: 5001 2 10 with dummy animal ID and ID for the pair of real animals. Then $F(5001) = 0.5 [a(2,10)]$, where $a(2,10)$ is the relationship between animal (2) and animal (10).

Dummy animals with a larger ID than animals in the original pedigree file and the sire and dam ID from the original pedigree file can be added to the end of original pedigree file. That procedure works well for a few or several pairs. Then the MTDFNRM program is run with the augmented pedigree file.

Theoretically, the full numerator relationship matrix can be obtained this way. For example, for row 1 with five columns, four dummy animals will be required, with parents being the combinations of animal 1 with animals 2, 3, 4, 5.

1002	1	2
1003	1	3
1004	1	4
1005	1	5

Row 2 would be similar, but with only the upper half calculated:

2003	2	3
2004	2	4
2005	2	5

This numbering system shows the ID of parents in the ID for the dummy animal.

The other rows would be done similarly. For example for $F(2004)$, $2\{.5 [a(2,4)]\}$ is the 2,4 element of the relationship matrix.

To add to the original pedigree file, dummy animals with "parents" being all pairs of animals in the A matrix can easily be programmed. This method may be time consuming for a large pedigree, but it needs only two DO loops in Fortran or FOR loops in MATLAB to create the file to augment the original pedigree file. A sample Fortran program to create a file combining the original pedigree file and the dummy animals is in Appendix. This may create a very large file for a large number of animals, as there will be $n(n - 1) / 2$ half stored pairs of animals in the numerator relationship matrix (A) for the original pedigree.

The program reads and writes the original pedigree file (animal, sire, dam) to begin the combined data file and then creates and writes to the end of the new file, large dummy IDs with "parents" being all pairs of animals in the original pedigree file. The last step is to run MTDFNRM with the combined file. The F for animals in the original pedigree file will be calculated and, more importantly, the F for the dummy animals will be calculated for all pairs of animals in the original pedigree file and will be written to MTDF13. The "parents" of the dummy animals can both be males or females, or a male and a female, as stated earlier. The dummy ID might be created so

the recoded "sire" and "dam" ID can be read easily, for example i00000j, where i and j are the recoded ID in the original pedigree. A run of MTDFNRM with the original pedigree will write the MTDF13 file that will match original ID with the recoded ID. This MTDF13 file should be copied to another file such as MTDF13.PED because the second run of MTDFNRM will write a new MTDF13 that includes the dummy animals.

Then one needs to double the calculated "F" to obtain the additive relationship between the "parents" of the dummy animal.

The three keys are:

1) Add dummy animals with "parents", the pairs of animals for which relationships are wanted to the original pedigree file.

2) Calculate the wanted relationship from the F for a dummy animal with pairs of animals as "parents" for which the relationship is wanted with MTDFNRM and the augmented pedigree file.

3) Multiply F by 2 to obtain the numerator relationship between the pair of animals.

In the example, both 2 and 5 can be females (or males). The algorithm does not work for clones, e.g., 2 2 2, as animal and pair of parents. The algorithm does work for the parents being the same animal, e.g., 5005 5 5.

The "F" comes out in MTDF13, when the option to write the original and recoded ID (1, 2, 3….) is chosen. The F for the dummy animal and its "sire" and "dam" are in the last three columns.

Obviously this method can also be used to calculate inbreeding coefficients for potential progeny of mating of a specified sire and dam. In that case, the dummy animal would represent the potential progeny. For example, a set of sires is available for mating with females in a herd or a subset of females in the herd, such as females that have not produced progeny. The ID of each of the set of sires would be matched with the ID of all of the potential dams, with corresponding dummy progeny, and would be added to the original pedigree file, including the potential sires and dams, as for example,

$$
\begin{array}{ccc}
100101 & 1 & 101 \\
100102 & 1 & 102 \\
200101 & 2 & 101 \\
200102 & 2 & 102
\end{array}
$$

where 1 and 2 are potential sires and 101 and 102 are possible mates.

In this case, the inbreeding coefficient of the potential progeny is wanted rather than the relationship between the two animals specified to be parents of a dummy animal.

REFERENCES

Boldman KG, Kriese LA, Van Vleck LD, Van Tassell CP, et al. (1995). A manual for the use of MTDFREML. A set of programs to obtain estimates of variances and covariances [Draft]. ARS, USDA, USMARC, Clay Center, Lincoln.
Cruden D (1949). The computation of inbreeding coefficients for closed populations. *J. Hered.* 40: 248-251.
Emik LO and Terrill CE (1949). Systematic procedures for calculating inbreeding coefficients. *J. Hered.* 40: 51-55.
Henderson CR (1949). Estimation of changes in herd environment. *J. Dairy Sci.* 32: 706.
Henderson CR (1963). Selection index and expected genetic advance. In: Statistical genetics and plant breeding, Publication 982 (Hanson WD and Robison HF, eds.). National Academy of Sciences, Washington, 141-163.
Henderson CR (1975). A rapid method for computing the inverse of a relationship matrix. *J. Dairy Sci.* 58: 1727-1730.

Henderson CR (1976). A simple way for calculating the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32: 69-83.

Henderson CR (1984). Applications of linear models in animal breeding. University of Guelph, Guelph.

Henderson CR, Kempthorne O, Searle SR and Von Krosigk CM (1959). The estimation of environmental and genetic traits from records subject to culling. *Biometrics* 15: 192-218.

Quaas RL (1976). Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics* 32: 949-953.

Wright S (1918). On the nature of size factors. *Genetics* 3: 367-374.

Wright S (1922). Coefficients of inbreeding and relationship. *Am. Nat.* 56: 330-338.

Wright S (1934). The methods of path coefficients. *Ann. Math. Stat.* 5: 161-215.

APPENDIX

Fortran code to create a pedigree file of real and dummy animals to use with MTDFNRM to calculate the numerator relationship between all pairs of real animals

```
c        combine real ped file and make and add dummy ped for all original pairs
c        combine.for
         character*25 filenam

         open(104,file='comped.dat')      ! combined ped file
         write(6,*) 'Enter number of lines in original pedigree file'
         read(5,*) n1
         write(6,*) ' '

         write(6,*) 'Enter name of original pedigree file'
         read(5,509) filenam
509      format(A)

         open(103,file=filenam)            !filenam will contain original pedigree
         write(6,*) ' '
         write(6,*) ' '
         write(6,*) 'combined file will be comped.dat'

99       continue
         read(103,*,end=98) ida,ids,idd    !read an sire dam list from real ped
         write(104,*) ida,ids,idd ! writes original file to comped.dat
         go to 99                          !go to read another line

98       continue                          !end of real ped file

c        create dummy animals with all pairs of animals in original ped file
         do i=1,n1-1
         do j=i+1,n1
         id=10000*i+j               ! dummy animal ID will show pair i and j
c                                   ! limit as set is 9999 original animals
c                                   ! no real animal can have ID > 9999
c                                   ! could run mtdfnrm with original ped file
```

```
c                                  ! and then use the mtdf13 file that is a
c                                  ! recoded and ordered pedigree file
c                                  ! as the ordered original pedigree file
        write(104,*)id,i,j ! writes dummy animal to comped.dat
       enddo
       enddo

       close(101)
       close(103)
       close(104)
       stop
       end
```