Livestock Breeding and Genomics - Solution 9

Peter von Rohr

2018-11-19

Problem 1: Decompositions

Given is the following pedigree.

Animal	Sire	Dam
1	NA	NA
2	NA	NA
3	1	2
4	1	NA
5	3	4
6	5	4

Your Tasks

- Do the simple decomposition of the above pedigree
- Do the recursive decomposition of the above pedigree until only m_i terms appear on the right-hand side of the decomposition.

Solution

- **Simple Decomposition**: For the simple decomposition, the true breeding values are decomposed into true breeding values of parents plus the respective mendelian sampling effect. For the pedigree given above this is

$$
\begin{aligned}
a_1 &= m_1 \\
a_2 &= m_2 \\
a_3 &= \frac{1}{2}a_1 + \frac{1}{2}a_2 + m_3 \\
a_4 &= \frac{1}{2}a_1 + m_4 \\
a_5 &= \frac{1}{2}a_3 + \frac{1}{2}a_4 + m_5 \\
a_6 &= \frac{1}{2}a_5 + \frac{1}{2}a_4 + m_6
\end{aligned}
$$

Converting the same decomposition into matrix-vector notation, we get

$$a = P \cdot a + m$$

Putting the information from the pedigree into the decomposition yields

$$
\left[\begin{array}{c} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{array}\right] = \left[\begin{array}{cccccc} 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.50 & 0.50 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.50 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.50 & 0.50 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.50 & 0.50 & 0.00 \end{array}\right] \cdot \left[\begin{array}{c} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{array}\right] + \left[\begin{array}{c} m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \\ m_6 \end{array}\right]
$$

- **Recursive Decomposition**: The recursive decomposition repeats simple decompositions of true breeding values of ancestors until the right-hand side of the decomposition consists only of mendelian sampling terms. In matrix-vector notation, the recursive decomposition can be written as

$$
a = L \cdot m
$$

The vectors a and m are defined as for the simple decomposition. The matrix L has the following structure.

$$
\left[\begin{array}{c} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{array}\right] = \left[\begin{array}{cccccc} 1.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.500 & 0.500 & 1.000 & 0.000 & 0.000 & 0.000 \\ 0.500 & 0.000 & 0.000 & 1.000 & 0.000 & 0.000 \\ 0.500 & 0.250 & 0.500 & 0.500 & 1.000 & 0.000 \\ 0.500 & 0.125 & 0.250 & 0.750 & 0.500 & 1.000 \end{array}\right] \cdot \left[\begin{array}{c} m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \\ m_6 \end{array}\right]
$$

Problem 2: Covariance Between Animals

So far we have assumed that the covariance between true breeding values a_i and a_j of animals i and j is defined as

$$
cov(a_i, a_j) = (A)_{ij} * \sigma_a^2
$$

where	$(A)_{ij}$	element of numerator relationship matrix A on row i and column j
	σ_a^2	additive genetic variance

Your Task

- Compute the covariance $cov(a_i, a_j)$ between true breeding values a_i and a_j for all pairs of animals in the pedigree given in Problem 1.
- The following procedure can be used to compute the covariances
  - Recursively decompose a_i and a_j until common ancestors between i and j are found or until unrelated founder animals occur in the decomposition
  - Expand the covariance between sums of true breeding values into sums of variances of true breeding values
  - Collect the results of all pairwise covariances and compare them to the entries in A

Hints

- The covariance $cov(a_i, a_j)$ between true breeding values a_i and a_j for unrelated animals i and j is set to 0. An example in the pedigree from Problem 1 would be the covarance $cov(a_1, a_2)$ between the true breeding values of animals 1 and 2 which is equal to 0, because animals 1 and 2 are unrelated, that means they do not have any known common ancestors.
- The covariances $cov(a_i, m_k)$ between any true breeding value a_i and any mendelian sampling constant m_k is 0.

- Any previously compute covariance can be re-used.
- The variances $var(a_i)$ of the true breeding values are $var(a_i) = (1 + F_i)\sigma_a^2$ where F_i is the inbreeding coefficient of animal i.
- The inbreeding coefficients for all animals in pedigree of Problem 1 are

Animal	Inbreeding Coefficient
1	0.0000
2	0.0000
3	0.0000
4	0.0000
5	0.1250
6	0.3125

- As an example, it is shown how to compute $cov(a_3, a_4)$. First replace a_3 and a_4 by their simple decomposition into true breeding values of parents plus mendelian sampling terms. Then expand the covariance of two sums into sums of pairwise covariances. Then known values are inserted and the result is computed.

$$
\begin{aligned}
cov(a_3, a_4) &= cov((\frac{1}{2}a_1 + \frac{1}{2}a_2 + m_3), (\frac{1}{2}a_1 + m_4)) \\
&= cov(\frac{1}{2}a_1, \frac{1}{2}a_1) + cov(\frac{1}{2}a_2, \frac{1}{2}a_1) \\
&= \frac{1}{4}var(a_1) = \frac{1}{4}\sigma_a^2
\end{aligned}
$$

Solution

$$cov(a_1, a_2) = 0$$

$$
\begin{aligned}
cov(a_1, a_3) &= cov(a_1, \frac{1}{2}a_1 + \frac{1}{2}a_2 + m_3) \\
&= cov(a_1, \frac{1}{2}a_1) + cov(a_1, \frac{1}{2}a_2) \\
&= \frac{1}{2}cov(a_1, a_1) = \frac{1}{2}var(a_1) = \frac{1}{2}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_1, a_4) &= cov(a_1, \frac{1}{2}a_1 + m_4) \\
&= cov(a_1, \frac{1}{2}a_1) \\
&= \frac{1}{2}cov(a_1, a_1) = \frac{1}{2}var(a_1) = \frac{1}{2}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_1, a_5) &= cov(a_1, \frac{1}{2}a_3 + \frac{1}{2}a_4 + m_5) \\
&= cov(a_1, \frac{1}{2}(\frac{1}{2}a_1 + \frac{1}{2}a_2 + m_3) + \frac{1}{2}(\frac{1}{2}a_1 + m_4) + m_5) \\
&= cov(a_1, \frac{1}{4}a_1) + cov(a_1, \frac{1}{4}a_2 + cov(a_1, \frac{1}{4}a_1) \\
&= \frac{1}{4}cov(a_1, a_1) + \frac{1}{4}cov(a_1, a_1) = \frac{1}{2}var(a_1) = \frac{1}{2}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_1, a_6) &= cov(a_1, (\frac{1}{2}a_5 + \frac{1}{2}a_4 + m_6)) \\
&= cov(a_1, (\frac{1}{2}(\frac{1}{2}a_3 + \frac{1}{2}a_4 + m_5) + \frac{1}{2}(\frac{1}{2}a_1 + m_4) + m_6)) \\
&= cov(a_1, (\frac{1}{2}(\frac{1}{2}(\frac{1}{2}a_1 + \frac{1}{2}a_2 + m_3) + \frac{1}{2}(\frac{1}{2}a_1 + m_4) + m_5) + \frac{1}{2}(\frac{1}{2}a_1 + m_4) + m_6)) \\
&= cov(a_1, \frac{1}{8}a_1) + cov(a_1, \frac{1}{8}a_2) + cov(a_1, \frac{1}{8}a_1) + cov(a_1, \frac{1}{4}a_1) \\
&= \frac{1}{8}var(a_1) + \frac{1}{8}var(a_1) + \frac{1}{4}var(a_1) = \frac{1}{2}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_2, a_3) &= cov(a_2, \frac{1}{2}(a_1 + a_2 + m_3)) \\
&= cov(a_2, \frac{1}{2}a_1) + cov(a_2, \frac{1}{2}a_2) = \frac{1}{2}var(a_2) = \frac{1}{2}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_2, a_4) &= cov(a_2, (\frac{1}{2}a_1 + m_4) \\
&= cov(a_2, \frac{1}{2}a_1) = 0
\end{aligned}
$$

$$
\begin{aligned}
cov(a_2, a_5) &= cov(a_2, (\frac{1}{2}a_3 + \frac{1}{2}a_4 + m_5)) \\
&= cov(a_2, (\frac{1}{2}(\frac{1}{2}a_1 + \frac{1}{2}a_2 + m_3) + \frac{1}{2}(\frac{1}{2}a_1 + m_4) + m_5)) \\
&= cov(a_2, \frac{1}{4}a_1) + cov(a_2, \frac{1}{4}a_2) + cov(a_2, \frac{1}{4}a_1) \\
&= \frac{1}{4}var(a_2) = \frac{1}{4}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_2, a_6) &= cov(a_2, (\frac{1}{2}a_5 + \frac{1}{2}a_4 + m_6))) \\
&= cov(a_2, (\frac{1}{2}(\frac{1}{2}a_3 + \frac{1}{2}a_4 + m_5) + \frac{1}{2}(\frac{1}{2}a_1 + m_4) + m_6)) \\
&= cov(a_2, \frac{1}{4}a_3) + cov(a_2, \frac{1}{4}a_4) + cov(a_2, \frac{1}{4}a_1) \\
&= \frac{1}{4}cov(a_2, a_3) = \frac{1}{8}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_3, a_4) &= cov((\frac{1}{2}a_1 + \frac{1}{2}a_2 + m_3), (\frac{1}{2}a_1 + m_4)) \\
&= cov(\frac{1}{2}a_1, \frac{1}{2}a_1) + cov(\frac{1}{2}a_2, \frac{1}{2}a_1) \\
&= \frac{1}{4}var(a_1) = \frac{1}{4}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_3, a_5) &= cov(a_3, (\frac{1}{2}a_3 + \frac{1}{2}a_4 + m_5))) \\
&= cov(a_3, \frac{1}{2}a_3) + cov(a_3, \frac{1}{2}a_4) \\
&= \frac{1}{2}var(a_3) + \frac{1}{2}cov(a_3, a_4) = \frac{1}{2}\sigma_a^2 + \frac{1}{8}\sigma_a^2 = \frac{5}{8}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_3, a_6) &= cov(a_3, (\frac{1}{2}a_5 + \frac{1}{2}a_4 + m_6)) \\
&= \frac{1}{2}cov(a_3, a_5) + \frac{1}{2}cov(a_3, a_4) \\
&= \frac{5}{16}\sigma_a^2 + \frac{1}{8}\sigma_a^2 = \frac{7}{16}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_4, a_5) &= cov(a_4, (\frac{1}{2}a_3 + \frac{1}{2}a_4 + m_5)) \\
&= cov(a_4, \frac{1}{2}a_3) + cov(a_4, \frac{1}{2}a_4) \\
&= \frac{1}{2}cov(a_4, a_3) + \frac{1}{2}var(a_4) = \frac{1}{2} * \frac{1}{4}\sigma_a^2 + \frac{1}{2}\sigma_a^2 = \frac{5}{8}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_4, a_6) &= cov(a_4, (\frac{1}{2}a_5 + \frac{1}{2}a_4 + m_6))) \\
&= cov(a_4, \frac{1}{2}a_5) + cov(a_4, \frac{1}{2}a_4) \\
&= \frac{1}{2}cov(a_4, a_5) + \frac{1}{2}var(a_4) = \frac{1}{2} * \frac{5}{8}\sigma_a^2 + \frac{1}{2}\sigma_a^2 = \frac{13}{16}\sigma_a^2
\end{aligned}
$$

$$
\begin{aligned}
cov(a_5, a_6) &= cov(a_5, (\frac{1}{2}a_5 + \frac{1}{2}a_4 + m_6))) \\
&= \frac{1}{2}cov(a_5, a_5) + \frac{1}{2}cov(a_5, a_4) \\
&= \frac{1}{2}var(a_5) + \frac{1}{2} * \frac{5}{8}\sigma_a^2 = \frac{1}{2} * (1 + \frac{1}{8})\sigma_a^2 + \frac{5}{16}\sigma_a^2 = \frac{7}{8}\sigma_a^2
\end{aligned}
$$

The results of all the covariances is verified by the fact that

$$var(a) = A * \sigma_a^2$$

and the factors besides σ_a^2 are the same as found in the numerator relationship matrix A.

$$A = \begin{bmatrix} 1.0000 & 0.0000 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.0000 & 1.0000 & 0.5000 & 0.0000 & 0.2500 & 0.1250 \\ 0.5000 & 0.5000 & 1.0000 & 0.2500 & 0.6250 & 0.4375 \\ 0.5000 & 0.0000 & 0.2500 & 1.0000 & 0.6250 & 0.8125 \\ 0.5000 & 0.2500 & 0.6250 & 0.6250 & 1.1250 & 0.8750 \\ 0.5000 & 0.1250 & 0.4375 & 0.8125 & 0.8750 & 1.3125 \end{bmatrix}$$