1
Matrix Algebra Review

Introduction

This is the matrix review for *A Student's Guide to Vectors and Tensors* (SGVT). It is not meant to be a thorough introduction to the theory and practice of matrix algebra, but rather a review of the matrix concepts you should understand when learning about vectors and tensors.

A matrix is defined as an array of numbers. Some authors restrict the term "matrix" to refer only to square arrays of real numbers, but in common usage a matrix is a rectangular (m x n) array of complex numbers or "elements," where "m" represents the number of rows, "n" represents the number of columns, and "complex" means that the numbers may be purely real, purely imaginary, or mixed. Thus

$$A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 8 \\ 0 \\ -2 \\ 1 \\ 5 \end{bmatrix} \quad C = \begin{bmatrix} 9 & 15 & 3 & -8 & 12 & -2 \end{bmatrix}$$

and

$$D = \begin{bmatrix} 5-2i & 3i \\ 0 & 15 \\ -4i & 2+8i \\ 12 & i-1 \end{bmatrix} \quad E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad F = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

are all matrices. In these examples, matrix A has "dimensions" or "order" (rows x columns) of 2x2, B of 5x1, C of 1x6, D of 4x2, E of 3x3, and F of 2x2.

If both m and n for a matrix are equal to one, that matrix is called a scalar (because it's just a single value), and if either (but not both) of m and n are one, that matrix may be called a vector. So in the examples shown above, B and C are vectors; in some texts you'll see B called a column vector and C called a row vector. A square matrix (that is, a matrix with m=n) with ones on the diagonal and zeroes in all off-diagonal elements (such as matrix E above) is called the "unit" or "identity" matrix, and a matrix with all elements equal to zero is called the "null" or "zero" matrix.

Two matrices are said to be equal only if they have the same number of rows as well as the same number of columns and if every element in one matrix is equal in value to the corresponding element in the other matrix. Matrices are usually denoted using uppercase letters or bold font, often surrounded by square brackets (such as [A]), and the elements are often written using lowercase letters with subscripts. So you may see the elements of matrix [A] written as $\mathrm{a_{ij}}$, although some authors use A_{ij} or $[A]_{ij}$ to refer to the elements of A (but be careful to note that A_{ij} may also refer to the "Matrix of Cofactors" of matrix A, which you can read about in Section E of this document).

So does terminology such as "row vector" mean that matrices, vectors, and tensors are the same thing? Not really. It's certainly true that the vectors and tensors in SGVT are often represented using matrices, but it's important to remember what those matrices stand for. Those arrays of values represent the *components* of vectors and tensors, and those components have meaning only when associated with the basis vectors of a particular coordinate system. Since the basis vectors are not always shown, it's convenient to think of the matrix as representing the vector or tensor itself, and that's fine as long as you remember that the actual vector or tensor has existence independent of any coordinate system.

A) Matrix addition, multiplication by a scalar, and subtraction

Matrices may be added only if both the row dimension (m) and their column dimension (n) are equal (such matrices are said to be "of the same order"). The addition is accomplished simply by adding each element of one matrix to the corresponding element of the other matrix. For example:

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

or in general

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

Note that the result is a matrix of the same order as the matrices being added. Note also that matrix addition is commutative, so [A]+[B]=[B]+[A].

Multiplication of a matrix by a scalar is straightforward; you simply multiply each element of the matrix by the scalar. Thus

$$3A = 3\begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 9 & -6 \\ 3 & 0 \end{bmatrix}$$

and generally

$$kA = k\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{bmatrix}$$

You can use the rules for addition of matrices and scalar multiplication to see that subtraction of matrices is accomplished simply by subtracting the corresponding elements. Thus [A]-[B]=[A]+(-1)[B]. So if

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

and

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

then

$$A-B = \begin{bmatrix} a_{11}+(-1)b_{11} & a_{12}+(-1)b_{12} \\ a_{21}+(-1)b_{21} & a_{22}+(-1)b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}-b_{11} & a_{12}-b_{12} \\ a_{21}-b_{21} & a_{22}-b_{22} \end{bmatrix}.$$

Just as for addition, subtraction of matrices only works for matrices of the same order.

B) Matrix multiplication

There are several different ways to multiply matrices; the most common and most relevant to the vector and tensor concepts in SGVT is to multiply two matrices (call them A and B) by multiplying the elements of each row in A by the elements of each column of B and then summing the results. So if matrix A is given by

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

and matrix B is given by

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

then the product A times B is given by

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \end{bmatrix}.$$

This result is achieved by multiplying the elements of the first row of A by the elements of the first column of B, summing the results, and placing the sum in the first row and first column of the result matrix, as shown in Figure 1.1.

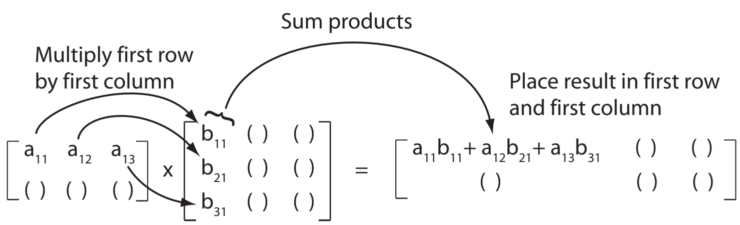

Figure 1.1 Multiplying first row by first column

The next step is to multiply the elements of the first row of A by the elements of the second column of B, summing the products, and placing the sum in the first row and second column of the result matrix, as shown in Figure 1.2.

After multiplying the first row of A by the third column of B and placing the sum in the first row and third column of the result matrix, the same procedure is done with the second row of A - those elements are multiplied by the first column of B, summed, and placed in the second

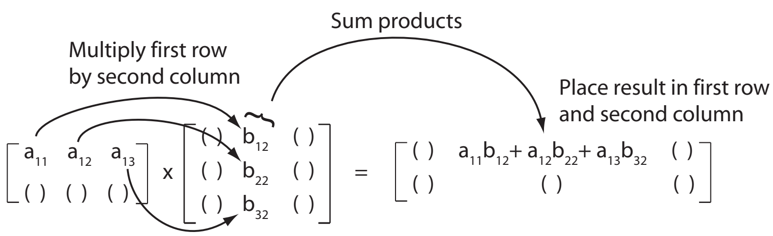

Figure 1.2 Multiplying first row by second column

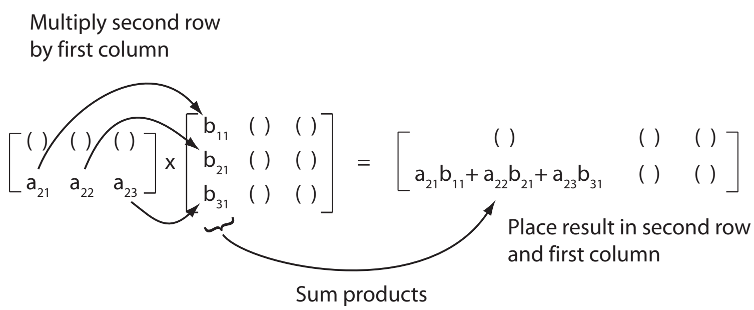

Figure 1.3 Multiplying second row by first column

row and first column of the result matrix, as shown in Figure 1.3.

Notice that the matrix that results from multipling a 2x3 matrix (A) by a 3x3 matrix (B) is a 2x3 matrix - the result has the number of rows of the first matrix and the number of columns of the second matrix. You should also note that this type of matrix multiplication only works when the number of columns of the first matrix (3 in the case of A) equals the number of rows of the second matrix (also 3 in the case of B).

To understand why matrices are multiplied in this way, consider the sorting of five types of toy marbles (Bowlers, Cat's Eyes, Steelies, Aggies, and Commons) into four sizes of package (Small, Medium, Large, and Extra Large). The number of each type of marble in each size package is shown in the following table:

	Bowlers	Cat's Eyes	Steelies	Aggies	Commons
Small package	1	3	2	5	4
Medium package	2	5	3	7	6
Large package	5	10	5	12	10
Extra Large package	10	15	12	16	14

This array of numbers can be put into a matrix - call it "P" the package matrix:

$$P = \begin{bmatrix} 1 & 3 & 2 & 5 & 4 \\ 2 & 5 & 3 & 7 & 6 \\ 5 & 10 & 5 & 12 & 10 \\ 10 & 15 & 12 & 16 & 14 \end{bmatrix}$$

Now imagine three containers (Cans, Boxes, and Crates) with different numbers of packages in each type:

	Small Pkgs	Medium Pkgs	Large Pkgs	Extra-Large Pkgs
Can	10	8	5	2
Box	15	12	7	3
Crate	40	25	15	10

Put this array into a matrix called "C" the container matrix:

$$C = \begin{bmatrix} 10 & 8 & 5 & 2 \\ 15 & 12 & 7 & 3 \\ 40 & 25 & 15 & 10 \end{bmatrix}$$

If you wish to find the number of each type of marble in each type of container, you could do it like this:

Bowlers per Can = 10 small packages $\times$ 1 Bowler per small package
+ 8 medium packages $\times$ 2 Bowlers per medium package
+ 5 large packages $\times$ 5 Bowlers per large package
+ 2 extra-large packages $\times$ 10 Bowlers per extra-large package
= (10)(1)+(8)(2)+(5)(5)+(2)(10)=71 Bowlers per Can

Likewise, you can find the number of Cat's Eyes per can using

Cat's Eyes per Can=10 small packages×3 Cat's Eyes per small package
+ 8 medium packages × 5 Cat's Eyes per medium package
+ 5 large packages × 10 Cat's Eyes per large package
+ 2 extra-large packages × 15 Cat's Eyes per extra-large package
= (10)(3)+(8)(5)+(5)(10)+(2)(15)= 150 Cat's Eyes per Can

And if you wished to find the number of Bowlers per Box, you could use

Bowlers per Box = 15 small packages × 1 Bowler per small package
+ 12 medium packages × 2 Bowlers per medium package
+ 7 large packages × 5 Bowlers per large package
+ 3 extra-large packages × 10 Bowlers per extra-large package
= (15)(1)+(12)(2)+(7)(5)+(3)(10)=104 Bowlers per Box

If you compare the numbers in these calculations to the values in the P and C matrices, you'll see that in each case you're multiplying the row elements of C by the column elements of P, which is exactly how matrix multiplication works:

$$CP = \begin{bmatrix} 10 & 8 & 5 & 2 \\ 15 & 12 & 7 & 3 \\ 40 & 25 & 15 & 10 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 & 5 & 4 \\ 2 & 5 & 3 & 7 & 6 \\ 5 & 10 & 5 & 12 & 10 \\ 10 & 15 & 12 & 16 & 14 \end{bmatrix} = \begin{bmatrix} 71 & 150 & 93 & 198 & 166 \\ 104 & 220 & 137 & 291 & 244 \\ 265 & 545 & 350 & 715 & 600 \end{bmatrix}$$

So the product matrix shows you how many of each type of marble are in each type of container:

	Bowlers	Cat's Eyes	Steelies	Aggies	Commons
Can	71	150	93	198	166
Box	104	220	137	291	244
Crate	265	545	350	715	600

In addition to showing why the rows of the first matrix are multiplied by the columns of the second matrix, this example also illustrates the point made earlier: when you multiply two matrices, the number of columns in the first matrix must equal the number of rows in the second matrix. So in this case, multiplication works only if you multiply C by P, not P by C. The larger point is that matrix multiplication is not commutative, so even if you're able to multiply two matrices in either order (which you can do, for example, with square matrices), the answer is not the same. So in matrix world, AB is in general not equal to BA.

Another difference between matrix multiplication and multiplication of numbers is that it's possible to get zero as the result of multiplying two matrices, even if neither matrix is zero. For example,

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

If you study vectors and tensors, you're likely to come across the multiplication of a row vector times a column vector or vice versa, and you should understand that such products use the same process described above for matrix multiplication, although that may not be immediately obvious when you look at the result. For example, multiplying a row vector A by a column vector B gives

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix} \times \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = [a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}] \, .$$

This scalar result (essentially the dot product between A and B) comes from multiplying the first (and only) row of A by the first (and only) column of B and adding the sums. Likewise, if A is a column vector and B a row vector, the product AB is

$$\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} & b_{13} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} \\ a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} \\ a_{31}b_{11} & a_{31}b_{12} & a_{31}b_{13} \end{bmatrix}$$

Once again, the same rule of matrix multiplication has been applied. The first row of A (just a_{11} in this case) is multiplied by the first column of B (just b_{11} in this case), and since there are no other elements in the first row of A or the first column of B, there is nothing to add, and the result ($a_{11}b_{11}$) is written in the first row and column of the result matrix. Then the first row of A (again, only a_{11}) is multiplied by the second row of B (only b_{12}), and the result is written in the first row, second column of the result matrix. After doing the same for the first row of A and the third column of B, you then multiply the second row of A (which is just a_{21} in this case) by the first column of B (again just b_{11}) and write the result in the second row, first column of the result matrix. So although you get a scalar when you multiply a row vector by a column vector (sometimes called the "inner product" of the matrics) and a matrix when you multiply a column vector by a row vector (sometimes called the "outer product" of the matrices), the process is exactly the same in both cases.

Although matrix multiplication is not commutative (so AB is not necessarily equal to BA), matrix multiplication is the associative and distributive over addition, so for matrices A, B, and C

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

and

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

as long as you remember not to reverse the order of any of the products.

You may be wondering if there's ever a need to multiply each element of one matrix by the corresponding element of an equal-size matrix. There certainly is (for example, when applying a two-dimensional filter funtion to an image). This process is called "element-by-element" or "entrywise" matrix multiplication and the result is sometimes called the "Hadamard product." In MATLAB, such multiplication is denoted A.*B, where the decimal point before the multiplication symbol signifies element-by-element multiplication.

C) Transpose and trace of a matrix

The transpose of a matrix is accomplished by interchanging the rows with the columns of that matrix. This is usually denoted by placing the superscript "T" after the matrix name, so the transpose of matrix A is written as A^T (MATLAB uses A'). So if the matrix A has elements

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix}$$

then the transpose of A is given by

$$A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \\ a_{14} & a_{24} \end{bmatrix}.$$

So $(A_{ij})^T = A_{ji}$; notice that the indices have been switched. The transpose of the product of two matrices is equal to the transpose of each matrix multiplied in reverse order, so

$$(AB)^T = B^T A^T.$$

so long as the dimensions of the matrices allow such products to be formed.

For square matrices, the trace of the matrix is given by the sum of the diagonal elements. The trace is usually denoted by "Tr" so the trace of matrix A is written as Tr(A). So if the matrix A is given by

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

the trace of A is

$$Tr(A) = a_{11} + a_{22} + a_{33} + a_{44} = \sum a_{ii}.$$

D) Determinant, minors, and cofactors of a matrix

The determinant of a square matrix is a scalar calculated by multiplying, adding, and subtracting various elements of the matrix. If A is a 2x2 matrix

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

the determinant of A (denoted using vertical bars on each side of A) is found by cross-multiplying the upper-left element times the lower-right element and then subtracting the product of the lower-left element times the upper-right element:

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Hence if A is

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

the determinant of A is

$$|A| = \begin{vmatrix} 4 & -2 \\ 3 & 1 \end{vmatrix} = (4)(1) - (3)(-2) = 4 - (-6) = 10.$$

To find the determinant of higher-order matrices, you must use the "minors" and "cofactors" of the determinant of the matrix. The minor for each element of the determinant of a square matrix is found by eliminating the entire row and the entire column in which the element appears and then writing the remaining elements as a new determinant, with row and column dimensions reduced by one (so the minor of each element of a 3x3 determinant is a 2x2 determinant). So for the 3x3 determinant of matrix A

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

the minor of element a_{11} is

$$\begin{vmatrix} - & - & - \\ - & a_{22} & a_{23} \\ - & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

obtained by elminating the first row and column of the determinant of matrix A (since the element for which this minor is taken (a_{11}) is in the first row and the first column). Likewise, the minor for element a_{12} is found by eliminating the row and column in which a_{12} appears (first row and second column):

$$\begin{vmatrix} - & - & - \\ a_{21} & - & a_{23} \\ a_{31} & - & a_{33} \end{vmatrix} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}.$$

For element a_{22}, the minor is

$$\begin{vmatrix} a_{11} & - & a_{13} \\ - & - & - \\ a_{31} & - & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}.$$

Before you can use minors to find the determinant of a higher-order matrix, you have turn each minor into a "cofactor." The cofactor of an element is just the minor of that element multiplied by either +1 or -1. To know which of these to use, just determine whether the sum of the element's row index and its column index is even or odd. If even (as it is, for example, for the element in the first row and first column, since 1+1=2 which is even), the cofactor equals the minor of that element mutiplied by +1. If odd (for example, for the element in the first row and second column, since 1+2=3 which is odd) the cofactor equals the minor of that element multiplied by -1. So for element a_{ij} with row index of i and column index of j,

$$\text{Cofactor of } (a_{ij}) = (-1)^{(i+j)} \text{Minor of } (a_{ij}).$$

So for the minors shown above, the cofactors are

$$\text{Cofactor of } (a_{11}) = (-1)^{(1+1)} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = (1)(a_{22}a_{33} - a_{32}a_{23})$$

$$\text{Cofactor of } (a_{12}) = (-1)^{(1+2)} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = (-1)(a_{21}a_{33} - a_{31}a_{23})$$

and

$$\text{Cofactor of } (a_{22}) = (-1)^{(2+2)} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = (1)(a_{11}a_{33} - a_{31}a_{13}).$$

With cofactors in hand, the determinant of a 3x3 matrix is straightforward to find. Simply pick one row or column, and multiply each of the elements of that row or column by its cofactor and sum the results. So, for example, if you choose to use the first row of the 3x3 matrix A, the determinant is given by

$$|A| = a_{11}(\text{Cofactor of } a_{11}) + a_{12}(\text{Cofactor of } a_{12}) + a_{13}(\text{Cofactor of } a_{13}).$$

or

$$|A| = a_{11}(1)(a_{22}a_{33} - a_{32}a_{23}) + a_{12}(-1)(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(1)(a_{21}a_{32} - a_{31}a_{22}).$$

The value of the determinant would have been the same if you had used a different row (or any of the columns) of matrix A, as long as you multiply each element in your selected row or column by its cofactor and sum the results.

The same approach can be used to find the determinant of 4x4 and higher-order matrices; you just have to expand each of the determinants using elements and cofactors until you get down to the 2x2 level. Of course, the calculation becomes a bit tedious, so it's worth your time to learn to find determinants using a program like MATLAB.

Some useful characteristics of determinants are that the determinant of the transpose of a matrix equals the determinant of the matrix (so $|A^T| = |A|$) and the determinant of the product of two matrices is the same as the determinant of the reverse product (so $|AB| = |BA|$) provided that the dimensions of A and B allow both of these products to be made.

You may also find it useful to know that if a matrix has two identical rows or columns, or if one row or column is an integer multiple of another row or column, the determinant of that matrix must be zero.

E) Inverse of a matrix

As indicated in Section C of this document, matrix multiplication is quite different from ordinary multiplication of numbers. One example of this is that the product of two matrices may be zero even if both of the matrices are non-zero (so you can't simply divide both sides of a matrix equation such as $AB = 0$ by A to get $B = 0$). Additionally, the matrix equation $AB = AC$ does not mean that matrix B equals matrix C (so dividing both sides of this equation by A does not work for matrices).

Differences such as this suggest that matrix division has little in common with ordinary division of numbers. About the closest you can get to a process similar to division for matrices comes about by considering the matrix equation

$$AX = B$$

where A, X, and B are all matrices. If you wish to find matrix X from this equation, here's one thing you definitely cannot do:

$$\begin{aligned} \frac{AX}{A} &= \frac{B}{A} \\ X &= \frac{B}{A} \end{aligned}$$

because this type of division does not work for matrices.

What you can do to find X is this: you can try to find a matrix A^{-1} (called the "inverse" or "reciprocal" of A) with the following property:

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

where I represents the identity matrix (ones on the diagonal and zeroes everywhere else). Now you can find X rather easily:

$$\begin{aligned} AX &= B \\ A^{-1}(AX) &= A^{-1}B \\ (A^{-1}A)X &= A^{-1}B \\ IX &= A^{-1}B \\ X &= A^{-1}B \end{aligned}$$

since $IX = X$ (multiplying the identity matrix times any matrix does not change that matrix).

So although you haven't really divided matrix B by matrix A, you've used the matrix equivalent of multiplying by the reciprocal of a number to achieve the same result as dividing.

The question is, does every matrix have an inverse (that is, another matrix with which it multiplies to give the identiry matrix), and if so, how do you find it? The first part of this question is easy to answer: any matrix that has a non-zero determinant ($|A| \neq 0$) has an inverse. Matrices that have no inverse (($|A| = 0$) are called "singular" matrices, and matrices that have an inverse are called "non-singular" matrices.

But if a certain matrix has an inverse, how do you find that inverse? There are several ways to go about this, but one approach is to use the concepts of cofactors, transpose, and determinant described above. Those concepts appear in the following equation for the inverse of matrix A:

$$A^{-1} = \frac{(\text{Matrix of cofactors of A})^T}{|A|}$$

in which the "Matrix of cofactors of A" is a matrix in which each element of A is replaced by that element's cofactor.

Here's how that works for a 3x3 matrix. If matrix A is the usual

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

then the matrix of cofactors (MoC) for A is

$$\text{MoC } (A) = \begin{bmatrix} a_{22}a_{33} - a_{32}a_{23} & -(a_{21}a_{33} - a_{31}a_{23}) & a_{21}a_{32} - a_{31}a_{22} \\ -(a_{12}a_{33} - a_{32}a_{13}) & a_{11}a_{33} - a_{31}a_{13} & -(a_{11}a_{32} - a_{31}a_{12}) \\ a_{12}a_{23} - a_{22}a_{13} & -(a_{11}a_{23} - a_{21}a_{13}) & a_{11}a_{22} - a_{21}a_{12} \end{bmatrix}.$$

The equation for the inverse of A shown above requires that you take the transpose of this matrix, which is

$$[\text{MoC } (A)]^T = \begin{bmatrix} a_{22}a_{33} - a_{32}a_{23} & -(a_{12}a_{33} - a_{32}a_{13}) & a_{12}a_{23} - a_{22}a_{13} \\ -(a_{21}a_{33} - a_{31}a_{23}) & a_{11}a_{33} - a_{31}a_{13} & -(a_{11}a_{23} - a_{21}a_{13}) \\ a_{21}a_{32} - a_{31}a_{22} & -(a_{11}a_{32} - a_{31}a_{12}) & a_{11}a_{22} - a_{21}a_{12} \end{bmatrix}.$$

Dividing this matrix by the determinant of A provides the inverse of A.

Fortunately, the inverse of non-singular matrices is easily found using most scientific calculators or computers, usually by defining a matrix A and then raising A to the power of minus one.

For a diagonal matrix, the inverse is simply another diagonal matrix in which each diagonal element is the reciprocal of the corresponding element in the original matrix. So

$$\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & 0 & 0 \\ 0 & \frac{1}{a_{22}} & 0 \\ 0 & 0 & \frac{1}{a_{33}} \end{bmatrix}.$$

F) Simultaneous linear equations and Cramer's Rule

The relationship between matrices and simultaneous linear equations can be understood by considering equations such as

$$\begin{aligned} 2x + 5y - z &= 12 \\ -3x - 3y + z &= -1 \\ x + y &= 1. \end{aligned}$$

Using the rules of matrix multiplication, this system of equations can be written as a single matrix equaton:

$$\begin{bmatrix} 2 & 5 & -1 \\ -3 & -3 & 1 \\ 1 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 12 \\ -1 \\ 1 \end{bmatrix}.$$

In general, three linear equations in three unknowns (x_1, x_2, x_3) can be written as

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

or

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

If you define the matrices

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

and

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

and

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

then the system of equations can be written as a single matrix equation:

$$Ax = b$$

which can sometimes be solved using the inverse of matrix A:

$$x = A^{-1}b.$$

The reason that the word "sometimes" appears in the previous sentence is that there are several conditions which will prevent this approach from succeeding. For example, the system of equations may be inconsistent - that is, there may be no set of values for x_1, x_2, x_3 that satisfy all three equations. In that case, you will not be able to find the inverse of A because it will be a singular matrix. And if matrix b equals zero, you have a system of homogeneous linear equations, which means there will be only the trivial solution ($x_1 = x_2 = x_3 = 0$) if A is non-singular or an infinite number of solutions if A is singular.

In cases for which A is non-singular and b does not equal zero, you can find the values of x_1, x_2, and x_3 by finding the inverse of matrix A and multiplying that inverse by matrix b, or you can use Cramer's Rule. In that approach, the first unknown (x_1 in this case) is found by replacing the values in the first column of the coefficient matrix (A) with the elements of matrix b and dividing the determinant of that matrix by the determinant of A. Here's how that looks:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

Likewise, to find the second unknonwn (x_2 in this case), replace the values in the second column of A with the elements of b:

$$x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

and to find x_3 use

$$x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

Thus for the equations given at the start of this section,

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

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 12 \\ -1 \\ 1 \end{bmatrix}.$$

Hence

$$x = x_1 = \frac{\begin{vmatrix} 12 & 5 & -1 \\ -1 & -3 & 1 \\ 1 & 1 & 0 \end{vmatrix}}{\begin{vmatrix} 2 & 5 & -1 \\ -3 & -3 & 1 \\ 1 & 1 & 0 \end{vmatrix}}$$

so

$$x = \frac{12[(-3)(0)-(1)(1)]+5(-1)[(-1)(0)-(1)(1)]-1[(-1)(1)-(1)(-3)]}{2[(-3)(0)-(1)(1)]+5(-1)[(-3)(0)-(1)(1)]-1[(-3)(1)-(1)(-3)]}$$
$$= \frac{-12+5-2}{-2+5-0} = -3.$$

Proceeding in the same way for y and z gives

$$y = x_2 = \frac{\begin{vmatrix} 2 & 12 & -1 \\ -3 & -1 & 1 \\ 1 & 1 & 0 \end{vmatrix}}{\begin{vmatrix} 2 & 5 & -1 \\ -3 & -3 & 1 \\ 1 & 1 & 0 \end{vmatrix}}$$

so

$$y = \frac{-2+12+2}{-2+5-0} = 4$$

and

$$z = x_3 = \frac{\begin{vmatrix} 2 & 5 & 12 \\ -3 & -3 & -1 \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 5 & -1 \\ -3 & -3 & 1 \\ 1 & 1 & 0 \end{vmatrix}}$$

so

$$z = \frac{-4+10+0}{-2+5-0} = 2.$$

G) Matrix diagonalization using eigenvectors and eigenvalues

In the matrix equation discussed in the previous section

$$Ax = b$$

the matrix A operates on matrix x to produce matrix b. If x and b are vectors, this equation represents the transformation by matrix A of vector x into vector b. In some cases, the operation of A on the components of x produces the components of a vector b that is a scaled (but not rotated) version of x. In such cases, the equation becomes

$$Ax = \lambda x$$

where λ represents a scalar multiplier (and scalar multipliers can change the length but not the direction of a vector).

Any vector x which satisfy this equation for a matrix A is called an "eigenvector" of matrix A, and the scalar λ is called the "eigenvalue" associated with that eigenvector.

The eigenvalues of a matrix can be very useful in finding a diagonal version of that matrix, so you may wish to understand how to find the eigenvalues of a given matrix. To do that, write the previous equation as

$$Ax - \lambda x = 0$$

which, since $Ix = x$, can be written as

$$Ax - \lambda(Ix) = 0$$

or

$$(A - \lambda I)x = 0.$$

which means that either $x = 0$ (which is a the trivial case) or

$$|A - \lambda I| = 0.$$

This equation is called the "characteristic equation" for matrix A, and for a 3x3 matrix it looks like this:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - \lambda \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 0$$

or

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0.$$

This expands to

$$\begin{aligned} (a_{11} - \lambda)&[(a_{22} - \lambda)(a_{33} - \lambda) - a_{32}a_{23}] \\ &+ a_{12}(-1)[a_{21}(a_{33} - \lambda) - a_{31}a_{23}] \\ &+ a_{13}[a_{21}a_{32} - a_{31}(a_{22} - \lambda)] = 0. \end{aligned}$$

Finding the roots of this polynomial provides the eigenvalues (λ) for matrix A, and substituting those values back into the matrix equation $Ax = \lambda x$ allows you to find eigenvectors corresponding to each eigenvalue. The process of finding the roots is less daunting than it may appear, as you can see by considering the following example. For the 3x3 matrix A given by

$$A = \begin{bmatrix} 4 & -2 & -2 \\ -7 & 5 & 8 \\ 5 & -1 & -4 \end{bmatrix}$$

the characteristic equation is

$$\begin{vmatrix} 4-\lambda & -2 & -2 \\ -7 & 5-\lambda & 8 \\ 5 & -1 & -4-\lambda \end{vmatrix} = 0$$

or

$$\begin{aligned}(4-\lambda)&[(5-\lambda)(-4-\lambda)-(-1)(8)] \\ &-2(-1)[(-7)(-4-\lambda)-(5)(8)] \\ &-2[(-7)(-1)-(5)(5-\lambda)]=0.\end{aligned}$$

Multiplying through and subtracting within the square brackets makes this

$$(4-\lambda)(\lambda^2-\lambda-12)+2(7\lambda-12)-2(5\lambda-18)=0$$

or

$$-\lambda^3+5\lambda^2+12\lambda-36=0.$$

Finding the roots of a polynomial like this is probably best left to a computer, but if you're lucky enough to have a polynomial with integer roots, you know that each root must be a factor of the term not involving λ (36 in this case). So (+/-) 2,3,4,6,9,12,and 18 are possibilities, and it turns out that +2 works just fine:

$$-(2)^3+5(2^2)+12(2)-36=-8+20+24-36=0.$$

So you know that one root of the characteristic equation (and hence one eigenvalue) must be +2. That means you can divide a factor of $(\lambda-2)$ out of the equation and try to see other roots in the remainder. That division yields this:

$$\frac{-\lambda^3+5\lambda^2+12\lambda-36}{(\lambda-2)}=-\lambda^2+3\lambda+18.$$

The roots remaining polynomial on the right-hand side of this equation are +6 and -3, so you now have

$$-\lambda^3+5\lambda^2+12\lambda-36=(\lambda-2)(6-\lambda)(\lambda+3)=0.$$

So matrix A has three distinct eigenvalues with values +6, -3, and +2; these are the factors by which matrix A scales its eigenvectors. You could find the eigenvectors of A by plugging each of the eigenvalues back into the characteristic equation for A, but as long as you can find N distinct eigenvalues for an NxN matrix, you can be sure that A can be diagonalized simply by constructing a new diagonal matrix with the eigenvalues as the diagonal elements. So in this case, the diagonal matrix (call it A') associated with matrix A is

$$A' = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

To see why this is true, consider the operation of matrix A on each of its three eigenvectors (call them e, f, and g):

$$\begin{aligned} Ae &= \lambda_1 e \\ Af &= \lambda_2 f \\ Ag &= \lambda_3 g \end{aligned}$$

Now imagine a matrix E whose columns are made up of the eigenvectors of matrix A:

$$E = \begin{bmatrix} e_1 & f_1 & g_1 \\ e_2 & f_2 & g_2 \\ e_3 & f_3 & g_3 \end{bmatrix}$$

where the components of eigenvector e are (e_1, e_2, e_3), the components of eigenvector f are (f_1, f_2, f_3), and the components of eigenvector g are (g_1, g_2, g_3). Mutiplying matrix A (the original matrix) by E (the matrix made up of the eigenvectors of A), you get

$$AE = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \times \begin{bmatrix} e_1 & f_1 & g_1 \\ e_2 & f_2 & g_2 \\ e_3 & f_3 & g_3 \end{bmatrix}$$

which is

$$AE = \begin{bmatrix} a_{11}e_1 + a_{12}e_2 + a_{13}e_3 & a_{11}f_1 + a_{12}f_2 + a_{13}f_3 & a_{11}g_1 + a_{12}g_2 + a_{13}g_3 \\ a_{21}e_1 + a_{22}e_2 + a_{23}e_3 & a_{21}f_1 + a_{22}f_2 + a_{23}f_3 & a_{21}g_1 + a_{22}g_2 + a_{23}g_3 \\ a_{31}e_1 + a_{32}e_2 + a_{33}e_3 & a_{31}f_1 + a_{32}f_2 + a_{33}f_3 & a_{31}g_1 + a_{32}g_2 + a_{33}g_3 \end{bmatrix}$$

The columns of this AE matrix are the result of multiplying A by each of the eigenvectors. But you know from the definition of eigenvectors and eigenvalues that

$$Ae = \lambda_1 e = \lambda_1 \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 e_1 \\ \lambda_1 e_2 \\ \lambda_1 e_3 \end{bmatrix}$$

and

$$Af = \lambda_2 f = \lambda_2 \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} \lambda_2 f_1 \\ \lambda_2 f_2 \\ \lambda_2 f_3 \end{bmatrix}$$

and

$$Ag = \lambda_3 g = \lambda_3 \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} \lambda_3 g_1 \\ \lambda_3 g_2 \\ \lambda_3 g_3 \end{bmatrix}.$$

This means that the product AE can be written

$$AE = \begin{bmatrix} \lambda_1 e_1 & \lambda_2 f_1 & \lambda_3 g_1 \\ \lambda_1 e_2 & \lambda_2 f_2 & \lambda_3 g_2 \\ \lambda_1 e_3 & \lambda_2 f_3 & \lambda_3 g_3 \end{bmatrix}.$$

But the matrix on the right-hand side can also be written like this:

$$\begin{bmatrix} \lambda_1 e_1 & \lambda_2 f_1 & \lambda_3 g_1 \\ \lambda_1 e_2 & \lambda_2 f_2 & \lambda_3 g_2 \\ \lambda_1 e_3 & \lambda_2 f_3 & \lambda_3 g_3 \end{bmatrix} = \begin{bmatrix} e_1 & f_1 & g_1 \\ e_2 & f_2 & g_2 \\ e_3 & f_3 & g_3 \end{bmatrix} \times \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

This means that you can write

$$AE = E \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

and multiplying both sides by the inverse of matrix E (E^{-1}) gives

$$E^{-1}AE = E^{-1}E \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

If you tracked the discussion of the similarity transform in Section 6.1, you'll recognize the expression $E^{-1}AE$ as the similarity transform of matrix A to a coordinate system with basis vectors that are the columns of matrix E. Those columns are the eigenvectors of matrix A, and the matrix that results from the similarity transform (call it A') is diagonal and has the eigenvalues of A as its diagonal elements.

References

Pettofrezzo, A., *Matrices and Transformations*, Prentice-Hall 1966.

Riley, K., Hobson, M.P., and Bence, S.J. *Mathematical Methods for Physics and Enginnering*, Cambridge 2006.

Schwartz, J.T., *Introduction to Matrices and Vectors*, McGraw-Hill 1961.