Vectors and Matrices: Basic Concepts

We begin by introducing some very basic concepts that are essential for algebraic explanations of subsequent statistical techniques. The intent is to provide a brief refresher of the concepts which will be used throughout the semester for those who have not been previously exposed to the rudiments of matrix algebra. Toward the end, derivations are presented for particular algebraic forms that repeatedly appear in papers dealing with mixed models.

VECTORS

Many concepts, such as observed measurements of a trait, factors of a model, or an animal's breeding value cannot be adequately quantified as a single number. Rather, several different measurements $y_1, y_2, \ldots, y_m$ are required.

Definition 1) An m-tuple of real numbers $(y_1, y_2, \ldots, y_m)$ arranged in a column is called a vector and is denoted by a boldfaced, lowercase letter. Other references might use an ~ underline in place of a boldface letter.

Examples of vectors are:

\[
\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}
\]

Vectors are said to be equal if their corresponding entries are the same.

Definition 2) (Scalar multiplication) Let $c$ be an arbitrary scalar. Then the product $c\mathbf{x}$ is a vector with $i^{th}$ entry $cx_i$.

To illustrate scalar multiplication, take $c_1=5$ and $c_2=-1.2$. Then
Definition 3) (Vector addition) The sum of two vectors $x$ and $y$, each having the same number of entries, is that vector
\[
z = x + y, \text{ with } i^{th} \text{ entry } z_i = x_i + y_i
\]
Thus,
\[
\begin{bmatrix}
3 \\
-1 \\
4
\end{bmatrix}
+ 
\begin{bmatrix}
1 \\
2 \\
-2
\end{bmatrix}
= 
\begin{bmatrix}
4 \\
1 \\
2
\end{bmatrix}
\]

\[
x + y = z
\]

Definition 4) The vector $y = a_1x_1 + a_2x_2 + \ldots + a_kx_k$ is a linear combination of the vectors $x_1$, $x_2$, ..., $x_k$.

Definition 5) A set of vectors $x_1$, $x_2$, ..., $x_k$ is said to be linearly dependent if there exist $k$ numbers $(a_1, a_2, \ldots, a_k)$, not all zero, such that
\[
a_1x_1 + a_2x_2 + \ldots + a_kx_k = 0.
\]
Otherwise the set of vectors is said to be linearly independent.

If one vector, for example $x_1$, is $0$, the set is linearly dependent (let $a_i$ be the only nonzero coefficient from Definition 5).

The familiar vectors with a one as an entry and zeros elsewhere are linearly independent.

For $m=4$,
\[
x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad x_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}
\]
so
\[
c_1y = 5 * \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}, \text{ and } c_2y = -1.2 * \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} -1.2 \\ -2.4 \end{bmatrix}
\]
0 = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = \begin{bmatrix}
a_1 + a_2 + a_3 + a_4 \\
a_1 + a_2 + a_3 + a_4 \\
a_1 + a_2 + a_3 + a_4 \\
a_1 + a_2 + a_3 + a_4 \\
\end{bmatrix}
implies a_1 = a_2 = a_3 = a_4 = 0.

As another example, let k=3 and m=3.

\begin{align*}
x_1 &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \\
x_2 &= \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \\
x_3 &= \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}
\end{align*}

Then

2x_1 - x_2 + 3x_3 = 0.

Thus x_1, x_2, x_3 are a linearly dependent set of vectors since any one of them can be written as a linear combination of the others (for example, x_2 = 2x_1 + 3x_3).

**Definition 6**  The inner (or dot) product of two vectors x and y with the same number of entries is defined as the sum of component products

\[ x_1y_1 + x_2y_2 + ... + x_my_m. \]

We use the notation x'y or y'x to denote this inner product.

**MATRICES**

**Definition 7**  A mxk matrix, generally denoted by boldface uppercase letters like A, R, and so forth, is a rectangular array of elements having m rows and k columns. As was the case for vectors an underline can be used in place of boldface letters (e.g. ).

**Definition 8**  The dimension of an mxk matrix is the ordered pair (m,k); m is the row dimension and k is the column dimension.

An mxk matrix, call it A, of arbitrary constants can be written
or more compactly as $A_{mxk} = \{a_{ij}\}$, where the index $i$ refers to the row and the index $j$ refers to the column.

**Definition 9** (Scalar multiplication). Let $c$ be an arbitrary scalar and $A = \{a_{ij}\}$. Then $cA_{mxk} = A_{mxk}c = B_{mxk} = \{b_{ij}\}$, where $b_{ij} = ca_{ij} = a_{ij}c$, $i=1,2,...,m$, and $j=1,2,...,k$. Multiplication of a matrix by a scalar produces a new matrix whose elements are the elements of the original matrix, each multiplied by the scalar.

**Definition 10** (Matrix multiplication). The product $AB$ of an $mxn$ matrix $A = \{a_{ij}\}$ and an $nxk$ matrix $B = \{b_{ij}\}$ is the $mxk$ matrix $C$ whose elements $c_{ij}$ are given by

$$c_{ij} = \sum_{t=1}^{n} a_{it} b_{tj}$$

Note that for the product $AB$ to be defined, the column dimension of $A$ must equal the row dimension of $B$. Then the row dimension of $AB$ equals the row dimension of $A$ and the column dimension of $AB$ equals the column dimension of $B$.

**Result 10a** For all matrices $A$, $B$, and $C$ (dimensions such that the indicated products are defined) and a scalar $c$,

a) $c(AB) = (cA)B$

b) $A(BC) = (AB)C$

c) $A(B+C) = AB + AC$

d) $(B+C)A = BA + CA$

e) $(AB)' = B'A'$

There are several important differences between the algebra of matrices and the algebra of real numbers. Two of these differences are:
a) Matrix multiplication is, in general, not commutative. That is, in general \( AB \neq BA \).

b) Let 0 denote the zero matrix, that is, the matrix with zero for every element. In the algebra of real numbers, if the product of two numbers, ab, is zero, then a=0 or b=0. In matrix algebra, however, the product of two nonzero matrices may be a zero matrix. Hence,

\[(A_{mXn})(B_{nXk}) = 0\]

does not imply that \( A=0 \) or \( B=0 \).

**Definition 11** Inverse of Partitioned Matrices (Searle, 1966 p.210). Obtaining the inverse of a matrix can often be simplified by partitioning it into four sub-matrices in a manner such that those at the top left and bottom right are square; i.e.,

\[
M = \begin{bmatrix}
A & B \\
C & D
\end{bmatrix}
\]

where \( A \) and \( D \) are square. If the corresponding partitioned form of \( M^{-1} \) is

\[
M^{-1} = \begin{bmatrix}
X & Y \\
Z & W
\end{bmatrix}
\]

Then

\[
Z = -D^{-1}CX \\
Y = -A^{-1}BW
\]

\[
X = (A-BD^{-1}C)^{-1} \\
W = (D-CA^{-1}B)^{-1}
\]

The procedures simplify when \( M \) is symmetric, \( M=M' \), for then \( A=A', C=B', D=D', \) and \( X=X', Z=Y', \) and \( W=W' \). That is,

\[
M = \begin{bmatrix}
A & B \\
B' & D
\end{bmatrix} \quad \text{and} \quad M^{-1} = \begin{bmatrix}
X & Y \\
Y' & W
\end{bmatrix}
\]

One will encounter alternate expressions depending on the existence of \( A^{-1} \) or \( D^{-1} \) and the ease of computation.

**Procedure a** (Requires the existence of \( D^{-1} \) but not \( A^{-1} \)).

\[
X = (A-BD^{-1}B)^{-1}
\]

\[
Y = -XB^{-1}\]

5
\[ W = D^{-1} - YB'BD^{-1} \]

**Procedure b** (Requires the existence of \( A^{-1} \) but not \( D^{-1} \)).

\[ W = (D-B'A^{-1}B)^{-1} \]
\[ Y = -A^{-1}BW \]
\[ X = A^{-1} - YB'A^{-1} \]

**Definition 12**) Direct Sum (Searle, 1982 p. 283). Direct sums and direct products are matrix operations defined in terms of partitioned matrices. The direct sum of two matrices \( A \) and \( B \) is defined as

\[
A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}
\]

and extends very simply to more than two matrices

\[
A \oplus B \oplus C = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{bmatrix}
\]

The definition and its extensions apply whether or not the submatrices are of the same order; and all null matrices are of appropriate order. For example,

\[
\begin{bmatrix} 1 & 2 & 3 \\ 6 & 7 \end{bmatrix} \oplus \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 & 7 \\ 0 & 0 & 0 & 8 & 9 \end{bmatrix}
\]

transposing a direct sum gives the direct sum of the transposes.

**Definition 13**) Direct Product (Searle, 1982 p. 265). The direct product of two matrices \( A_{p \times q} \) and \( B_{m \times n} \) is defined as

\[
A_{p \times q} \otimes B_{m \times n} = \begin{bmatrix} a_{11}B & \cdots & a_{1q}B \\ \vdots & \ddots & \vdots \\ a_{p1}B & \cdots & a_{pq}B \end{bmatrix}
\]
A direct product matrix is partitioned into as many submatrices as there are elements of \( A \), each submatrix being \( B \) multiplied by an element of \( A \). It has order \( m \times p \times q \). For example,

\[
\begin{bmatrix}
1 & 2 & 3 \\
8 & 9 \\
\end{bmatrix} \otimes 
\begin{bmatrix}
6 & 7 \\
12 & 14 & 18 & 21 \\
8 & 9 & 16 & 18 & 24 & 27 \\
\end{bmatrix}
\]

Direct products have many useful and interesting properties:

a) In contrast to \((AB)' = B'A'\), we have \((A \otimes B)' = A' \otimes B'\)

b) For vectors \( x \) and \( y \), \( x \otimes y = yx' = yx' \)

c) For a scalar \( l \), \( l \otimes A = A \otimes l \)

d) For partitioned matrices, although

\[
\begin{bmatrix}
A_1 & A_2 \\
B_1 & B_2 \\
\end{bmatrix} \otimes B = \begin{bmatrix}
A_1 \otimes B & A_2 \otimes B \\
A \otimes B_1 & A \otimes B_2 \\
\end{bmatrix}
\]

**Definition 14**) Differentiation.

a) \( w = Ax \) \( \quad \frac{\partial w}{\partial x} = A' = A \) if \( A \) is symmetric.

b) \( w = x'Ax \) \( \quad \frac{\partial w}{\partial x} = Ax + A'x = 2Ax \) if \( A \) is symmetric.

c) \( V = \sum_{i=1}^{s} V_i \sigma_i^2 \) \( \quad \frac{\partial V}{\partial \sigma_k^2} = V_k \) for \( k \leq s \) and \( k \geq 1 \).

d) \( w = \log |V|, V = \sum_{i=1}^{s} V_i \sigma_i^2 \) \( \quad \frac{\partial w}{\partial \sigma_k^2} = \text{tr} (V^{-1} \frac{\partial V}{\partial \sigma_k^2}) = \text{tr} (V^{-1} V_k) \).

**Definition 15**) Inverse of \( V \).

a) \( V = ZGZ' + R \) then \( V^{-1} = R^{-1} - R^{-1}Z(ZR^{-1}Z + G^{-1})^{-1}ZR^{-1} \)

Proof: Show that \( VV^{-1} = I \)

b) \( ZV^{-1} = G^{-1}(ZR^{-1}Z + G^{-1})^{-1}ZR^{-1} \)

c) The above equalities are used many times in variance component estimation methods. Try to memorize these results.
There are a number of ways to find the inverse of a matrix. Two common approaches are:

a) Row-wise reduction

b) Determinants

1. Find the determinant of a matrix.
2. Divide the determinant into the determinant of each principal cofactor matrix.
   The cofactor matrix is the matrix remaining after deleting all row and column elements with the same subscripts for the element you seek.
3. If the sum of the row and column subscripts is an odd number, multiply the coefficient by minus one.
4. Repeat b and c for all elements and take the transpose of the matrix.

Whichever method is best depends on the elements in the original matrix. I have found the methods of determinants works well for 2x2 and 3x3 matrixes, but I do not recommend it for larger matrices.

Example 1.

$$A_{2x2} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

The determinant, D, is $D = 1*4 - 3*2 = -2$.

$$A^{-1} = -\frac{1}{2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix}$$

Check that $AA^{-1}=I$. The determinant is negative and diagonal elements of the inverse are negative. This is not a general result, a least-squares coefficient matrix will have a positive determinant and all diagonal elements of the inverse are positive.
Example 2. \( A_{3\times3} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{bmatrix} \)

The determinant, \( D = c_1 + c_2 + c_3 - d_1 - d_2 - d_3 \) where

\[
\begin{align*}
c_1 &= a_{11} \cdot a_{22} \cdot a_{33} \\
c_2 &= a_{12} \cdot a_{23} \cdot a_{31} \\
c_3 &= a_{13} \cdot a_{21} \cdot a_{32} \\
d_1 &= a_{31} \cdot a_{22} \cdot a_{13} \\
d_2 &= a_{32} \cdot a_{23} \cdot a_{11} \\
d_3 &= a_{33} \cdot a_{21} \cdot a_{12}
\end{align*}
\]

as defined by the arrows. \( D = 0 \), therefore there is a linear dependency among columns of the matrix (i.e. \( 2x_2-x_1=x_3 \), where \( x_1, x_2, \) and \( x_3 \) are columns of the matrix \( A \)).

Example 3. \( A_{3\times3} = \begin{bmatrix} 1 & 2 & 5 \\ 6 & 7 & 9 \\ 10 & 11 & 13 \end{bmatrix} \)

The determinant, \( D \), is \( 1 \cdot 7 \cdot 13 + 2 \cdot 9 \cdot 10 + 5 \cdot 6 \cdot 11 - 10 \cdot 7 \cdot 5 - 11 \cdot 9 \cdot 1 - 13 \cdot 6 \cdot 2 = 91 + 180 + 330 - 350 - 99 - 156 = -4 \).

\[
A^{-1} = -\frac{1}{4} \begin{bmatrix} -8 & 12 & -4 \\ 29 & -37 & 9 \\ -17 & 21 & -5 \end{bmatrix}' = -\frac{1}{4} \begin{bmatrix} -8 & 29 & -17 \\ 12 & -37 & 21 \\ -4 & 9 & -5 \end{bmatrix} = \begin{bmatrix} 2 & -7.5 & 4.25 \\ -3 & 9.25 & -5.25 \\ 1 & -2.25 & 1.25 \end{bmatrix}
\]

References


Purpose: to give a thorough description, and a rigorous mathematical analysis, of some of the most commonly used methods in Numerical Linear Algebra and Optimisation.


Purpose: can be used in both undergraduate and beginning graduate courses, as well as by working scientists and engineers for self study and reference. Special features: applications- computational problems accompanied by real-life examples; recent research- ; explanation of concepts; fundamental linear algebra problems; examples with MATLAB.


Purpose: programming with MATLAB.