Matrices

A matrix can be a square or rectangular array of values, enclosed in brackets. The notation for a matrix is a capital bolded letter.

We have already seen matrices in the review of statistics on Day 1.

A data matrix will contain rows of cases and columns of variables; for example:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | SAT | GPA | Self-Esteem | IQ |
| 1 | 560 | 3.0 | 11 | 112 |
| 2 | 780 | 3.9 | 10 | 143 |
| 3 | 620 | 2.9 | 19 | 124 |
| 4 | 600 | 2.7 | 7 | 129 |

**y** =  **X = ** where y is a vector and **X** is a matrix.

A data matrix will contain rows of cases and columns of variables; for example:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | SAT | GPA | Gender | IQ |
| 1 | 560 | 3.0 | 1 | 112 |
| 2 | 780 | 3.9 | 0 | 143 |
| 3 | 620 | 2.9 | 0 | 124 |
| 4 | 600 | 2.7 | 1 | 129 |

**y** =  **X = **

Other examples of matrices include:

 **A** =  **B** = 

We define matrices in terms in the number of rows and columns within the matrix; matrices are identified in terms of their *dimensions*, number of rows × number of columns.

 **A** is a 3 × 3 matrix; **B** is a 3 × 2 matrix.

A matrix with an equal number of rows and columns is a *square matrix*.

A vector is a special case of a matrix, where a row vector has one row and *n* columns and a column vector has *n* rows and one column.

 a = (a, b, c) is a 1 × 3 matrix or a 3-element row vector.

 b =  is a 3 × 1 matrix or a 3-element column vector.

A vector can always be described as a matrix, but a matrix cannot always be described as a vector.

*Group Exercises:*

Describe the following vectors and/or matrices:

**A** = 

b′ = (0, 1, 0, 0)

**Y** = 

# Matrix Elements

Each element of a matrix has two subscripts to denote its location in terms of row and column

**A** =  a r × c matrix

A square matrix is a matrix where the number of rows equals the number of columns, where *n* = *p*, and can be described in terms of its *order*. A 4 × 4 matrix is of order 4.

A square matrix also has a diagonal that goes from the upper left corner to the lower right corner that is called the *principle* or *major* diagonal. Elements not in the diagonal are called *off-diagonal* elements.

X =  The principle diagonal of X is a′ = (32, 23, 54, 24).

Using our matrix X as data from four course exams for four students, consider the following descriptive tools.

1. Which student obtained a 35 and on which test? Report the row and column:
2. What did student 4 receive on exam 1?

# Equality of Matrices

Two matrices can be equal if

1. they have the same dimensions
2. all corresponding elements are equal

A =  and B =  are not equal (different dimension)

*Transpose of a Matrix*

Just as when we transpose a vector, by taking the column and making it a row, we do so with a matrix one column at a time – we interchange each column and row, so the first column becomes the first row, the second column becomes the second row, etc.

X = , so that X**′** = 

Notice with a square matrix, the principle diagonal remains the same.

Does (X′)′ = X ?

# Matrix Addition and Subtraction

Two or more matrices can be added if they all have the same dimensions; if not, matrix addition and subtraction is undefined. Just as in vector addition, each corresponding element is added or subtracted and placed in the corresponding location in the new matrix.

 A =  and B = , then

A + B =  = 

The general case for matrix addition is

 *cij* = *aij* + *bij* (for all *i*, *j*).

The commutative and associative laws of addition work here just as they do in vector addition.

 A + B = B + A and A + (B + C) = (A + B) + C

# Deviation Matrix

We have seen that deviation scores are particularly useful in statistics. We can create a deviation matrix by taking the original matrix and subtracting from it a matrix of means where each column contains the mean for each corresponding column in the original matrix.

D = A – M

# Scalar, Matrix Multiplication

Any matrix can be multiplied by a scalar, where each element in the matrix is multiplied by the value of the scalar.

A =  and λ = -2, so that λ A = -2 = 

The general case is where λ A, resulting in λ *aij*. Because λ *aij* = *aij* λ, then λ A = A λ .

# Matrix Multiplication

The process of multiplying matrices follows the same basic principle as vector multiplication, where we consider matrices to be collections of vectors. When we multiply vectors, they must have the same number of elements because we multiple corresponding elements and add the resulting products, obtaining a scalar product. The first vector is written as a row vector and the second is a column vector, as in our familiar a′a.

Consider an example:

 X =  and Y = .

To complete the multiplication of XY, we take the first row of X and the first column of Y.

 (2, 3)  = (2)(3) + (3)(2) = 12, which gives us the first element of the new matrix Z.

 Next we again take the first row, but now times the second column.

 This continue row by column for each corresponding element in the new matrix.

  

  

 Z = = 

This process must always be carefully followed:

 taking the product of the row of the first matrix by the column of the second matrix.

The general rule on matrix multiplication for AB = C, for each element *cij* in C, *cij* is the scalar product of the *i*th row of A and the *j*th column of B.

So, the element c34 is the scalar product of row 3 in A and column 4 in B.

*Exercises*

 A = , B = , C = , D = 

 Compute AB, BC, and CD

In matrix multiplication, rarely does AB = BA.

The order of matrix multiplication is important. Because of this we have specific language to describe the order of multiplication.

For AB, B is *premultiplied* by A and A is *postmultiplied* by B.

We can also multiply matrices that are not of equal size, as long as they are *conformable*. To be conformable, the matrix must have inner dimensions that are equal – that is, the number of columns in the first matrix must equal the number of rows in the second matrix.

 A =  B = 

2 × 4

4 × 3

Conformable

Because the process is the same as the scalar product of two vectors, the number of elements in a row in A must equal the number of elements in a column of B.

In this example, BA would be undefined, where 4 × 3 does not conform to 2 × 4.

 Compute AB =

 Consider a =  and A = 

 Can we compute a A or a′ A ?

*The Example of Weights*

One example of matrix multiplication involves a matrix of data and a matrix of weights. Consider the earlier matrix of test scores for four students.

X =  w = 

where the first and second exams were each worth 10%, the third was worth 30% and the fourth was worth 50%.

To compute a weighted combined score, since X is (4 × 4) and w is (4 × 1), c will be (4 × 1).

 c = X w

 This multiplication takes each student’s scores and multiplies them by the weights, where for the first student: (32)(0.10) + (54)(0.10) + (56)(0.30) + (21)(0.50).

 c = 

There are a couple of special properties that can be demonstrated through the following examples of matrix multiplication.

 AB = = =

 *Note*: The result is a *null matrix*. In scalar arithmetic, if ab = 0, then one of the two (a or b) must be zero. This is not the case in matrix multiplication.

 EF =  =  and

 EG = = 

 *Note*: In scalar arithmetic, as long as a ≠ 0, when ab = ac, then b = c. This is not necessarily true in matrix multiplication.

 In the first example above, we noted that AB resulted in a null-matrix.

 BA = = 

 *Note*: In scalar arithmetic, if ab = 0, then ba = 0. Not true in matrix multiplication.

Multiple matrices may be multiplied, when they are conformable: A(BC) = (AB)C = ABC.

 EFH = = = 

When employing a scalar, it is true that λ(AB) = (λA)B = A(λB).

Additional relations include:

 (AB)′ = B′A′ and (ABC)′ = C′B′A′

*Symmetric Matrices*

If A = A′, the matrix is symmetric. This suggests that they have the same dimensions and all corresponding elements are equal. So, only square matrices can be symmetric.

 A very common symmetric matrix is a correlation matrix, such as:

 R =  = R′ = 

In general, a matrix is symmetric when *cij* = *cji*.

When any single matrix is multiplied by its transpose, it creates a symmetric (square) matrix. This matrix also provides an interesting statistical tool, a matrix of sums of squares and cross products.

Consider X = , a 4 × 3 matrix.

When we compute X′ X, we obtain *columns product matrix*, a 3 × 3 matrix. This is because the off-diagonal elements are the cross products of the columns.

 X′ X = =

 From this example, we can see that the diagonal values are the sums of squares:

 

 The off diagonal elements are the cross products of the columns:

 

When we compute X X′, we obtain rows product matrix, a 4 × 4 matrix. This is because the off-diagonal elements are the cross products of the rows.

 X X′ ==

Any rectangular matrix can be used to create a rows product or columns product matrix. This will always result in a symmetric matrix.

*Diagonal Matrices*

Any square matrix where all off-diagonal elements are zero is a diagonal matrix.

D is a diagonal matrix if *dij* = 0 for all *i* ≠ *j*.

 D1 =  and D2 =  are diagonal matrices.

Consider pre- and post-multiplication by a diagonal matrix D = 

 DX = =

 Pre-multiplication by the diagonal multiplied each corresponding row in X by the value in the corresponding row of D.

 XD = =

 Post-multiplication by the diagonal multiplied each column in X by the value in the corresponding column of D.

Using diagonal matrices is one way to accomplish division by a scalar. For instance, you could employ the matrix R:

R = 

Through pre-multiplication, you would divide each row in X by the value in the corresponding row of D.

RX = =

 Through post-multiplication, you would then divide each column in X by the value in the corresponding column of R.

 RXR ==

 If X was a covariance matrix and R was a diagonal matrix containing standard deviations, then the resulting matrix would be a correlation matrix – a standardized covariance matrix.

Note that the product of any two diagonal matrices will result in a diagonal matrix which is the simple product of each corresponding element. Try to prove this on your own.

*Scalar Matrix*

A diagonal matrix with all diagonal elements of the same value is called a *scalar matrix*. The following two matrices are scalar matrices.

S1 =  and S2 = , where *sii* = *k*, for *i* = 1 to *n*.

When any matrix is pre- or post-multiplied by a scalar matrix, each element in the matrix is multiplied by the scalar value of the scalar matrix.

 So if AK = B, then *aijk* = *bij*.

 This is equivalent to multiplying a matrix by a scalar.

*Identity Matrix*

A special scalar matrix is one where all diagonal elements are the value of one (1). *Identity matrices* are signified by I.

 I = 

Pre- or post-multiplication of a matrix by an identity matrix results in the original matrix unchanged. This is the same as multiplying any value by 1 in scalar arithmetic.