Learning Objectives

Lesson Map

- Review of Linear Algebra
    - Definitions and notation
    - Operations
    - Motivation

Review of Linear (Matrix) Algebra

A general linear model for a trait evaluated on a continuous scale, can be written:

$$Y_i = \mu + \varepsilon_i \quad (1.1),$$

where Y_i represents the observed (measured) trait value of a sample or experimental unit, i, μ is a parameter representing the true (mean) value for all experimental and sample units, and ε_i represents residual variability not accounted for by the mean value in the measured value of Y_i. For example, this model might be used to represent the yield of an area of land that has been divided into $i = 1, 2, 3 \ldots n$ smaller sampled areas.

If we were interested in whether the proportion of an introgressed genome affected yield of a cultivar planted in the plots, we might write the model as:

$$Y_{ij} = \beta_0 + \beta_1 G_i + \varepsilon_{ij} \quad (1.2).$$

In (1.2) the parameter β_o represents the intercept and β_1 represents the slope of a line that summarizes the proportion of the genome of interest. These two parameters are **un**known and need to be estimated based on data for Y_i and G_i. The latter indicates the proportion of the introgressed genome. It is a continuous and known (i.e., measured without error) variable, while the yield are measured for each plot (experimental unit).
If we were interested in whether different cultivars affect yield we might write the model as:

$$Y_{ij} = C_i + \varepsilon_{ij} \quad (1.3)$$

where Y_{ij} represents the observed (measured) trait value of a cultivar, i, evaluated in the experimental unit (plot j), C_i represents the genetic value of the cultivar and ε_{ij} represents residual variability not accounted for by the cultivar values in the measured value of Y_{ij}.

By organizing data using row x column data models and modeling phenotypes using general linear models are based on theory developed for linear (a.k.a. matrix) algebra. For this brief review we assume that all of the measured data for the variables, Yi, Gi, Ci in models 1.1, 1.2 and 1.3 are measured without error and the model for the phenotype is complete and accurate. Thus, all of the variability in Y_i is completely and correctly described by the parameters of the model. In other words, there is no need for a parameter ε to be included in models 1.1, 1.2 or 1.3 or any other general linear model. In statistics this assumption is relaxed.

All linear models can be represented in matrix form as $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta}$. As previously described, the variables in $\boldsymbol{y}$ and $\mathbf{X}$ are measured without error and the parameters in $\boldsymbol{\beta}$ are not known. Our motivation in this review is to find the unknown values for the parameters in $\boldsymbol{\beta}$.

Definitions and notation.

A *matrix* is a collection of numerical values arranged in rows and columns. Herein, the elements of a matrix are enclosed in brackets. For example, $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is a matrix with 4 elements arranged in two rows and two columns. Matrices with two or more rows and columns are denoted with upper case bold letters. *Vectors* are a special type of matrix with only one column, or one row. Vector matrices are denoted with lower case bold italicized letters.

For example, $\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$ or $\boldsymbol{y} = \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix}$.

A matrix consisting of only one row and one column is referred to as a *scalar*. A square matrix has the same number of rows and columns. A *diagonal matrix* is a *square matrix* with off-diagonal elements equal to 0. An *identity matrix* is a diagonal matrix with diagonal elements = 1. The identity matrix is almost always denoted **I**.

Matrix Operations. Matrices must be *conformable*, i.e., matrix operations have requirements on the numbers of rows and columns.

It is possible to add or subtract matrices, but only if they have the same numbers of rows and columns. For example,

$$\mathbf{C} = \mathbf{A} - \mathbf{B} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = \begin{pmatrix} a_{11} - b_{11} & a_{12} - b_{12} & a_{13} - b_{13} \\ a_{21} - b_{21} & a_{22} - b_{22} & a_{23} - b_{23} \\ a_{31} - b_{31} & a_{32} - b_{32} & a_{33} - b_{33} \end{pmatrix}.$$

The *transpose* of a matrix, denoted as $\mathbf{A}'$ (or $\mathbf{A}^t$ or $\mathbf{A}^T$) is a useful operation in which the first row of a matrix becomes the first column of its transpose, while the second, third, … etc rows become the second, third, … etc columns of its transpose. For example, let $\mathbf{A} = \begin{pmatrix} 2 & 8 & -1 \\ 3 & 6 & 4 \end{pmatrix}$, then $\mathbf{A}' = \begin{pmatrix} 2 & 3 \\ 8 & 6 \\ -1 & 4 \end{pmatrix}$.

TRY THIS M1: Let $\boldsymbol{x} = \begin{pmatrix} 2 \\ 8 \\ -1 \end{pmatrix}$ and $\boldsymbol{y} = \begin{pmatrix} 3 & 6 & 4 \end{pmatrix}$. If we obtain the transpose of one of these vectors what will be the result of adding them? What will be the result if we obtain the transpose of the other vector before adding them?

Scalar multiplication: It is possible to multiply a matrix by a scalar or a scalar by a matrix by simply multiplying all elements of the matrix by the scalar value, v. Thus $\mathbf{D} = v\mathbf{A} = \mathbf{A}v = \begin{pmatrix} va_{11} & va_{12} & va_{13} \\ va_{21} & va_{22} & va_{23} \\ va_{31} & va_{32} & va_{33} \end{pmatrix}$.

It is possible to multiply two vectors, but only if 1) one of the vectors is a row vector, 2) the second is a column vector, 3) the row vector has as many elements as the column vector. For example, if $\boldsymbol{v}$ = $\begin{pmatrix} 1 & 3 & 5 \end{pmatrix}$ and $\boldsymbol{w} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$, then $\boldsymbol{v} \bullet \boldsymbol{w}$ is a legal operation, whereas $\begin{pmatrix} 1 \\ 3 \end{pmatrix}\begin{pmatrix} 2 & 4 & 6 \end{pmatrix}$ is not. The operation of vector multiplication in the first instance indicates that we have a 1x3 vector multiplied by a 3x1 vector. The way we carry out the *vector multiplication* is to multiply the elements from each matrix in a pairwise manner, then sum the results of all pairs:

$\begin{pmatrix} 1 & 3 & 5 \end{pmatrix}\begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} = 1\text{x}2+3\text{x}4+5\text{x}6=44$. This operation is also known as the *dot product* and *inner product* of two vectors. If the dot product of two vectors is zero, then the two vectors are *perpendicular* to each other. The *length of a vector*, denoted $||\boldsymbol{v}||$ is the square root of the dot product of the vector with itself. $||\boldsymbol{v}|| = \sqrt{\boldsymbol{v} \bullet \boldsymbol{v}}$.

We could also apply the rule of multiplying and summing pairs of elements to the reverse arrangement $\boldsymbol{w} \bullet \boldsymbol{v} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} \begin{pmatrix} 1 & 3 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 6 & 10 \\ 4 & 12 & 20 \\ 6 & 18 & 30 \end{pmatrix}$. Notice that the order of arrangement of vectors matters.

Likewise, the arrangement of matrices that are to be multiplied matters.

It is also possible to multiply matrices, if we are careful. Virtually all types of matrix multiplication involve the multiplication of a row vector by a column vector. In essence we partition each matrix into a set of row and column vectors, then apply the rules of vector multiplication. Let's consider **C=AB.** $c_{ij} = \mathbf{a}_{i}.\mathbf{b}_{.j}$, where $\mathbf{a}_{i.}$ is the i^{th} row vector of **A** and $\mathbf{b}_{.j}$ is the j^{th} column vector of **B**.

For example,

$\mathbf{A} = \begin{pmatrix} 2 & 8 & -1 \\ 3 & 6 & 4 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 1 & 7 \\ 9 & -2 \\ 6 & 3 \end{pmatrix}$, then $c_{11} = a_1.b._1 = \begin{pmatrix} 2 & 8 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 9 \\ 6 \end{pmatrix} = 2x1+8x9-1*6=68$ and

$c_{12} = a_1.b._2 = 1$, $c_{21} = a_2.b._1 = 81$, $c_{21} = a_2.b._1 = 21$ and $\mathbf{AB} = \begin{pmatrix} 68 & 1 \\ 81 & 21 \end{pmatrix} = \mathbf{C}$.

An interesting (and useful) operation is $\mathbf{A'A}$ because it creates a *square matrix*.

Notice that matrix multiplication requires that the first matrix must have as many columns as the second matrix has rows. A consequence is that (**AB**)**C** = **A**(**BC**). Notice also that **AB** is usually not equal to **BA**. Indeed, while **AB** may be a possible operation, **BA** may not be a possible operation.

TRY THIS M2: Verify that **IA**, **IB** and **I*v*** = **A**, **B** and ***v*** respectively.

Matrix operations in EXCEL.

TRY THIS M3: Let $\mathbf{A} = \begin{pmatrix} 1 & 7 & 8 \\ 2 & 5 & 3 \\ 3 & 3 & 15 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, Use EXCEL to calculate **A+B, A-B, AB, BA, AA′, A′A, BB′, B′B**

The *inverse of a matrix* $\mathbf{A}$ is denoted $\mathbf{A}^{-1}$ and can be understood by recalling that in scalar algebra the inverse of a real number multiplied by the real number will be = 1; the inverse of a is a^{-1}. In matrix algebra the inverse of a matrix is a matrix when multiplied by the original matrix is $\mathbf{I}$, formally a matrix $\mathbf{A}$ is invertible if $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. Only square matrices will have an inverse, although not all square matrices have an inverse.

Our motivation for finding the inverse of a matrix is that we want to calculate values for the unknown parameters in linear models.

A solution, if there is a solution, is to utilize the inverse of $\mathbf{X}$ and apply it to both sides of the equation $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta}$ to produce

$$\mathbf{X}^{-1}\boldsymbol{y} = \boldsymbol{\beta},$$

This operation is possible because $\mathbf{X}^{-1}\mathbf{X}\,\boldsymbol{\beta} = \mathbf{I}\boldsymbol{\beta} = \boldsymbol{\beta}$.

However, the variables and observations included in $\mathbf{X}$ seldom create a square matrix, so unless $\mathbf{X}$ is square we cannot directly find a solution for the parameters in $\boldsymbol{\beta}$. With what we already know about the transpose of a matrix we can obtain an algebraic solution:

Remember $\mathbf{X}'\mathbf{X}$ is a square matrix and if we multiply both sides of the equation $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta}$ by $\mathbf{X}'$, then we have the equivalent relationship $\mathbf{X}'\boldsymbol{y} = \mathbf{X}'\mathbf{X}\boldsymbol{\beta}$. Now we can solve for $\boldsymbol{\beta}$ by taking the inverse of $\mathbf{X}'\mathbf{X}$:

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{y} = \boldsymbol{\beta},$$

which has a solution as long as $(\mathbf{X}'\mathbf{X})$ can be inverted. If it cannot be inverted, and we attempt to solve for the unknown parameters using some piece of software, the software will produce a cryptic message that may or may not help us understand the nature of the computational problem.

Invertible square matrices consisting of n rows and n columns have characteristics that can be used to determine if there is an inverse:

- All n columns of $(\mathbf{X'X})$ are independent
- All n rows of $(\mathbf{X'X})$ are independent
- The column space of $(\mathbf{X'X})$ has dimension = n
- The row space of $(\mathbf{X'X})$ has dimension = n
- The determinant of the matrix is not zero
- $(\mathbf{X'X})$ has full rank, r, = n
- All eigenvalues of $(\mathbf{X'X})$ are not zero
- $(\mathbf{X'X})$ has n (nonzero) pivots

All of these characteristics are equivalent and indicate that $(\mathbf{X'X})$ is a nonsingular nxn matrix.

Matrix inversion in EXCEL.

TRY THIS M4: : Let $\mathbf{A} = \begin{pmatrix} 1 & 7 & 8 \\ 2 & 5 & 3 \\ 3 & 3 & 15 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, Use EXCEL to determine $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$, $(\mathbf{AA'})^{-1}$, $(\mathbf{A'A})^{-1}$, $(\mathbf{BB'})^{-1}$, $(\mathbf{B'B})^{-1}$