Unbalanced Analysis of Variance, Design, and Regression
Ronald Christensen, fletcher@stat.unm.edu

Appendix A: Matrices

A matrix is a rectangular array of numbers. Such arrays have *rows* and *columns*. The numbers of rows and columns are referred to as the *dimensions* of a matrix. A matrix with, say, 5 rows and 3 columns is referred to as a 5×3 matrix.

EXAMPLE A.0.1. Three matrices are given below along with their dimensions.

$$\underset{3\times 2}{\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}}, \quad \underset{2\times 2}{\begin{bmatrix} 20 & 80 \\ 90 & 140 \end{bmatrix}}, \quad \underset{4\times 1}{\begin{bmatrix} 6 \\ 180 \\ -3 \\ 0 \end{bmatrix}}.$$

□

Let r be an arbitrary positive integer. A matrix with r rows and r columns, i.e., an $r \times r$ matrix, is called a *square matrix*. The second matrix in Example A.0.1 is square. A matrix with only one column, i.e., an $r \times 1$ matrix, is a *vector*, sometimes called a *column vector*. The third matrix in Example A.0.1 is a vector. A $1 \times r$ matrix is sometimes called a *row vector*.

An arbitrary matrix A is often written

$$A = [a_{ij}]$$

where a_{ij} denotes the element of A in the ith row and jth column. Two matrices are equal if they have the same dimensions and all of their elements (entries) are equal. Thus for $r \times c$ matrices $A = [a_{ij}]$ and $B = [b_{ij}]$, $A = B$ if and only if $a_{ij} = b_{ij}$ for every $i = 1, \dots, r$ and $j = 1, \dots, c$.

EXAMPLE A.0.2. Let

$$A = \begin{bmatrix} 20 & 80 \\ 90 & 140 \end{bmatrix} \text{ and } B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

If $B = A$, then $b_{11} = 20, b_{12} = 80, b_{21} = 90$, and $b_{22} = 140$. □

The *transpose* of a matrix A, denoted A', changes the rows of A into columns of a new matrix A'. If A is an $r \times c$ matrix, the transpose A' is a $c \times r$ matrix. In particular, if we write $A' = [\tilde{a}_{ij}]$, then the element in row i and column j of A' is defined to be $\tilde{a}_{ij} = a_{ji}$.

EXAMPLE A.0.3.

$$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}' = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

and

$$\begin{bmatrix} 20 & 80 \\ 90 & 140 \end{bmatrix}' = \begin{bmatrix} 20 & 90 \\ 80 & 140 \end{bmatrix}.$$

The transpose of a column vector is a row vector,

$$\begin{bmatrix} 6 \\ 180 \\ -3 \\ 0 \end{bmatrix}' = [6 \quad 180 \quad -3 \quad 0]. \qquad \square$$

A.1 Matrix addition and subtraction

Two matrices can be added (or subtracted) if they have the same dimensions, that is, if they have the same number of rows and columns. Addition and subtraction is performed elementwise.

EXAMPLE A.1.1.

$$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} + \begin{bmatrix} 2 & 8 \\ 4 & 10 \\ 6 & 12 \end{bmatrix} = \begin{bmatrix} 1+2 & 4+8 \\ 2+4 & 5+10 \\ 3+6 & 6+12 \end{bmatrix} = \begin{bmatrix} 3 & 12 \\ 6 & 15 \\ 9 & 18 \end{bmatrix}.$$

$$\begin{bmatrix} 20 & 80 \\ 90 & 140 \end{bmatrix} - \begin{bmatrix} -15 & -75 \\ 80 & 130 \end{bmatrix} = \begin{bmatrix} 35 & 155 \\ 10 & 10 \end{bmatrix}. \qquad \square$$

In general, if A and B are $r \times c$ matrices with $A = [a_{ij}]$ and $B = [b_{ij}]$, then

$$A + B = [a_{ij} + b_{ij}] \text{ and } A - B = [a_{ij} - b_{ij}].$$

A.2 Scalar multiplication

Any matrix can be multiplied by a scalar. Multiplication by a scalar (a *real number*) is elementwise.

EXAMPLE A.2.1. Scalar multiplication gives

$$\frac{1}{10}\begin{bmatrix} 20 & 80 \\ 90 & 140 \end{bmatrix} = \begin{bmatrix} 20/10 & 80/10 \\ 90/10 & 140/10 \end{bmatrix} = \begin{bmatrix} 2 & 8 \\ 9 & 14 \end{bmatrix}.$$

$$2[6 \quad 180 \quad -3 \quad 0] = [12 \quad 360 \quad -6 \quad 0]. \qquad \square$$

In general, if λ is any number and $A = [a_{ij}]$, then

$$\lambda A = [\lambda a_{ij}].$$

A.3 Matrix multiplication

Two matrices can be multiplied together if the number of columns in the first matrix is the same as the number of rows in the second matrix. In the process of multiplication, the rows of the first matrix are matched up with the columns of the second matrix.

EXAMPLE A.3.1.

$$\begin{aligned} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 20 & 80 \\ 90 & 140 \end{bmatrix} &= \begin{bmatrix} (1)(20)+(4)(90) & (1)(80)+(4)(140) \\ (2)(20)+(5)(90) & (2)(80)+(5)(140) \\ (3)(20)+(6)(90) & (3)(80)+(6)(140) \end{bmatrix} \\ &= \begin{bmatrix} 380 & 640 \\ 490 & 860 \\ 600 & 1080 \end{bmatrix}. \end{aligned}$$

The entry in the first row and column of the product matrix, $(1)(20)+(4)(90)$, matches the elements in the first row of the first matrix, (1 4), with the elements in the first column of the second matrix, $\begin{pmatrix} 20 \\ 90 \end{pmatrix}$. The 1 in (1 4) is matched up with the 20 in $\begin{pmatrix} 20 \\ 90 \end{pmatrix}$ and these numbers are multiplied. Similarly, the 4 in (1 4) is matched up with the 90 in $\begin{pmatrix} 20 \\ 90 \end{pmatrix}$ and the numbers are multiplied. Finally, the two products are added to obtain the entry $(1)(20)+(4)(90)$. Similarly, the entry in the third row, second column of the product, $(3)(80)+(6)(140)$, matches the elements in the third row of the first matrix, (3 6), with the elements in the second column of the second matrix, $\begin{pmatrix} 80 \\ 140 \end{pmatrix}$. After multiplying and adding we get the entry $(3)(80)+(6)(140)$. To carry out this matching, the number of columns in the first matrix must equal the number of rows in the second matrix. The matrix product has the same number of rows as the first matrix and the same number of columns as the second because each row of the first matrix can be matched with each column of the second.

□

EXAMPLE A.3.2. We illustrate another matrix multiplication commonly performed in statistics, multiplying a matrix on its left by the transpose of that matrix, i.e., computing $A'A$.

$$\begin{aligned} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}' \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} &= \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \\ &= \begin{bmatrix} 1+4+9 & 4+10+18 \\ 4+10+18 & 16+25+36 \end{bmatrix} \\ &= \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix}. \end{aligned}$$

□

Notice that in matrix multiplication the roles of the first matrix and the second matrix are *not* interchangeable. In particular, if we reverse the order of the matrices in Example A.3.1, the matrix product

$$\begin{bmatrix} 20 & 80 \\ 90 & 140 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

is undefined because the first matrix has two columns while the second matrix has three rows. Even when the matrix products are defined for both AB and BA, the results of the multiplication typically differ. If A is $r \times s$ and B is $s \times r$, then AB is an $r \times r$ matrix and BA is and $s \times s$ matrix. When $r \neq s$, clearly $AB \neq BA$, but even when $r = s$ we still can not expect AB to equal BA.

EXAMPLE A.3.3. Consider two square matrices, say,

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 2 \\ 1 & 2 \end{bmatrix}.$$

Multiplication gives

$$AB = \begin{bmatrix} 2 & 6 \\ 4 & 14 \end{bmatrix}$$

and

$$BA = \begin{bmatrix} 6 & 8 \\ 7 & 10 \end{bmatrix},$$

so $AB \neq BA$.

□

In general if $A = [a_{ij}]$ is an $r \times s$ matrix and $B = [b_{ij}]$ is a $s \times c$ matrix, then

$$AB = [d_{ij}]$$

is the $r \times c$ matrix with

$$d_{ij} = \sum_{\ell=1}^{s} a_{i\ell} b_{\ell j}.$$

A useful result is that the transpose of the product AB is the product, in reverse order, of the transposed matrices, i.e. $(AB)' = B'A'$.

EXAMPLE A.3.4. As seen in Example A.3.1,

$$AB \equiv \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 20 & 80 \\ 90 & 140 \end{bmatrix} = \begin{bmatrix} 380 & 640 \\ 490 & 860 \\ 600 & 1080 \end{bmatrix} \equiv C.$$

The transpose of this matrix is

$$C' = \begin{bmatrix} 380 & 490 & 600 \\ 640 & 860 & 1080 \end{bmatrix} = \begin{bmatrix} 20 & 90 \\ 80 & 140 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = B'A'.$$

□

Let $a = (a_1, \ldots, a_n)'$ be a vector. A very useful property of vectors is that

$$a'a = \sum_{i=1}^{n} a_i^2 \geq 0.$$

A.4 Special matrices

If $A = A'$, then A is said to be *symmetric*. If $A = [a_{ij}]$ and $A = A'$, then $a_{ij} = a_{ji}$. The entry in row i and column j is the same as the entry in row j and column i. Only square matrices can be symmetric.

EXAMPLE A.4.1. The matrix

$$A = \begin{bmatrix} 4 & 3 & 1 \\ 3 & 2 & 6 \\ 1 & 6 & 5 \end{bmatrix}$$

has $A = A'$. A is symmetric about the diagonal that runs from the upper left to the lower right. □

For any $r \times c$ matrix A, the product $A'A$ is always symmetric. This was illustrated in Example 24.3.2. More generally, write $A = [a_{ij}]$, $A' = [\tilde{a}_{ij}]$ with $\tilde{a}_{ij} = a_{ji}$, and

$$A'A = [d_{ij}] = \left[\sum_{\ell=1}^{c} \tilde{a}_{i\ell} a_{\ell j} \right].$$

Note that

$$d_{ij} = \sum_{\ell=1}^{c} \tilde{a}_{i\ell} a_{\ell j} = \sum_{\ell=1}^{c} a_{\ell i} a_{\ell j} = \sum_{\ell=1}^{c} \tilde{a}_{j\ell} a_{\ell i} = d_{ji}$$

so the matrix is symmetric.

Diagonal matrices are square matrices with all off diagonal elements equal to zero.

EXAMPLE A.4.2. The matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 20 & 0 \\ 0 & -3 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

are diagonal. □

In general, a diagonal matrix is a square matrix $A = [a_{ij}]$ with $a_{ij} = 0$ for $i \neq j$. Obviously, diagonally matrices are symmetric.

An *identity matrix* is a diagonal matrix with all 1s along the diagonal, i.e., $a_{ii} = 1$ for all i. The third matrix in Example A.4.2 above is a 3×3 identity matrix. The identity matrix gets it name because any matrix multiplied by an identity matrix remains unchanged.

EXAMPLE A.4.3.

$$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

□

An $r \times r$ identity matrix is denoted I_r with the subscript deleted if the dimension is clear.

A *zero matrix* is a matrix that consists entirely of zeros. Obviously, the product of any matrix multiplied by a zero matrix is zero.

EXAMPLE A.4.4.

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

□

Often a zero matrix is denoted by 0 where the dimension of the matrix, and the fact that it is a matrix rather than a scalar, must be inferred from the context.

A matrix M that has the property $MM = M$ is called *idempotent*. A symmetric idempotent matrix is a *perpendicular projection operator*. indexidempotent

EXAMPLE A.4.5. The following matrices are both symmetric and idempotent,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}, \quad \begin{bmatrix} .5 & .5 & 0 \\ .5 & .5 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

□

A.5 Linear dependence and rank

Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & 5 & 1 \\ 2 & 2 & 10 & 6 \\ 3 & 4 & 15 & 1 \end{bmatrix}.$$

Note that each column of A can be viewed as a vector. The *column space* of A, denoted $C(A)$, is the collection of all vectors that can be written as *a linear combination of the columns of* A. In other words, $C(A)$ is the set of all vectors that can be written as

$$\lambda_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} + \lambda_3 \begin{bmatrix} 5 \\ 10 \\ 15 \end{bmatrix} + \lambda_4 \begin{bmatrix} 1 \\ 6 \\ 1 \end{bmatrix} = A \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = A\lambda$$

for some vector $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)'$.

The columns of any matrix A are *linearly dependent* if they contain redundant information. Specifically, let x be some vector in $C(A)$. The columns of A are linearly dependent if we can find two distinct vectors λ and γ such that $x = A\lambda$ and $x = A\gamma$. Thus two distinct linear combinations of the columns of A give rise to the same vector x. Note that $\lambda \neq \gamma$ because λ and γ are distinct. Note also that, using a distributive property of matrix multiplication, $A(\lambda - \gamma) = A\lambda - A\gamma = 0$, where $\lambda - \gamma \neq 0$. This condition is frequently used as an alternative definition for linear dependence, i.e., the columns of A are linearly dependent if there exists a vector $\delta \neq 0$ such that $A\delta = 0$. If the columns of A are not linearly dependent, they are *linearly independent*.

EXAMPLE A.5.1. Observe that the example matrix A given at the beginning of the section has

$$\begin{bmatrix} 1 & 2 & 5 & 1 \\ 2 & 2 & 10 & 6 \\ 3 & 4 & 15 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

so the columns of A are linearly dependent. □

The *rank* of A is the smallest number of columns of A that can generate $C(A)$. It is also the maximum number of linearly independent columns in A.

EXAMPLE A.5.2. The matrix

$$A = \begin{bmatrix} 1 & 2 & 5 & 1 \\ 2 & 2 & 10 & 6 \\ 3 & 4 & 15 & 1 \end{bmatrix}$$

has rank 3 because the columns

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \\ 1 \end{bmatrix}$$

generate $C(A)$. We saw in Example A.5.1 that the column $(5, 10, 15)'$ was redundant. None of the other three columns are redundant; they are linearly independent. In other words, the only way to get

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 2 & 6 \\ 3 & 4 & 1 \end{bmatrix} \delta = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

is to take $\delta = (0, 0, 0)'$. □

A.6 Inverse matrices

The *inverse* of a square matrix A is the matrix A^{-1} such that

$$AA^{-1} = A^{-1}A = I.$$

The inverse of A exists only if the columns of A are linearly independent. Typically, it is difficult to find inverses without the aid of a computer. For a 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

the inverse is given by

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}. \tag{A.6.1}$$

To confirm that this is correct, multiply AA^{-1} to see that it gives the identity matrix. Moderately complicated formulae exist for computing the inverse of 3×3 matrices. Inverses of larger matrices become very difficult to compute by hand. Of course computers are ideally suited for finding such things.

One use for inverse matrices is in solving systems of equations.

EXAMPLE A.6.1. Consider the system of equations

$$\begin{aligned} 2x + 4y &= 20 \\ 3x + 4y &= 10. \end{aligned}$$

We can write this in matrix form as

$$\begin{bmatrix} 2 & 4 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 20 \\ 10 \end{bmatrix}.$$

Multiplying on the left by the inverse of the coefficient matrix gives

$$\begin{bmatrix} 2 & 4 \\ 3 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 4 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 3 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 20 \\ 10 \end{bmatrix}.$$

Using the definition of the inverse on the left-hand side of the equality and the formula in (A.6.1) on the right-hand side gives

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 3/4 & -1/2 \end{bmatrix} \begin{bmatrix} 20 \\ 10 \end{bmatrix}$$

or

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -10 \\ 10 \end{bmatrix}.$$

Thus $(x,y) = (-10,10)$ is the solution for the two equations, i.e., $2(-10)+4(10) = 20$ and $3(-10)+4(10) = 10$. □

More generally a system of equations, say,

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 + a_{13}y_3 &= c_1 \\ a_{21}y_1 + a_{22}y_2 + a_{23}y_3 &= c_2 \\ a_{31}y_1 + a_{32}y_2 + a_{33}y_3 &= c_3 \end{aligned}$$

in which the a_{ij}s and c_is are known and the y_is are variables, can be written in matrix form as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

or

$$AY = C.$$

To find Y simply observe that $AY = C$ implies $A^{-1}AY = A^{-1}C$ and $Y = A^{-1}C$. Of course this argument assumes that A^{-1} exists, which is not always the case. Moreover, the procedure obviously extends to larger sets of equations.

On a computer, there are better ways of finding solutions to systems of equations than finding the inverse of a matrix. In fact, inverses are often found by solving systems of equations. For example, in a 3×3 case the first column of A^{-1} can be found as the solution to

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

For a special type of square matrix, called an *orthogonal matrix*, the transpose is also the inverse. In other words, a square matrix P is an orthogonal matrix if

$$P'P = I = PP'.$$

To establish that P is orthogonal, it is enough to show either that $P'P = I$ or that $PP' = I$. Orthogonal matrices are particularly useful in discussions of eigenvalues and principal component regression.

A.7 A list of useful properties

The following proposition summarizes many of the key properties of matrices and the operations performed on them.

Proposition A.7.1. Let A, B, and C be matrices of appropriate dimensions and let λ be a scalar.

$$\begin{aligned} A+B &= B+A \\ (A+B)+C &= A+(B+C) \\ (AB)C &= A(BC) \\ C(A+B) &= CA+CB \\ \lambda(A+B) &= \lambda A+\lambda B \\ (A')' &= A \\ (A+B)' &= A'+B' \\ (AB)' &= B'A' \\ \left(A^{-1}\right)^{-1} &= A \\ (A')^{-1} &= \left(A^{-1}\right)' \\ (AB)^{-1} &= B^{-1}A^{-1}. \end{aligned}$$

The last equality only holds when A and B both have inverses. The second to the last property implies that the inverse of a symmetric matrix is symmetric because then $A^{-1} = (A')^{-1} = (A^{-1})'$. This is a very important property.

A.8 Eigenvalues and eigenvectors

Let A be a square matrix. A scalar ϕ is an eigenvalue of A and $x \neq 0$ is an eigenvector for A corresponding to ϕ if

$$Ax = \phi x.$$

EXAMPLE A.8.1. Consider the matrix

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{bmatrix}.$$

The value 3 is an eigenvalue and any nonzero multiple of the vector $(1,1,1)'$ is a corresponding eigenvector. For example,

$$\begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Similarly, if we consider a multiple, say, $4(1,1,1)'$,

$$\begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 12 \\ 12 \\ 12 \end{bmatrix} = 3 \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix}.$$

The value 2 is also an eigenvalue with eigenvectors that are nonzero multiples of $(1,-1,0)'$.

$$\begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Finally, 6 is an eigenvalue with eigenvectors that are nonzero multiples of $(1,1,-2)'$. □

Proposition A.8.2. Let A be a symmetric matrix, then for a diagonal matrix $D(\phi_i)$ consisting of eigenvalues there exists an orthogonal matrix P whose columns are corresponding eigenvectors such that

$$A = PD(\phi_i)P'.$$

EXAMPLE A.8.3. Consider again the matrix

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{bmatrix}.$$

In writing $A = PD(\phi_i)P'$, the diagonal matrix is

$$D(\phi_i) = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

The orthogonal matrix is

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{bmatrix}.$$

We leave it to the reader to verify that $PD(\phi_i)P' = A$ and that $P'P = I$.

Note that the columns of P are multiples of the vectors identified as eigenvectors in Example A.8.1; hence the columns of P are also eigenvectors. The multiples of the eigenvectors were chosen so that $PP' = I$ and $P'P = I$. Moreover, the first column of P is an eigenvector corresponding to 3, which is the first eigenvalue listed in $D(\phi_i)$. Similarly, the second column of P is an eigenvector corresponding to 2 and the third column corresponds to the third listed eigenvalue, 6.

With a 3×3 matrix A having three *distinct* eigenvalues, any matrix P with eigenvectors for columns would have $P'P$ a diagonal matrix, but the multiples of the eigenvectors must be chosen so that the diagonal entries of $P'P$ are all 1. □

EXAMPLE A.8.4. Consider the matrix

$$B = \begin{bmatrix} 5 & -1 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & 5 \end{bmatrix}.$$

This matrix is closely related to the matrix in Example A.8.1. The matrix B has 3 as an eigenvalue with corresponding eigenvectors that are multiples of $(1,1,1)'$, just like the matrix A. Once again 6 is an eigenvalue with corresponding eigenvector $(1,1,-2)'$ and once again $(1,-1,0)'$ is an eigenvector, but now, unlike A, $(1,-1,0)$ also corresponds to the eigenvalue 6. We leave it to the reader to verify these facts. The point is that in this matrix, 6 is an eigenvalue that has two linearly independent eigenvectors. In such cases, any nonzero linear combination of the two eigenvectors is also an eigenvector. For example, it is easy to see that

$$3\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ -4 \end{bmatrix}$$

is an eigenvector corresponding to the eigenvalue 6.

To write $B = PD(\phi)P'$ as in Proposition A.8.2, $D(\phi)$ has 3, 6, and 6 down the diagonal and one choice of P is that given in Example A.8.3. However, because one of the eigenvalues occurs more than once in the diagonal matrix, there are many choices for P. □

Generally, if we need eigenvalues or eigenvectors we get a computer to find them for us.

Two frequently used functions of a square matrix are the determinant and the trace.

Definition A.8.5.

a) The determinant of a square matrix is the product of the eigenvalues of the matrix.

b) The trace of a square matrix is the sum of the eigenvalues of the matrix.

In fact, one can show that the trace of a square matrix also equals the sum of the diagonal elements of that matrix.