Unit 2: Introduction to Matrix Algebra

4.2.1

The main aim of this exercise is to help you become more at ease with the idea of the "game" of matrices. To this end, we wish to use our basic definitions to arrive at certain structural properties of matrices. Rather than run the risk of becoming too abstract we have chosen the special case of 2 x 2 matrices since in this case the actual computations are not overwhelming (although they are a bit tedious) and at the same time the procedure used in the 2 x 2 case generalizes rather easily to the case of n x n matrices. Hopefully, the interested reader will be able to see this on his own if he so desires.

To begin with, let us denote A, B, and C as specifically (and yet as generally) as possible by letting

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \qquad (1)$$

Our approach will be to compute A(B + C) and AB + AC, then show that these two matrices are equal term by term; hence equal by the definition of equality.

In a similar way we shall compare A(BC) with (AB)C.

Using the notation in (1) we have

$$A(B + C) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \left[\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} + \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}\right]. \qquad (2)$$

By our definition of matrix addition we have that

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} + \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} b_{11} + c_{11} & b_{12} + c_{12} \\ b_{21} + c_{21} & b_{22} + c_{22} \end{pmatrix}^*$$

* Notice that this verifies the rule of closure that the sum of two 2 x 2 matrices is a 2 x 2 matrix.

4.2.1 (continued)

and substituting this result into (2) yields

$$A(B + C) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} + c_{11} & b_{12} + c_{12} \\ b_{21} + c_{21} & b_{22} + c_{22} \end{pmatrix}$$

$$= \begin{bmatrix} (a_{11}, a_{12}) \cdot (b_{11} + c_{11}, b_{21} + c_{21})^* & (a_{11}, a_{12}) \cdot (b_{12} + c_{12}, b_{22} + c_{22}) \\ (a_{21}, a_{22}) \cdot (b_{11} + c_{11}, b_{21} + c_{21}) & (a_{21}, a_{22}) \cdot (b_{12} + c_{12}, b_{22} + c_{22}) \end{bmatrix}$$

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

$$= \begin{bmatrix} a_{11}b_{11} + a_{11}c_{11} + a_{12}b_{21} + a_{12}c_{21} & a_{11}b_{12} + a_{11}c_{12} + a_{12}b_{22} + a_{12}c_{22} \\ a_{21}b_{11} + a_{21}c_{11} + a_{22}b_{21} + a_{22}c_{21} & a_{21}b_{12} + a_{21}c_{12} + a_{22}b_{22} + a_{22}c_{22} \end{bmatrix} . \qquad (3)$$

Therefore,

$$A(B + C) = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{11}c_{11} + a_{12}c_{21} & a_{11}b_{12} + a_{12}b_{22} + a_{11}c_{12} + a_{12}c_{22} \\ a_{21}b_{11} + a_{22}b_{21} + a_{21}c_{11} + a_{22}c_{21} & a_{21}b_{12} + a_{22}b_{22} + a_{21}c_{12} + a_{22}c_{22} \end{bmatrix} . \qquad (4)$$

[Where the only difference between (4) and (3) is that in (4) we have regrouped terms to suggest the form AB + AC which we shall investigate next]

* Remember that $b_{11} + c_{11}$ is _a_ number as is $b_{21} + c_{21}$. Thus we are simply employing our usual rule for matrix multiplication; in this case, dotting the first row of A with the first column of B + C.

4.2.1 (continued)

On the other hand

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

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}^{*}, \qquad (5)$$

while

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

$$= \begin{bmatrix} a_{11}c_{11} + a_{12}c_{21} & a_{11}c_{12} + a_{12}c_{22} \\ a_{21}c_{11} + a_{22}c_{21} & a_{21}c_{12} + a_{22}c_{22} \end{bmatrix}. \qquad (6)$$

Combining (5) and (6), we have

$$AB + AC = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} + \begin{bmatrix} a_{11}c_{11} + a_{12}c_{21} & a_{11}c_{12} + a_{12}c_{22} \\ a_{21}c_{11} + a_{22}c_{21} & a_{21}c_{12} + a_{22}c_{22} \end{bmatrix} \qquad (7)$$

From our definition of matrix addition, (7) becomes

$$AB + AC = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{11}c_{11} + a_{12}c_{21} & a_{11}b_{12} + a_{12}b_{22} + a_{11}c_{12} + a_{12}c_{22} \\ a_{21}b_{11} + a_{22}b_{21} + a_{21}c_{11} + a_{22}c_{21} & a_{21}b_{12} + a_{22}b_{22} + a_{21}c_{12} + a_{22}c_{22} \end{bmatrix}. \qquad (8)$$

Comparing (8) with (4) and recalling our definition of matrix equality (i.e., the matrices must be equal entry by entry) it follows that

$$A(B + C) = AB + AC. \qquad (9)$$

* From now on we shall omit the step of writing a matrix product in terms of dot products of n tuples and write the result directly. The student who still feels insecure should supply the missing step.

4.2.1 (continued)

[Notice that the more experienced student could have derived (9) directly from (4) by observing that

$$\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{11}c_{11} + a_{12}c_{21} & a_{11}b_{12} + a_{12}b_{22} + a_{11}c_{12} + a_{12}c_{22} \\ a_{21}b_{11} + a_{22}b_{21} + a_{21}c_{11} + a_{22}c_{21} & a_{21}b_{12} + a_{22}b_{22} + a_{21}c_{12} + a_{22}c_{22} \end{pmatrix}$$

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

$$+ \begin{pmatrix} a_{11}c_{11} + a_{12}c_{21} & a_{11}c_{12} + a_{12}c_{22} \\ a_{21}c_{11} + a_{22}c_{21} & a_{21}c_{12} + a_{22}c_{22} \end{pmatrix}$$

$$= AB + AC].$$

As for verifying that $A(BC) = (AB)C$, we have

$$BC = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} b_{11}c_{11} + b_{12}c_{21} & b_{11}c_{12} + b_{12}c_{22} \\ b_{21}c_{11} + b_{22}c_{21} & b_{21}c_{12} + b_{22}c_{22} \end{pmatrix}. \qquad (10)$$

Hence

$$A(BC) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11}c_{11} + b_{12}c_{21} & b_{11}c_{12} + b_{12}c_{22} \\ b_{21}c_{11} + b_{22}c_{21} & b_{21}c_{12} + b_{22}c_{22} \end{pmatrix}$$

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

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

4.2.1 (continued)

On the other hand from (5),

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

Hence,

$$(AB)C = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

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

$$= \begin{bmatrix} a_{11}b_{11}c_{11} + a_{12}b_{21}c_{11} + a_{11}b_{12}c_{21} + a_{12}b_{22}c_{21} & a_{11}b_{11}c_{12} + a_{12}b_{21}c_{12} + a_{11}b_{12}c_{22} + a_{12}b_{22}c_{22} \\ a_{21}b_{11}c_{11} + a_{22}b_{21}c_{11} + a_{21}b_{12}c_{21} + a_{22}b_{22}c_{21} & a_{21}b_{11}c_{12} + a_{22}b_{21}c_{12} + a_{21}b_{12}c_{22} + a_{22}b_{22}c_{22} \end{bmatrix} \quad (12)$$

or more suggestively,

$$(AB)C = \begin{bmatrix} a_{11}b_{11}c_{11} + a_{11}b_{12}c_{21} + a_{12}b_{21}c_{11} + a_{12}b_{22}c_{21} & a_{11}b_{11}c_{12} + a_{11}b_{12}c_{22} + a_{12}b_{21}c_{12} + a_{12}b_{22}c_{22} \\ a_{21}b_{11}c_{11} + a_{21}b_{12}c_{21} + a_{22}b_{21}c_{11} + a_{22}b_{22}c_{21} & a_{21}b_{11}c_{12} + a_{21}b_{12}c_{22} + a_{22}b_{21}c_{12} + a_{22}b_{22}c_{22} \end{bmatrix}. \quad (13)$$

Comparing (11) with (13) shows us that

$$A(BC) = (AB)C \quad (14)$$

4.2.2

a. Letting $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ we see that $AB = 0$ means

4.2.2 (continued)

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \qquad (1)$$

Now

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} b_{11}+ b_{21} & b_{12} + b_{22} \\ b_{11}+ b_{21} & b_{12} + b_{22} \end{pmatrix}. \qquad (2)$$

Substituting the result of (2) into (1) yields

$$\begin{pmatrix} b_{11} + b_{21} & b_{12} + b_{22} \\ b_{11} + b_{21} & b_{12} + b_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \qquad (3)$$

Since matrix equality means that the matrices must be equal entry by entry, equation (3) implies that

$$b_{11} + b_{21} = 0 \qquad (4)$$

and

$$b_{12} + b_{22} = 0 \ . \qquad (5)$$

Equation (4) is obeyed provided only that $b_{21} = -b_{11}$ while equation (5) is obeyed provided only that $b_{22} = - b_{12}$.

This tells us, for example, that b_{11} and b_{12} may be chosen <u>completely</u> <u>at</u> <u>random</u> while b_{21} and b_{22} must then be given by $b_{21} = - b_{11}$ and $b_{22} = - b_{12}$.

In other words, if AB = 0, then

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ -b_{11} & -b_{12} \end{pmatrix}. \qquad (6)$$

To check this result, observe that with B as in (6)

4.2.2 (continued)

$$AB = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ -b_{11} & -b_{12} \end{pmatrix}$$

$$= \begin{pmatrix} b_{11} - b_{11} & b_{12} - b_{12} \\ b_{11} - b_{11} & b_{12} - b_{12} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

What this Exercise shows us is that it is quite possible for the product of two matrices to be the zero matrix even if neither of the factors is the zero matrix. In this particular example, we showed that if

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

then if B was any matrix, such that, entry by entry, the second row was the negative of the first row, the product AB would be the zero matrix. In other words, there are infinitely many different matrices B for which AB = 0, even though neither A nor B need be the zero matrix.

b. Here we wish to emphasize the problems that occur because matrix multiplication is not commutative. We have just seen that for AC to be the zero matrix, then C must have the form

$$C = \begin{pmatrix} x & y \\ -x & -y \end{pmatrix} \tag{7}$$

where x and y are any real numbers chosen at random.

If we now use (7) to compute CA, we obtain

$$CA = \begin{pmatrix} x & y \\ -x & -y \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} x + y & x + y \\ -x - y & -x - y \end{pmatrix} \tag{8}$$

4.2.2 (continued)

and from (8) we see that for CA to be the zero matrix, both x + y and -x -y must be 0, and this in turn means that y must be -x.

In other words, if CA is also to be the zero matrix, we need the additional constraint that only one entry of C can be chosen at random. In particular, in terms of the notation in (7), we must have that

$$C = \begin{pmatrix} x & -x \\ -x & x \end{pmatrix}. \tag{9}$$

With C as defined in (9) both AC and CA are equal to the zero matrix. As a check we have

$$AC = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x & -x \\ -x & x \end{pmatrix} = \begin{pmatrix} x & -x & -x + x \\ x & -x & -x + x \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

while

$$CA = \begin{pmatrix} x & -x \\ -x & x \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} x - x & x - x \\ -x + x & -x + x \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

The important thing to observe, however, in this case, is that the set of matrices B for which AB = 0 is not the same as the set of matrices C for which CA = 0. This result is quite different from the corresponding result of ordinary multiplication and also of, say, dot product multiplication where it is true that $\underline{A} \cdot \underline{B}$ can be 0 even though neither $\underline{A}$ nor $\underline{B}$ is $\underline{0}$, but in this case it is also true that $\underline{A} \cdot \underline{B} = \underline{B} \cdot \underline{A}$.

4.2.3

In this case letting

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

4.2.3 (continued)

we have that AB = 0 means

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} b_{11} + b_{21} & b_{12} + b_{22} \\ b_{11} - b_{21} & b_{12} - b_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus by definition of matrix equality, we have

$$\left.\begin{aligned} b_{11} + b_{21} &= 0 \\ b_{11} - b_{21} &= 0 \end{aligned}\right\} \qquad (1)$$

and

$$\left.\begin{aligned} b_{12} + b_{22} &= 0 \\ b_{12} - b_{22} &= 0 \end{aligned}\right\} \qquad (2)$$

The system (1) [as may be verified by first adding and then substracting the two equations] is solvable if and only if b_{11} and $b_{21} = 0$. Similarly the system (2) is solvable if and only if b_{12} and $b_{22} = 0$.

Thus, in this example B must itself be the zero matrix

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

A comparison of this exercise and the previous one reveals the following interesting (and complicating) fact, Given a square matrix A which is not the zero matrix, there are situations in which it is possible to find infinitely many matrices B such that AB = 0. On the other hand, there are other situations in which for the given matrix A there is only one matrix B (namely the zero matrix itself) such that AB = 0.

4.2.3 (continued)

This result is different from most arithmetic we have studied. For example, in numerical arithmetic if $a \neq 0$ then $ab = 0$ if and only if $b = 0$, while in dot-product-arithmetic if $\underline{a}$ is a non-zero vector $\underline{a} \cdot \underline{b} = 0$ always has infinitely many solutions for $\underline{b}$.

In the next unit we shall discuss this property of matrix multiplication in more detail.

Notice also that in the case in which $B = 0$ is the only matrix for which $AB = 0$, then $CA = 0$ yields only the solution $C = 0$. For example, using A as in this exercise, we see that if

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

then

$$\begin{pmatrix} c_{11} + c_{12} & c_{11} - c_{12} \\ c_{21} + c_{22} & c_{21} - c_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

whereupon we conclude as before that $c_{11} = c_{12} = 0$ and $c_{21} = c_{22} = 0$.

The results of these last two exercises should not be confused with the converse problem. That is, while $AB = 0$ does not tell us whether either A or B must be the zero matrix, it should be clear that if 0 denotes the zero matrix then $A0 = 0$. Namely, in the 2 x 2 case we have

$$A0 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= 0$$

That is, $A0 = 0$, but $AB = 0$ does not always mean that $B = 0$.

4.2.4

This exercise is actually a companion, or a corollary, of the previous two exercises. Recall that the fact that ab = 0 implied that either a = 0 or b = 0 hinged on the fact that a^{-1} existed. Our claim is that in part (a) the matrix A has an inverse (that is, we intend to show that there is a matrix X such that AX = I, and X is then denoted by A^{-1}), while the matrix B in part (b) does not have an inverse.

a. Letting

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$$

we have

$$AX = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$$

$$= \begin{pmatrix} x_{11} + x_{21} & x_{12} + x_{22} \\ x_{11} - x_{21} & x_{12} - x_{22} \end{pmatrix}.$$

Hence AX = I implies that

$$\begin{pmatrix} x_{11} + x_{21} & x_{12} + x_{22} \\ x_{11} - x_{21} & x_{12} - x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore,

$$\left.\begin{aligned} x_{11} + x_{21} &= 1 \\ x_{11} - x_{21} &= 0 \end{aligned}\right\} \qquad (1)$$

and

$$\left.\begin{aligned} x_{12} + x_{22} &= 0 \\ x_{12} - x_{22} &= 1 \end{aligned}\right\} \qquad (2)$$

Solving the pair of equations (1) for x_{11} and x_{21} yields

4.2.4 (continued)

$$x_{11} = x_{21} = \frac{1}{2} \tag{3}$$

Similarly solving (2) yields

$$x_{12} = -x_{22} = \frac{1}{2} \tag{4}$$

Putting (3) and (4) into our definition of X yields

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

As a check

$$AX = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} + \frac{1}{2} & \frac{1}{2} - \frac{1}{2} \\ \frac{1}{2} - \frac{1}{2} & \frac{1}{2} + \frac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore,

$$X = A^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

b. Letting

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$$

we have

$$BX = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{11} + x_{21} & x_{12} + x_{22} \\ x_{11} + x_{21} & x_{12} + x_{22} \end{pmatrix}. \tag{5}$$

Then since BX = I, equation (5) implies that

4.2.4 (continued)

$$x_{11} + x_{21} = 1 \tag{6}$$

and also that

$$x_{11} + x_{21} = 0 . \tag{7}$$

Equations (6) and (7) together imply that 1 = 0 (since both equal $x_{11} + x_{21}$).

This contradiction establishes the fact that there is no matrix X such that BX = I.

Thus B^{-1} does not exist (and this is closely related to why there are infinitely many matrices Y such that B = 0 while only the zero matrix satisfies AX = 0). In other words, the matrix equation BX = I has no solutions.

4.2.5

a. $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

Therefore,

$$A^2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1+1 & 1+2 \\ 1+2 & 1+4 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \tag{1}$$

$$A^3 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

$$= \left[\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \right] \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \tag{2}$$

[since matrix multiplication is associative (see Exercise 4.2.1)]

Substituting the result of (1) into (2) we obtain

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

4.2.5 (continued)

$$= \begin{pmatrix} 2+3 & 2+6 \\ 3+5 & 3+10 \end{pmatrix}$$

$$= \begin{pmatrix} 5 & 8 \\ 8 & 13 \end{pmatrix}$$

Aside: The interested reader may observe that multiplying a matrix X on the right by

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

yields

$$XA = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} x_{11} + x_{12} & x_{11} + 2x_{12} \\ x_{21} + x_{22} & x_{21} + 2x_{22} \end{pmatrix}.$$

Thus, in any event the first column of XA is obtained by adding the entries in each row of X. In particular in computing A^2 we pick X = A so that

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1+1 & \_\_ \\ 1+2 & \_\_ \end{pmatrix} = \begin{pmatrix} 2 & \_\_ \\ 3 & \_\_ \end{pmatrix}$$

The specific choice of X = A also tells us that the entry in the first row, second column is the same as the entry in the second row, first column, while the entry in the second row, second column is the sum of the entries in the first column. Thus, the various powers of A will be given by

$$\underset{A}{\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}} \to \underset{A^2}{\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}} \to \underset{A^3}{\begin{pmatrix} 5 & 8 \\ 8 & 13 \end{pmatrix}} \to \underset{A^4}{\begin{pmatrix} 13 & 21 \\ 21 & 34 \end{pmatrix}} \to \underset{A^5}{\begin{pmatrix} 34 & 55 \\ 55 & 89 \end{pmatrix}}$$

4.2.5 (continued)

$$\to \underset{A^6}{\begin{pmatrix} 89 & 144 \\ 144 & 233 \end{pmatrix}}$$

b. $$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad (1)$$

$$A^2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1+1 & 1+1 \\ 1+1 & 1+1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \qquad (2)$$

$$A^3 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$= \left[\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right] \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2+2 \\ 2+2 \end{pmatrix} \begin{pmatrix} 2+2 \\ 2+2 \end{pmatrix}$$

$$= \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} \qquad (3)$$

Comparing (1), (2), and (3) we may conjecture that

$$A^n = \begin{pmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{pmatrix}.$$

Since

$$A^1 = A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2^\circ & 2^\circ \\ 2^\circ & 2^\circ \end{pmatrix}$$

$$A^2 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 2^1 & 2^1 \\ 2^1 & 2^1 \end{pmatrix}$$

4.2.5 (continued)

and

$$A^3 = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} = \begin{pmatrix} 2^2 & 2^2 \\ 2^2 & 2^2 \end{pmatrix}.$$

One way of confirming our conjecture is by use of mathematical induction. Namely we already know that

$$A^k = \begin{pmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{pmatrix}$$

is true when $k = 1,2,3$.

We must then show that the assumption that

$$A^k = \begin{pmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{pmatrix}$$

implies that $A^{k+1} = \begin{pmatrix} 2^k & 2^k \\ 2^k & 2^k \end{pmatrix}$.

To this end,

$$A^{k+1} = A^kA = \begin{pmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2^{k-1} + 2^{k-1} & 2^{k-1} + 2^{k-1} \\ 2^{k-1} + 2^{k-1} & 2^{k-1} + 2^{k-1} \end{pmatrix}$$

$$= \begin{bmatrix} 2(2^{k-1}) & 2(2^{k-1}) \\ 2(2^{k-1}) & 2(2^{k-1}) \end{bmatrix}.$$

Hence

$$A^{k+1} = \begin{pmatrix} 2^k & 2^k \\ 2^k & 2^k \end{pmatrix}$$

and our proof by mathematical induction is complete.

4.2.5 (continued)

c.

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

Therefore,

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

$$= \begin{bmatrix} 1 + 5 - 6 & 1 + 2 - 3 & 3 + 6 - 9 \\ 5 + 10 - 12 & 5 + 4 - 6 & 15 + 12 - 18 \\ -2 - 5 + 6 & -2 - 2 + 3 & -6 - 6 + 9 \end{bmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 3 & 3 & 9 \\ -1 & -1 & -3 \end{pmatrix}.$$

Therefore,

$$\begin{aligned} A^3 &= AAA \\ &= (AA)A \\ &= A^2A \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 3 & 3 & 9 \\ -1 & -1 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{pmatrix}. \end{aligned}$$

Therefore,

$$A^3 = \begin{pmatrix} 0 + 0 + 0 & 0 + 0 + 0 & 0 + 0 + 0 \\ 3 + 15 - 18 & 3 + 6 - 9 & 9 + 18 - 27 \\ -1 - 5 + 6 & -1 - 2 + 3 & -3 - 6 + 9 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This example shows that it is possible for a power of a non-zero matrix to be the zero matrix. Such a matrix is given a special name. Namely if the non-zero matrix A has the

4.2.5 (continued)

property that there is a positive whole number n such that $A^n = 0$ (where 0 denotes the zero matrix), then A is called a <u>nilpotent matrix</u>.

4.2.6

$$A = \begin{pmatrix} 5 & 3 \\ 3 & -3 \end{pmatrix}$$

a. Therefore,

$$A^2 = \begin{pmatrix} 5 & 3 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} 5 & 3 \\ 3 & -3 \end{pmatrix}$$

$$= \begin{pmatrix} 25 + 9 & 15 - 9 \\ 15 - 9 & 9 + 9 \end{pmatrix}$$

$$= \begin{pmatrix} 34 & 6 \\ 6 & 18 \end{pmatrix}. \qquad (1)$$

Now recalling that cA means the matrix obtained by multiplying each entry of A by c, we also have

$$-2A = \begin{pmatrix} -10 & -6 \\ -6 & +6 \end{pmatrix}. \qquad (2)$$

Then $-24I =$

$$-24 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence

$$-24I = \begin{pmatrix} -24 & 0 \\ 0 & -24 \end{pmatrix} \qquad (3)$$

Combining equations (1), (2), and (3) we have

4.2.6 (continued)

$$A^2 - 2A - 24\,I = \begin{pmatrix} 34 & 6 \\ 6 & 18 \end{pmatrix} + \begin{pmatrix} -10 & -6 \\ -6 & 6 \end{pmatrix} + \begin{pmatrix} -24 & 0 \\ 0 & -24 \end{pmatrix}$$

or,

$$A^2 - 2A - 24\,I = \begin{pmatrix} 34 - 10 - 24 & 6 - 6 + 0 \\ 6 - 6 + 0 & 18 + 6 - 24 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

That is,

$A^2 - 2A - 24\,I = 0$ (where 0 is the zero matrix). (4)

Note:

Although the proof is beyond the scope of our course, there is a very remarkable theorem that applies to square matrices. We shall discuss the meaning of the theorem in more detail in Part 3 when we talk about eigenvalues but the "mechanics" are the following. Suppose we have a 2 x 2 matrix (and corresponding results hold for matrices of higher dimensions, but the computations are messier)

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

We then form the matrix $A - xI$, which is given by

$$A - xI = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$$

4.2.6 (continued)

$$= \begin{bmatrix} a_{11} - x & a_{12} \\ a_{21} & a_{22} - x \end{bmatrix}.$$

We then look at the equation obtained by setting the determinant of the matrix A - xI equal to 0. This yields

$$(a_{11} - x)(a_{22} - x) - a_{12}a_{22} = 0$$

or

$$x^2 - (a_{11} + a_{22})x + (a_{11}a_{22} - a_{12}a_{22}) = 0 . \qquad \text{(i)}$$

Without worrying here about why one would want to use equation (i) the amazing fact is that if the matrix A is used in place of x in equation (i) and matrix arithmetic replaces numerical arithmetic, equation (i) is still obeyed.

For example, in this particular exercise

$$A = \begin{pmatrix} 5 & 3 \\ 3 & -3 \end{pmatrix}$$

Hence,

$$A-xI = \begin{pmatrix} 5 & 3 \\ 3 & -3 \end{pmatrix} + \begin{pmatrix} -x & 0 \\ 0 & -x \end{pmatrix}$$

$$= \begin{pmatrix} 5 - x & 3 \\ 3 & -3 - x \end{pmatrix}$$

Therefore,

det(A - xI) = 0 implies

$$(5 - x)(-3-x) - (3)(3) = 0,$$

or

$$x^2 - 2x - 24 = 0 \qquad \text{(ii)}$$

4.2.6 (continued)

Letting A replace x in (ii) we obtain

$$A^2 - 2A - 24\ I^* = 0$$

which agrees with the result of this exercise.

b. We observe, of course, that just as in the previous exercise we could compute A^3 directly. Namely, from (1) in part (a),

$$A^2 = \begin{pmatrix} 5 & 3 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} 5 & 3 \\ 3 & -3 \end{pmatrix}$$

$$= \begin{pmatrix} 34 & 6 \\ 6 & 18 \end{pmatrix}.$$

So that

$$A^3 = A^2 A$$

$$= \begin{pmatrix} 34 & 6 \\ 6 & 18 \end{pmatrix} \begin{pmatrix} 5 & 3 \\ 3 & -3 \end{pmatrix}$$

$$= \begin{pmatrix} 170 + 18 & 102 - 18 \\ 30 + 54 & 18 - 54 \end{pmatrix}$$

$$= \begin{pmatrix} 188 & 84 \\ 84 & -36 \end{pmatrix}. \qquad (5)$$

We wish, however, to emphasize the result that for our choice of A, $A^2 - 2A - 24\ I = 0$. To this end we write A^3 in a form which emphasizes $A^2 - 2A - 24\ I$.

Namely, if we multiply $A^2 - 2A - 24\ I$ by A we obtain an A^3 term but also the term $-2A^2 - 24\ AI$, or since I is the

*We must be careful to observe that while 24 is a number it is not a matrix. The interpretation used in the theorem is that 24 is replaced by 24 I.

4.2.6 (continued)

identity matrix, $-2A^2-24A$. Hence if we add $2A^2 + 24A$ to this expression we have the identity

$$A^3 = A(A^2 - 2A - 24\ I) + 2A^2 + 24A\ . \tag{6}$$

Since $A^2 - 2A - 24\ I = 0$, it follows from (6) that

$$A^3 = 2A^2 + 24A^*. \tag{7}$$

Notice that (7) shows us that we can compute A^3 in terms of the lower powers A^2 and A.

In fact we may "reduce" A^3 even further by observing that

$$2A^2 + 24A = 2(A^2 - 2A - 24\ I) + 28A + 48\ I. \tag{8}$$

and, again, since $A^2 - 2A - 24\ I = 0$, (8) implies that $2A^2 + 24A = 28A + 48\ I$.

Putting this result into (7) yields

$$A^3 = 28A + 48\ I. \tag{9}$$

Equation (9) is particularly convenient for computing A^3 since

$$\begin{aligned} 28A + 48\ I &= 28\begin{pmatrix} 5 & 3 \\ 3 & -3 \end{pmatrix} + 48\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 140 & 84 \\ 84 & -84 \end{pmatrix} + \begin{pmatrix} 48 & 0 \\ 0 & 48 \end{pmatrix} \\ &= \begin{pmatrix} 140 + 48 & 84 + 0 \\ 84 + 0 & -84 + 48 \end{pmatrix} \end{aligned}$$

*This follows from the fact that if 0 is the zero matrix then $A0 = 0$. Therefore $A(A^2 - 2A - 24\ I) = A0 = 0$ and since 0 is the identity matrix with respect to addition $A^3 = 0 + 2A^2 + 24A = 2A^2 + 24A$.

4.2.6 (continued)

$$= \begin{pmatrix} 188 & 84 \\ 84 & -36 \end{pmatrix}. \qquad (10)$$

Equation (10) checks with the result in (5).

c. The main aim of part (b) was to emphasize that we may "reduce" powers of matrices by a technique which is essentially the same as the recipe for performing long division of polynomials. For example, when we want to divide x^3 by $x^2 - 2x - 24$, the recipe is nothing more than a convenient form for factoring $x^2 - 2x - 24$ from x^3. Without going through the specific details, the point is that

$$\begin{aligned} x^3 &= x(x^2 - 2x - 24) + 2x^2 + 24x \\ &= x(x^2 - 2x - 24) + 2(x^2 - 2x - 24) + 28x + 48 \\ &= (x + 2)(x^2 - 2x - 24) + 28x + 48. \end{aligned}$$

Consequently, it is easy to see that when x^3 is divided by $x^2 - 2x - 24$, the quotient is $x + 2$ and the remainder is $28x + 48$. In particular, if $x^2 - 2x - 24$ happens to be zero then the remainder is the quotient.

At any rate, if we now go through the procedure in (b) but use the convenience of the long division notation, we have:

$$\begin{array}{l|l} A^7 & A^2 - 2A - 24I \\ \underline{A^7 - 2A^6 - 24A^5} & \overline{A^5 + 2A^4 + 28A^3 + 104A^2 + 880A + 4256} \\ \quad 2A^6 + 24A^5 & \\ \quad \underline{2A^6 - 4A^5 - 48A^4} & \\ \qquad 28A^5 + 48A^4 & \\ \qquad \underline{28A^5 - 56A^4 - 672A^3} & \\ \qquad\quad 104A^4 + 672A^3 & \\ \qquad\quad \underline{104A^4 - 208A^3 - 2496A^2} & \\ \qquad\qquad 880A^3 + 2496A^2 & \\ \qquad\qquad \underline{880A^3 - 1760A^2 - 21120A} & \\ \qquad\qquad\quad 4256A^2 + 21120A & \\ \qquad\qquad\quad \underline{4256A^2 - 8512A - 102144} & \\ \qquad\qquad\qquad 29632A + 102144 & \end{array}$$

4.2.6 (continued)

Hence,

$$A^7 = (A^5 + 2A^4 + 28A^3 + 104A^2 + 880A + 4256\ I)(A^2 - 2A - 24\ I)$$
$$+ 29{,}632A + 102{,}144\ I.$$

Then, since $A^2 - 2A - 24\ I = 0$.

$$A^7 = 29{,}632A + 102{,}144\ I \qquad (11)$$

While the division in arriving at (11) was a bit cumbersome and while the right side of (11) may look a bit cumbersome as well, the point is that it would have been even more cumbersome to compute A^7 directly in the manner that we computed A^3 in part (b).

The important point is that since $A^2 - 2A - 24\ I = 0$, any polynomial in A can be reduced to the form $cA + kI$, where c and k are real numbers. That is, we can factor out $A^2 - 2A - 24\ I$ from the polynomial which means that the remainder can be no more than a first degree polynomial in A.

At any rate, with respect to our specific problem, equation (11) shows that

$$A^7 = 29{,}632A + 102{,}144\ I$$
$$= 29{,}632\begin{pmatrix} 5 & 3 \\ 3 & -3 \end{pmatrix} + 102{,}144\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 148{,}160 & 88{,}896 \\ 88{,}896 & -88{,}896 \end{pmatrix} + \begin{pmatrix} 102{,}144 & 0 \\ 0 & 102{,}144 \end{pmatrix}$$
$$= \begin{pmatrix} 250{,}304 & 88{,}896 \\ 88{,}896 & 13{,}248 \end{pmatrix}$$

4.2.7

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

4.2.7 (continued)

$$AB = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$= \begin{pmatrix} b_{11} - b_{21} & b_{12} - b_{22} \\ 2b_{21} & 2b_{22} \end{pmatrix} , \qquad (1)$$

while

$$BA = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} b_{11} & -b_{11} + 2b_{12} \\ b_{21} & -b_{21} + 2b_{22} \end{pmatrix} . \qquad (2)$$

Comparing (1) and (2) we see that AB = BA if and only if

$$\begin{cases} b_{11} = b_{11} - b_{21} & (3) \\ b_{21} = 2b_{21} & (4) \\ -b_{11} + 2b_{12} = b_{12} - b_{22} & (5) \\ -b_{21} + 2b_{22} = 2b_{22} . & (6) \end{cases}$$

Both (3) and (4) imply that

$$b_{21} = 0. \qquad (7)$$

With $b_{21} = 0$, equation (3) says that $b_{11} = b_{11}$, and since this is an identity, b_{11} may be arbitrarily chosen. Simiarly with $b_{21} = 0$, equation (6) becomes $2b_{22} = 2b_{22}$ which is satisfied by any value of b_{22}. So b_{22} may also be arbitrarily chosen.

Finally, equation (5) implies that

$$b_{12} = b_{11} - b_{22} . \qquad (8)$$

4.2.7 (continued)

Combining (7) and (8) with the fact that b_{11} and b_{22} may be chosen at random, we have

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

$$= \begin{pmatrix} b_{11} & b_{11} - b_{22} \\ 0 & b_{22} \end{pmatrix} \qquad (9)$$

As a check, (9) yields

$$AB = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} b_{11} & b_{11} - b_{22} \\ 0 & b_{22} \end{pmatrix}$$

$$= \begin{pmatrix} b_{11} & b_{11} - b_{22} - b_{22} \\ 0 & 2b_{22} \end{pmatrix} .$$

Therefore,

$$AB = \begin{pmatrix} b_{11} & b_{11} - 2b_{22} \\ 0 & 2b_{22} \end{pmatrix} \qquad (10)$$

while

$$BA = \begin{pmatrix} b_{11} & b_{11} - b_{22} \\ 0 & b_{22} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$$

$$= \begin{bmatrix} b_{11} & -b_{11} + 2(b_{11} - b_{22}) \\ 0 & 2b_{22} \end{bmatrix}$$

$$\begin{pmatrix} b_{11} & b_{11} - 2b_{22} \\ 0 & 2b_{22} \end{pmatrix}. \qquad (11)$$

Comparing (10) and (11) shows that with B as in (9) AB = BA.

As specific examples of matrices which satisfy (9) we have

$$\begin{pmatrix} 5 & 2 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 8 & 1 \\ 0 & 7 \end{pmatrix} \quad \text{etc.}$$

4.2.7 (continued

Since 5 - 3 = 2 and 8 - 7 = 1.

4.2.8

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

then

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

b. The matrix $(A^T)^T$ is precisely A. Namely we form A^T by interchanging the rows and columns of A. We form $(A^T)^T$ by interchanging the rows and columns of A^T. In other words we obtain $(A^T)^T$ by interchanging the rows and columns of A <u>twice</u>, but this is the same as leaving the rows and columns of A in place.

For example, with $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ we saw that

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

Hence,

$(A^T)^T$ is $\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}^T$ or $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, which is A.

c. Letting, as usual,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

we have

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

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

4.2.8 (continued)

Therefore,

$$(AB)^T = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{21}b_{11} + a_{22}b_{21} \\ a_{11}b_{12} + a_{12}b_{22} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \qquad (1)$$

On the other hand,

$$B^TA^T = \begin{pmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$$

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

$$= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{21}b_{11} + a_{22}b_{21} \\ a_{11}b_{12} + a_{12}b_{22} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \qquad (2)$$

Comparing (1) and (2) shows that $(AB)^T = B^TA^T$.

From a learning point of view, the crucial part of this exercise is the observation that since matrix multiplication is not commutative, we must keep the order of the factors intact.

Thus, while it might seem "natural" to say that $(AB)^T = A^TB^T$, the fact is that in general this result is not true. What is true is that $(AB)^T = B^TA^T$. In other words, the transpose of a product is the product of the transposes, with the order of multiplication reversed.

There are many situations in which one is interested in the transpose of a matrix rather than in the given matrix itself. We shall not pursue this idea further in this exercise but we shall in our later work.