Variances, covariances, matrices, genetic relationships

Ole F. Christensen

Center for Quantitative Genetics and Genomics, Aarhus University

BovReg course September 2023

Definition of variance and covariance

- Variance: $\texttt{Var}(x_1) = \texttt{E}[(x_1 - \texttt{E}[x_1])^2]$

- Covariance:
  $\texttt{Cov}(x_1, x_2) = \texttt{E}[(x_1 - \texttt{E}[x_1])(x_2 - \texttt{E}[x_2])]$.

- Variance is covariance with itself:
  $\texttt{Cov}(x_1, x_1) = \texttt{Var}(x_1)$.

- Correlation: $\texttt{Cor}(x_1, x_2) = \frac{\texttt{Cov}(x_1, x_2)}{\texttt{Var}(x_1)\texttt{Var}(x_2)}$.

Properties of variance

- $\texttt{Var}(x_1 + x_2) = \texttt{Var}(x_1) + \texttt{Var}(x_2) + 2\texttt{Cov}(x_1, x_2)$

- $\texttt{Var}(b_1 x_1) = b_1^2 \texttt{Var}(x_1)$.

- $\texttt{Var}(b_1 x_1 + b_2 x_2) =$
  $b_1^2 Var(x_1) + b_2^2 \texttt{Var}(x_2) + 2 b_1 b_2 \texttt{Cov}(x_1, x_2)$.

- Variance is always non-negative!

Properties of Covariance

- $\texttt{Cov}(x_1, x_2) = \texttt{Cov}(x_2, x_1)$
- $\texttt{Cov}(x_1 + x_3, x_2) = \texttt{Cov}(x_1, x_2) + \texttt{Cov}(x_3, x_2)$
- $\texttt{Cov}(a_1 x_1, x_2) = a_1 \texttt{Cov}(x_1, x_2)$.
- $\texttt{Cov}(a_1 x_1 + a_3 x_3, a_2 x_2 + a_4 x_4) =$
  $a_1 a_2 \texttt{Cov}(x_1, x_2) + a_1 a_4 \texttt{Cov}(x_1, x_4) +$
  $a_3 a_2 \texttt{Cov}(x_3, x_2) + a_3 a_4 \texttt{Cov}(x_3, x_4)$.
- Covariance is bi-linear

Variance-covariance matrix

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$\texttt{Var}(\mathbf{x}) = \begin{bmatrix} \texttt{Var}(x_1) & \texttt{Cov}(x_1, x_2) & \texttt{Cov}(x_1, x_3) & \texttt{Cov}(x_1, x_4) \\ \texttt{Cov}(x_2, x_1) & \texttt{Var}(x_2) & \texttt{Cov}(x_2, x_3) & \texttt{Cov}(x_2, x_4) \\ \texttt{Cov}(x_3, x_1) & \texttt{Cov}(x_3, x_2) & \texttt{Var}(x_2) & \texttt{Cov}(x_3, x_4) \\ \texttt{Cov}(x_4, x_1) & \texttt{Cov}(x_4, x_2) & \texttt{Cov}(x_4, x_3) & \texttt{Var}(x_4) \end{bmatrix}$$

- $\texttt{Var}(\mathbf{x})$ is symmetric and positive definite.

Are these valid variance-covariance matrices?

$$\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} ; \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} ; \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} ; \begin{bmatrix} 3 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 1 \end{bmatrix}$$

Property of Variance-covariance

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \end{bmatrix}$$

$$\mathbf{Bx} = \begin{bmatrix} b_{11}x_1 + b_{12}x_2 + b_{13}x_3 + b_{14}x_4 \\ b_{21}x_1 + b_{22}x_2 + b_{23}x_3 + b_{24}x_4 \end{bmatrix}$$

- Rule: $\texttt{Var}(\mathbf{Bx}) = \mathbf{B}\texttt{Var}(\mathbf{x})\mathbf{B}^T$
- Matrices and vectors: higher level of abstraction!

Genetic covariance between individuals

- Additive genetic covariances are important for breeding value estimation.
- Additive genetic covariance = additive genetic variance $\times$ additive genetic relationship

Additive genetic relationship (pedigree)

- Matrix **A**.
- Two times coancestry (probability of alleles being identical by descent)
- A recursive formula:
  - Base individuals: $a_{ij} = 0$, $a_{ii} = 1$
  - $a_{ij} = (a_{is} + a_{id})/2$
  - $a_{ii} = 1 + a_{sd}/2$
- Inverse $\mathbf{A}^{-1}$ can be computed directly (sparse)
- Additive genetic variance-covariance matrix: $\texttt{Var}(\mathbf{a}) = \sigma_a^2\mathbf{A}$