

MATH 223: Some results for  $2 \times 2$  matrices.

Richard Anstee

One can preview a lot of the theory in this course by looking at the special case of  $2 \times 2$  matrices. The proofs are relatively easy in this limited context and not all the complexity is seen. In particular, the Gaussian Elimination algorithm doesn't show its full complexity in this context nor do vector spaces. But  $2 \times 2$  matrices are very concrete and we will be able to introduce many topics such as determinants, inverses, linear transformations, diagonalization.

Here are some  $2 \times 2$  matrices:

$$\begin{array}{ccc} \begin{bmatrix} a & b \\ c & d \end{bmatrix}, & \begin{bmatrix} 1 & 3 \\ 5 & -2 \end{bmatrix} \\ & & \\ & \text{column 1} & \text{column 2} \\ & \downarrow & \downarrow \\ \text{row 1} \rightarrow & a_{11} & a_{12} & , & \begin{bmatrix} \pi & 0 \\ 6.7 & e \end{bmatrix} \\ \text{row 2} \rightarrow & a_{21} & a_{22} & & \end{array}$$

The matrices have two rows and two columns, hence  $2 \times 2$ . We always give indices in (row,column) form.

What are matrices good for? They are remarkably good for many things. The first use is as a function taking 2-tuples (vectors in the plane, for example) to other 2-tuples.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$A\mathbf{x} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.$$

Please note that in this course vectors are always given as columns (so called *column vectors*) and this standard notation makes manipulation of vectors as matrices ( $2 \times 1$  matrices) more straightforward. We will see that many interesting functions of vectors in the plane, such as rotations, can be represented in the above way through matrices. If we have  $b = c = 0$ , then

$$A\mathbf{x} = \begin{bmatrix} ax \\ dy \end{bmatrix}$$

which is like a kind of scalar multiplication. In that case

$$A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

is called a *diagonal* matrix.

If we have a matrix  $B$ ,

$$B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

then we can think of  $B$  as two column vectors

$$\begin{bmatrix} e \\ g \end{bmatrix}, \quad \begin{bmatrix} f \\ h \end{bmatrix}$$

and can try to give the product of two matrices using the formula for  $A\mathbf{x}$ :

$$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

(Note: we could have written this as  $A \cdot B$ ). With a definition of matrix multiplication, we can think about  $2 \times 2$  matrices as numbers but we will be wary about the differences. Addition has a straightforward definition:

$$A + B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a + e & b + f \\ c + g & d + h \end{bmatrix}$$

One can verify the following:

$$A + B = B + A; \text{ (commutativity)}$$

$$A + (B + C) = (A + B) + C; \text{ (associativity)}$$

$$A + 0 = A \text{ where } 0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$A + (-A) = 0 \text{ where } -A = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}$$

Not everything is OK since in general  $AB \neq BA$ :

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

This example even has  $AB = 0$  with neither  $A = 0$  or  $B = 0$ . This not our familiar multiplication.

If we consider diagonal matrices special things happen:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & 0 \\ 0 & h \end{bmatrix} = \begin{bmatrix} ae & bh \\ ce & dh \end{bmatrix}$$

If we take  $e = h$  we get

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & 0 \\ 0 & e \end{bmatrix} = \begin{bmatrix} ae & be \\ ce & de \end{bmatrix} = e \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Of course I am inventing the notation  $eA$  to denote this multiplication of  $A$  by the scalar  $e$  in analogy with notation  $e\mathbf{x}$  denoting multiplying vector  $\mathbf{x}$  by the scalar  $e$ . Special things happen with  $e = h = 1$ :

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

It makes sense to define

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

so that  $I$  acts as 1 does in ordinary multiplication, namely  $AI = A$ . Of course we need to check  $IA = A$ . Now

$$\begin{bmatrix} e & 0 \\ 0 & h \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ea & eb \\ hc & hd \end{bmatrix}$$

namely that right multiplication by a diagonal matrix is multiplying the rows by scalars and so  $IA = A$ .

There are other laws to be verified

$A(BC) = (AB)C$ ; *associativity*

$A(B + C) = AB + AC$ ; *distributive law*

$(A + B)C = AC + BC$ ; *distributive law*

I'll delay demonstrating the associative law, but you could try verifying it using three arbitrary matrices, which is not so difficult, just tedious. The latter two distributive laws suggest the following idea. Let  $E_{ij}$  = matrix with a 1 in position  $i, j$  and 0's elsewhere. Then we have

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = aE_{11} + bE_{12} + cE_{21} + dE_{22},$$

and so using matrix multiplication of  $AB$  with  $B = eE_{11} + fE_{12} + gE_{21} + hE_{22}$ , we need only consider products of the form  $E_{ij}E_{kl}$  in the product  $(aE_{11} + bE_{12} + cE_{21} + dE_{22})(eE_{11} + fE_{12} + gE_{21} + hE_{22})$ . What are the products  $E_{ij}E_{kl}$ ? Can you deduce the general formula?