

A Brief Primer on Matrix Algebra

A matrix is a rectangular array of numbers whose individual entries are called *elements*. Each horizontal array of elements is called a *row*, while each vertical array is referred to as a *column*. The number of rows and columns of the matrix are its *dimensions*, and they are expressed as $r \times c$, with the number of rows always being expressed first. Matrices will be identified by an upper case, boldfaced letter.

For example, the matrix **A** below has r rows and c columns. Each element within the matrix is identified by a double subscript indicating the row and column in which the element is located.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1c} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2c} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{ic} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rj} & \cdots & a_{rc} \end{bmatrix}$$

If the generic elements are replaced with actual numbers, the result is matrix **A** below. In this example, the element $a_{23} = 6$ and the element $a_{44} = 7$. The primary thing to remember about subscripting these elements is that the row identifier always precedes that for the column.

$$\mathbf{A} = \begin{bmatrix} 3 & 8 & 2 & 4 \\ 1 & 9 & 6 & 3 \\ 9 & 4 & 5 & 1 \\ 7 & 3 & 2 & 7 \end{bmatrix}$$

When a matrix has only one row or one column, it is called a vector. When a matrix has only a single element, it is called a scalar. A column vector will be identified by a lower case, boldfaced letter, while a row vector will be labeled similarly but with a prime, \mathbf{a}' . Below are examples of a column vector and a row vector.

$$\mathbf{a} = \begin{bmatrix} 7 \\ 9 \\ 5 \\ 3 \end{bmatrix} \quad \mathbf{a}' = [7 \quad 9 \quad 5 \quad 3]$$

Addition and Subtraction

Matrices can be added and subtracted just the way that individual numbers can be added and subtracted. In matrix algebra, this is referred to as elementwise addition and subtraction. When performing elementwise operations, each element in matrix \mathbf{A} is added to or subtracted from its corresponding element in matrix \mathbf{B} . Hence the elements of the new matrix, \mathbf{AB} , are $ab_{11} = a_{11} + b_{11}$, $ab_{12} = a_{12} + b_{12}$, and so forth. In the example below,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \\ a_{31} + b_{31} & a_{32} + b_{32} \end{bmatrix}.$$

Subtraction is done the same way.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} - \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} \\ a_{21} - b_{21} & a_{22} - b_{22} \\ a_{31} - b_{31} & a_{32} - b_{32} \end{bmatrix}$$

Note that just as is the case for regular arithmetic (the technical term is “scalar algebra,” but that seems pretentious), the order in which matrices are added does not matter. Hence, $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$. Similarly, if both addition and subtraction are involved, the order of the operations does not matter: $\mathbf{A} + \mathbf{B} - \mathbf{C} = \mathbf{A} - \mathbf{C} + \mathbf{B} = \mathbf{B} - \mathbf{C} + \mathbf{A}$, and so on. This lack of concern about order does not hold for subtraction only, where what is subtracted from what obviously does make a difference. Thus, $\mathbf{A} - \mathbf{B} \neq \mathbf{B} - \mathbf{A}$.

Multiplication of Matrices

Although elementwise multiplication of the matrix elements can be and is occasionally performed in more advanced matrix calculations that is not the type of multiplication

that is generally encountered in matrix manipulations. Rather, in standard matrix multiplication, one must follow a row-by-column rule that strictly dictates how the multiplication is performed. This will be demonstrated shortly.

Because this row-by-column rule governs matrix multiplication, only matrices with certain eligible dimensions can be multiplied. Specifically, the columns of the first matrix (technically called the “prefactor”) must be identical in number to the rows for the second matrix (“postfactor”). When this condition is met, the matrices are said to be conformable. In Eq. A.1, the two inner dimensions of matrices **A** and **B** are the same, thus fulfilling the condition for conformability. Note also that the product matrix **C** has dimensions which are equal to the two outer dimensions of **A** and **B**. This will always be the case.

$$\begin{matrix} \mathbf{A} & \mathbf{B} & = & \mathbf{C} \\ r \times c & c \times r & & r \times r \end{matrix} \quad (\text{A.1})$$

The row-by-column rule dictates that the first row of **A** should be elementwise multiplied by the first column **B** and the c product terms should be summed. Then the first row of **A** is elementwise multiplied by the second column of **B** and the products are summed. This continues until the first row of **A** has exhausted all of the columns of **B**. Then the second row of **A** continues in the same fashion. This is an instance where an example will do much to clarify the previous sentences.

Assume that matrix **A** is 3×3 and matrix **B** is 3×3 . They are obviously conformable. For simplicity, letters without subscripts will be used to illustrate this multiplication.

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} j & k & l \\ m & n & o \\ p & q & r \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} aj + bm + cp & ak + bn + cq & al + bo + cr \\ dj + em + hp & dk + en + hq & dl + eo + hr \\ gj + hm + ip & gk + hn + iq & gl + ho + ir \end{bmatrix}$$

Any number of matrices can be multiplied in this manner as long as all of the matrices are conformable in the sequence in which they are multiplied. Thus,

$$\begin{matrix} \mathbf{E} & = & \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{D} \\ r \times r & & r \times c & c \times p & p \times q & q \times r \end{matrix} .$$

Vectors of numbers can be multiplied in much the same way. When a column vector is premultiplied by a row vector, the result is a scalar because both of the vectors' outer dimensions are 1. Hence,

$$\begin{aligned} \mathbf{a}'\mathbf{b} &= \mathbf{c} \\ [3 \quad 4 \quad 5] \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} &= 23 \end{aligned} \tag{A.2}$$

In contrast, when a column vector is postmultiplied by a row vector, the result is a full matrix. Equation A.3 displays this as:

$$\begin{aligned} \mathbf{a}\mathbf{b}' &= \mathbf{C} . \\ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} [3 \quad 4 \quad 5] &= \begin{bmatrix} 3 & 4 & 5 \\ 6 & 8 & 10 \\ 9 & 12 & 15 \end{bmatrix} \end{aligned} \tag{A.3}$$

Finally, if a matrix is postmultiplied by a column vector, the result is another column vector. Another numerical example is not necessary. Nevertheless, if the letters are changed a bit, Eq. 9.4 is easily recognizable as the matrix version of the familiar multiple regression equation:

$$\hat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta} , \tag{A.4}$$

where \mathbf{X} is a $N \times p$ matrix containing the predictor variables, $\boldsymbol{\beta}$ ($p \times 1$) is a column vector containing the regression coefficients, and $\hat{\mathbf{y}}$ ($n \times 1$) is a column vector containing the predicted values of Y (note: $\boldsymbol{\beta}$ is the lower case Greek beta).

Matrix Division and Inversion

The term “division” is somewhat misleading in matrix algebra. Although elementwise division is a legitimate operation, this procedure is used infrequently in statistical calculations. Even though element-by-element division is acceptable (if seldom used), division of an entire matrix by another matrix is not possible. Do not despair for there is a work around, and most people learned it in elementary school.

Whenever our fourth grade teacher wanted 4 to be divided by $\frac{1}{4}$, the obvious way to proceed was to invert the $\frac{1}{4}$ and multiply the 4. The principle is exactly the same for matrices. If the calculations call for \mathbf{A} to be divided by \mathbf{B} , \mathbf{B} is first inverted and then the matrices are multiplied as outlined in the previous section.

Matrix inversion is both difficult and tedious to do with only a calculator, so one of the major advantages of having statistical software is that the computer can do the heavy lifting. Despite not having to calculate an inverted matrix by hand, several of their features are important. First, the inverse of a matrix is symbolized by an exponent of -1 . Hence, the inverse of \mathbf{A} is \mathbf{A}^{-1} . Sometimes, a scalar will have the -1 exponent, which means exactly what it says, invert the number. Hence, 9^{-1} should be read as $1/9$. If the negative exponent is larger than 1, it means to square or cube the number before inverting it, e.g., $6^{-2} = 1/36$.¹

Finally, if a matrix is multiplied by its inverse, the product is a special type of matrix called an identity matrix. Thus,

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}, \tag{A.5}$$

which is the equivalent of dividing a matrix by itself. An identity matrix has 1s in the diagonal elements and 0s in all of the off-diagonal elements.

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

An identity matrix is analogous to the number 1 in regular arithmetic. To reinforce this concept, the reader should demonstrate for himself or herself that:

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A} . \tag{Eq. A.6}$$

Spending a half an hour or so calculating the above equality, as well as double checking some of the examples presented earlier, will go a long way to instantiating these important concepts.

¹ One final point on exponents. If the exponent is a fraction, it means to take the root of the principal number. Hence, $16^{1/2} = 4$, while $625^{1/2} = 25$.