# Math From Scratch Lesson 43: Matrices

W. Blaine Dowler

April 27, 2014

## Contents



## 1 Introducing Matrices

Now that we have established vectors, one-forms, and other objects that can be defined as having multiple elements along a single column or row, it is not difficult to imagine a more complicated object which has multiple rows or columns. When each row has the same number of elements, and each column has the same number of elements, this object is a matrix.[1] A column with $n$ rows and $m$ columns is an $n \times m$ matrix. For example,

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

is a $3 \times 2$ matrix, as it has 3 (horizontal) rows and 2 (vertical) columns. If you wish to refer to a specific element, or number within, a matrix, then you use two subscripts. Traditionally, the entire matrix is labeled with a capital Latin

[1] To the best of my knowledge, objects whose rows and columns differ in length have no specific name and are unwieldy beasts with few, if any, practical usage.

letter, such as $A$, while the elements are the corresponding lowercase letters with subscripts listing first row and then column. In this example, $a_{21} = 3$, as

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

Three and higher dimensional objects are entirely possible, but are not considered matrices and will not be considered here. If $a_{ij} = b_{ij} \forall i, j$ then $A = B$.

# 2 Operations with Matrices

If matrices are going to be useful in any way, they must satisfy two criteria:

1. They must have applications to... something. It doesn't matter what, and it doesn't need to be anything in the physical realm, but there needs to be something you can do with matrices that you can't do with other math. (These are actually of vital importance to physics, engineering, computer science, and more.)
2. They must have some sort of sensible rules for being manipulated. These manipulations can be our classic operations or new ones, but we have to be able to do something with them.

The rest of this section will deal with the second criterion. We will eventually address both.

## 2.1 Addition and Subtraction

The definitions of addition and subtraction will be fairly intuitive, and very similar to those used for vectors. We will require that any matrices being added are of the same dimensions. Thus, if $A$ is an $n \times m$ matrix and $B$ is also an $n/timesm$ matrix, then we can define $A + B = C$ if $C$ is an $n \times m$ matrix such that $c_{ij} = a_{ij} + b_{ij}$. For example, with

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, B = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

then

$$A + B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 1+1 & 2+4 \\ 3+2 & 4+5 \\ 5+3 & 6+6 \end{pmatrix} = \begin{pmatrix} 2 & 6 \\ 5 & 9 \\ 8 & 12 \end{pmatrix} = C$$

Similarly, with two identically shaped matrices $A$ and $B$, we define $A - B = C$ as $c_{ij} = a_{ij} - b_{ij}$.

For simplicity, we define the 0 matrix as the matrix such that $0_{ij} = 0 \forall i, j$. In this manner,

$$A$$

as expected.

## 2.2 Multiplication and Division

We have options for how we define multiplication and division. We could, for example, define $C = A \times B$ as the matrix whose elements are given by $c_{ij} = a_{ij} \cdot b_{ij}$. In this case, we would then see that each element of a matrix operates in complete isolation, so that every element of every matrix with subscripts, say, 21, would "interact" with every other element with subscripts 21 and only those in all other matrices. This would reduce the matrices to being nothing more than $n \times m$ different calculations being performed in a single dimension. This sounds easy, but it is surprisingly inconvenient, as we'll be missing out on a few applications while really gaining nothing in the process. This intuitive option works for addition, but we will have a far more versatile toolkit to work with if we change the way we define multiplication.

### 2.2.1 Requirements and Definition

To be considered "multiplication," an operation must satisfy the algebraic requirements of multiplication that apply to a particular type of algebra. We will define multiplication in a manner that, with matrices of a certain size or class of sizes, satisfies all of the axioms for an algebraic ring. We can even multiply matrices of different sizes, provided certain conditions are met.

If $A$ is an $n \times m$ matrix, and $B$ is a $p \times q$ matrix, then we can define $A \cdot B = A \times B$ if and only if $m = p$, such that $C = A \cdot B$ is defined by $c_{ij} = \sum_{k=1}^{m} a_{ik} \cdot b_{kj}$. Note that this requires finite dimensional matrices, as we must have finite sums for the axioms of algebra to hold. Infinite matrices are possible, but will have entirely different properties than finite matrices, and will be handled specifically later.

With this definition, the number of columns in the matrix on the left has to match the number of rows in the matrix on the right. This makes it not only possible, but also means that matrices are not only not commutative in all cases, but there are cases in which $A \times B$ is perfectly valid but $B \times A$ is undefined. For example, if $C = A \times B$ then $C_{23}$ is calculated by multiplying each element from row 2 of $A$ and column 3 of $B$, and summing those products as follows:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, B = \begin{pmatrix} 7 & 9 & 11 \\ 8 & 10 & 12 \end{pmatrix}$$

then

$$\begin{aligned} A \times B &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \times \begin{pmatrix} 7 & 9 & 11 \\ 8 & 10 & 12 \end{pmatrix} \\ &= \begin{pmatrix} 1 \times 7 + 2 \times 8 & 1 \times 9 + 2 \times 10 & 1 \times 11 + 2 \times 12 \\ 3 \times 7 + 4 \times 8 & 3 \times 9 + 4 \times 10 & 3 \times 11 + 4 \times 12 \\ 5 \times 7 + 6 \times 8 & 5 \times 9 + 6 \times 10 & 5 \times 11 + 6 \times 12 \end{pmatrix} \\ &= \begin{pmatrix} 7 + 16 & 9 + 20 & 11 + 24 \\ 21 + 32 & 27 + 40 & 33 + 48 \\ 35 + 48 & 45 + 60 & 55 + 72 \end{pmatrix} \\ &= \begin{pmatrix} 23 & 29 & 35 \\ 53 & 67 & 81 \\ 83 & 105 & 127 \end{pmatrix} = \begin{pmatrix} 23 & 29 & 35 \\ 53 & 67 & 81 \\ 83 & 105 & 127 \end{pmatrix} \end{aligned}$$

In contrast,

$$\begin{aligned} B \times A &= \begin{pmatrix} 7 & 9 & 11 \\ 8 & 10 & 12 \end{pmatrix} \times \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 7 \times 1 + 9 \times 3 + 11 \times 5 & 7 \times 2 + 9 \times 4 + 11 \times 6 \\ 8 \times 1 + 10 \times 3 + 12 \times 5 & 8 \times 2 + 10 \times 4 + 12 \times 6 \end{pmatrix} \\ &= \begin{pmatrix} 7 + 27 + 55 & 14 + 36 + 66 \\ 8 + 30 + 60 & 16 + 40 + 72 \end{pmatrix} \\ &= \begin{pmatrix} 89 & 116 \\ 98 & 128 \end{pmatrix} \end{aligned}$$

We can see that $A \times B \neq B \times A$, as the two matrices aren't even the same size, let alone having the same components. This also means that we are limited in terms of setting up matrices with exponent operations. To have $A^2$, we must be able to compute $A \times A$. If $A$ is an $n \times m$ matrix, then computing $A \times A$ only makes sense if $n = m$ and the matrix is square. It is with square matrices that we are able to produce the most robust set of rules and opportunities. With some other groundwork laid, we can start to look at division.

### 2.2.2 Identity

Let $A$ by an arbitrary $n \times n$ matrix. Then $I$ is the $n \times n$ *identity matrix* if $AI = IA = A$. This will operate just as the number 1 operates in multiplication over real numbers. If we are unable to define such an identity for all possible matrices, then we will be unable to use our definition of multiplication for ring or field algebras. (We could use addition to define an additive group, but doing so again eliminates the need for a new object when a collection of existing objects would suffice.)

Let us examine our original definition of multiplication: $C = A \cdot B$ is defined by $c_{ij} = \sum_{k=1}^{m} a_{ik} \cdot b_{kj}$. Thus, if we want $c_{ij} = a_{ij} \forall i, j$, then we need to make sure that $B = I$ is carefully defined. The easiest definition of $I$ (using uppercase $I$ for the component so it is not confused with the index $i$) is this:

$$I_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$$

With this choice of $I$, our multiplication reduces to $c_{ij} = \sum_{k=1}^{m} a_{ik} \cdot I_{kj} = a_{ij}$ and our sum is ensured. We can confirm this as follows for an arbitrary $2 \times 2$ matrix $A$:

$$\begin{aligned} A \times I &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} + 0 & 0 + a_{12} \\ a_{21} + 0 & 0 + a_{22} \end{pmatrix} = A \\ I \times A &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} + 0 & a_{12} + 0 \\ 0 + a_{21} & 0 + a_{22} \end{pmatrix} = A \end{aligned}$$

Note that it is with the summation notation that we have formally proven the form of the identity matrix. The explicit example above only proves the general case for the $2 \times 2$ matrix. The appearance of the identity matrix is consistent at all levels, with 1 along the diagonal joining the upper left corner and the lower right hand corner, and 0 for every other entry. For example,

$$I_{3\times 3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$I_{4\times 4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$I_{n\times n} = \begin{pmatrix} \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & \cdots & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \vdots & \ddots & \vdots \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \cdots & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \end{pmatrix}$$

Now that we have an identity matrix, we can look at the possibility of division.

### 2.2.3 Inverses

Division is always defined through multiplicative inverses. When we cannot guarantee commutativity, then we say $B$ and $A$ are inverses if and only if $AB = BA = I$. What options for division do we have with matrices? First, if $A$ is going to have an inverse, we need to determine the shape of $A$. Let us first demonstrate that the inverse of $A$ would need to be unique. Assume $B$ is an inverse to $A$ from the left so that $BA = I$, and that $C$ is inverse to the right so that $AC = I$. We can then demonstrate that $B = C$ as follows:

$$B = BI = BAC = IC = C$$

If $A$ is a $j \times k$ matrix, $B$ is a $m \times n$ matrix, $C$ is a $p \times q$ matrix, and $I$ is an $r \times r$ matrix (as $I$ must be square), then the multiplications $BI$, $BA$, $AC$ and $IC$ imply that $n = r$, $n = j$, $l = p$ and $r = p$. As we end up with $B = C$, we also know that $m = p$ and $n = q$. Combining all of these equalities, we see that $n = r = j = p = l = m = q$, so all invertible matrices are square matrices.

Now that this is established, we can look at inverses in more detail. We currently have no way to identify the inverse of a matrix, but if we start with a simple $2 \times 2$ matrix and look for a general solution, we may find a pattern.

We begin with

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

and

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$B$ will be the inverse of $A$ if

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{12} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

This is equivalent to simultaneously solving these four equations:

$$\begin{array}{ll} a_{11}b_{11} + a_{12}b_{21} = 1 & a_{11}b_{12} + a_{12}b_{22} = 0 \\ a_{21}b_{11} + a_{22}b_{12} = 0 & a_{21}b_{12} + a_{22}b_{22} = 1 \end{array}$$

which produces the following solution:

$$B = \begin{pmatrix} \frac{a_{22}}{a_{11}a_{22}-a_{12}a_{21}} & \frac{-a_{21}}{a_{11}a_{22}-a_{12}a_{21}} \\ \frac{-a_{12}}{a_{11}a_{22}-a_{12}a_{21}} & \frac{a_{11}}{a_{11}a_{22}-a_{12}a_{21}} \end{pmatrix} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}$$

where we import the definition of scalar multiplication from our work in vectors. This all seems fine and dandy, except there is a caveat: we can now find the inverse of any $2 \times 2$ matrix for which $a_{11}a_{22} - a_{12}a_{21} \neq 0$, or $a_{11}a_{22} \neq a_{12}a_{21}$. Individual entries may be 0, as we see by the fact that the identity matrix is its own inverse, but if this combination (which we will come to know as the *determinant* many, many lessons down the road) is zero, then there is no hope of inverting this matrix. This sounds like an obstacle, but counter-intuitively, the most useful matrices in many practical applications are those which do not have inverses.

# 3 Next Lesson

In our next lesson, we shall explore the idea that matrices can be roots of polynomials in much the same way regular numbers are in the fields of real and complex numbers. This will lead us to operators, eigenvalues, eigenvectors, and ultimately, at long last, to graphing.