Unit 3: Inverse Matrices

4.3.1

The main aim of this exercise is to stress the structure of matrix arithmetic.

a. If we had been given the equation

$$\frac{1}{3}X - AB = C$$

where X, A, B, and C were numbers, we would have at once concluded that

$$\frac{1}{3}X = AB + C$$

or

$$X = 3(AB + C). \qquad (1)$$

The important point is that the rules which were used to arrive at equation (1) are true in matrix arithmetic as well as in numerical arithmetic. In other words, equation (1) is still valid when X, A, B, and C are matrices.

In our particular exercise, the specific choices of A, B, and C yield

$$AB = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 3 + 2 & 4 + 3 \\ 6 + 6 & 8 + 9 \end{pmatrix}$$

$$= \begin{pmatrix} 5 & 7 \\ 12 & 17 \end{pmatrix}.$$

Hence,

4.3.1 continued

$$AB + C = \begin{pmatrix} 5 & 7 \\ 12 & 17 \end{pmatrix} + \begin{pmatrix} 5 & 4 \\ 6 & 5 \end{pmatrix}.$$

Therefore,

$$AB + C = \begin{pmatrix} 10 & 11 \\ 18 & 22 \end{pmatrix}.$$

If we then recall that 3(AB + C) means to multiply each entry of AB + C by 3, equation (1) becomes

$$X = 3(AB + C) = \begin{pmatrix} 30 & 33 \\ 54 & 66 \end{pmatrix}. \qquad (2)$$

As a check, we have that with X as defined by equation (2),

$$\frac{1}{3} X - AB = \frac{1}{3}\begin{pmatrix} 30 & 33 \\ 54 & 66 \end{pmatrix} - \begin{pmatrix} 5 & 7 \\ 12 & 17 \end{pmatrix}$$

$$= \begin{pmatrix} 10 & 11 \\ 18 & 22 \end{pmatrix} - \begin{pmatrix} 5 & 7 \\ 12 & 17 \end{pmatrix}$$

$$= \begin{pmatrix} 5 & 4 \\ 6 & 5 \end{pmatrix}$$

$$= C.$$

In summary, when it comes to addition and scalar multiplication of matrices, the arithmetic has the same structure as that of ordinary numerical arithmetic.

b. If X, A, B, C, and 0 had been numbers, then the equation

$$AX - BC = 0$$

would have possessed the solution

4.3.1 continued

$$X = A^{-1}(BC) \tag{3}$$

provided only that $A \neq 0$.

Now in arriving at equation (3), the only step that was not covered by either addition or scalar multiplication was in assuming the existence of A^{-1}. That is, we know that depending on the specific choice of A, A^{-1} may not exist even though A is not the zero matrix. However, assuming that A^{-1} does exist, the equation $AX - BC = 0$ is solved precisely by the same structure as in the numerical case. More specifically, we have

$$AX - BC = 0$$

implies that

$$AX = BC$$

and if A^{-1} exists, this, in turn, implies that

$$A^{-1}(AX) = A^{-1}(BC)$$

or, since matrix multiplication is commutative,

$$(A^{-1}A)X = A^{-1}(BC).$$

By definition of A^{-1}, $A^{-1}A = I$ and $IX = X$, hence, our last equation implies that

$$X = A^{-1}(BC).$$

Thus, all that remains to be done in this part of the exercise is to see whether A^{-1} exists. If it does, we can use A^{-1} to solve equation (3) by using the ordinary matrix operations. If A^{-1} does not exist, then we cannot find the required matrix X. In other words, if A^{-1} doesn't exist, then the equation $AX - BC = 0$ cannot be solved for X.

4.3.1 continued

Now, one way of trying to compute A^{-1} is to use the augmented matrix technique discussed in our supplementary notes. That is,

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 1 \end{bmatrix}$$

so that A^{-1} does exist and is in fact given by

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

Using equation (4) in (3) with B and C as given in this exercise, we see from (3) that

$$X = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \left[\begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 5 & 4 \\ 6 & 5 \end{pmatrix} \right]$$

$$= \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 15 + 24 & 12 + 20 \\ 10 + 18 & 8 + 15 \end{pmatrix}$$

$$= \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 39 & 32 \\ 28 & 23 \end{pmatrix}$$

$$= \begin{pmatrix} 117 - 28 & 96 - 23 \\ -78 + 28 & -64 + 23 \end{pmatrix}$$

$$= \begin{pmatrix} 89 & 73 \\ -50 & -41 \end{pmatrix}. \tag{5}$$

As a check, the value of X in equation (5) implies that

4.3.1 continued

$$AX - BC = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 89 & 73 \\ -50 & -41 \end{pmatrix} - \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 5 & 4 \\ 6 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} 89 - 50 & 73 - 41 \\ 178 - 150 & 146 - 123 \end{pmatrix} - \begin{pmatrix} 15 + 24 & 12 + 20 \\ 10 + 18 & 8 + 15 \end{pmatrix}$$

$$= \begin{pmatrix} 39 & 32 \\ 28 & 23 \end{pmatrix} - \begin{pmatrix} 39 & 32 \\ 28 & 23 \end{pmatrix}$$

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

$$= 0$$

which checks with equation (3).

In passing, it is worth making the aside that in the case of 2 x 2 matrices, it is not too difficult (in fact, it might even be easier than our augmented matrix technique) to find A^{-1} directly. Namely, letting

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

we have

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

implies that

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This tells us that

4.3.1 continued

$$\begin{pmatrix} x_{11} + x_{21} & x_{12} + x_{22} \\ 2x_{11} + 3x_{21} & 2x_{12} + 3x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and, therefore that

$$\left.\begin{aligned} x_{11} + x_{21} &= 1 \\ 2x_{11} + 3x_{21} &= 0 \end{aligned}\right\} \qquad (6)$$

while

$$\left.\begin{aligned} x_{12} + x_{22} &= 0 \\ 2x_{12} + 3x_{22} &= 1 \end{aligned}\right\} . \qquad (7)$$

Equations (6) may be solved to yield

$x_{11} = 3,\ x_{21} = -2$

while equations (7) yield

$x_{12} = -1$ and $x_{22} = 1$

so that

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

as obtained in equation (4).

The drawback to this latter technique lies in the fact that when n is large (even when n is 3, 4, or 5) finding A^{-1} for the n x n matrix A becomes extremely cumbersome since we obtain several systems of equations in several unknowns.

4.3.2

Let

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

We shall see what conditions are imposed on x_{11}, x_{12}, x_{21}, and x_{22} by the equation

$$AX = I. \tag{1}$$

Since

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

equation (1) becomes

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} ax_{11} + bx_{21} & ax_{12} + bx_{22} \\ cx_{11} + dx_{21} & cx_{12} + dx_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore,

$$\left.\begin{aligned} ax_{11} + bx_{21} &= 1 \\ cx_{11} + dx_{21} &= 0 \end{aligned}\right\} \tag{1'}$$

and

$$\left.\begin{aligned} ax_{12} + bx_{22} &= 0 \\ cx_{12} + dx_{22} &= 1 \end{aligned}\right\}. \tag{2}$$

4.3.2 continued

In the "usual way" we solve (1') by multiplying the top equation by d and the bottom equation by -b to yield

$$\left.\begin{aligned} adx_{11} + bdx_{21} &= d \\ -bcx_{11} - bdx_{21} &= 0 \end{aligned}\right\}. \qquad (3)$$

If we add the two equations in (3), we obtain

$$(ad - bc)x_{11} = d. \qquad (4)$$

From (4), we see that x_{11} is uniquely determined <u>provided ad - bc ≠ 0</u>. For in this case,

$$x_{11} = \frac{d}{ad - bc}. \qquad (5)$$

[If ad - bc = 0, then if d ≠ 0, there is no value of x_{11} which can satisfy equation (4) since in this case, equation (4) says

$$0x_{11} = d \neq 0$$

which is impossible since $0x_{11} = 0$ for all x_{11}.
On the other hand, if both ad - bc = 0 and d = 0, then any value of x_{11} satisfies equation (4).]
Thus, for the number x_{11} to be uniquely determined, it is necessary that

$$ad - bc \neq 0. \qquad (6)$$

The question now is whether this condition is also sufficient. The best way to find out is to assume that condition (6) holds and then see whether x_{12}, x_{21}, and x_{22} are uniquely determined.
To this end, we return to equations (1') and now multiply the top equation by c and the bottom equation by -a to obtain

$$\begin{aligned} acx_{11} + bcx_{21} &= c \\ -acx_{11} - adx_{21} &= 0 \end{aligned}$$

4.3.2 continued

whereupon adding these equations, we obtain

$$(bc - ad)x_{21} = c$$

or to emphasize condition (6),

$$(ad - bc)x_{21} = -c \tag{7}$$

then since $ad - bc \neq 0$, equation (7) yields that x_{21} is uniquely given by

$$x_{21} = \frac{-c}{ad - bc}. \tag{8}$$

In a similar way, we try to use equations (2) to see if both x_{12} and x_{22} are uniquely determined. To solve for x_{12}, we multiply the top equation by d and the bottom equation by -b to obtain

$$adx_{12} + bdx_{22} = 0$$
$$-bcx_{12} - bdx_{22} = -b$$

whereupon adding these two equations yields

$$(ad - bc)x_{12} = -b$$

or since we are assuming that $ad - bc \neq 0$, this means that

$$x_{12} = \frac{-b}{ad - bc}. \tag{9}$$

Finally, to solve for x_{22}, we multiply the top equation in (2) by -c and the bottom equation by a to obtain

$$-acx_{12} - bcx_{22} = 0$$
$$acx_{12} + adx_{22} = a$$

and adding these two equations yields

$$(ad - bc)x_{22} = a,$$

4.3.2 continued

and, again, since $ad - bc \neq 0$, this implies that

$$x_{22} = \frac{a}{ad - bc}. \tag{10}$$

Equations (5), (8), (9), and (10) show us that as long as $ad - bc \neq 0$, A^{-1} exists and is in fact given by

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

$$= \begin{bmatrix} \frac{d}{ad - bc} & \frac{-b}{ad - bc} \\ \frac{-c}{ad - bc} & \frac{a}{ad - bc} \end{bmatrix}. \tag{11}$$

Recalling our rule for scalar multiplication, (11) may be rewritten in the form

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{12}$$

Recalling that

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{13}$$

a comparison of (12) and (13) reveals that to obtain A^{-1} from A, where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

we interchange the two entries on the main diagonal, change the sign of the other two entries, and multiply the resulting matrix

4.3.2 continued

by $\frac{1}{ad - bc}$* provided $ad - bc \neq 0$. If $ad - bc = 0$, A^{-1} does not exist.

While, in the interest of computational simplicity, we have concentrated solely on 2 x 2 matrices in this exercise, the fact is that these results generalize to n x n matrices, in the following form.

Notice that if

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then $ad - bc$ is the determinant of A. In other words, what we showed in this problem was that if $\det(A) \neq 0$, then A^{-1} existed. We further went on to compute A^{-1} in this case. In general, although determinants are "nastier" for square matrices of dimension greater than 2 x 2, the same result holds for all n x n matrices. Namely, if A is any n x n matrix, then A^{-1} exists if and only if $\det(A) \neq 0$. We shall wait until Block 8 before we probe the question of what a determinant looks like in the case of an n x n matrix with n greater than 3, but for now we thought it an interesting observation to show how the existence of A^{-1} was related to the determinant of A.

4.3.3

Given that $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a diagonal matrix, we have that $b = c = 0$. Consequently, the result of the previous exercise that

*Arithmetically, a particularly simple case arises if $ad - bc = 1$ (or even -1) since then the factor $\frac{1}{ad - bc} = 1$ (or -1) so that in this case

$$A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad \left[\text{or } -\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} -d & b \\ c & -a \end{pmatrix}\right].$$

For this reason (and others as well), special study is made of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for which $ad - bc = \pm 1$.

4.3.3 continued

$$A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

yields

$$A^{-1} = \frac{1}{ad}\begin{pmatrix} d & 0 \\ 0 & a \end{pmatrix}. \qquad (1)$$

Obviously, for (1) to be mathematically meaningful we must have that $ad \neq 0$, which in turn means that

$a \neq 0$ and $d \neq 0$.

(If either $a = 0$ or $d = 0$, the fact that $b = c = 0$ implies that $ad - bc = 0$, so that A^{-1} does not exist. In other words, a diagonal 2 x 2 matrix has an inverse provided that none of its entries on the main diagonal is 0.)

At any rate, if $ad \neq 0$, equation (1) becomes

$$A^{-1} = \begin{pmatrix} \frac{d}{ad} & 0 \\ 0 & \frac{a}{ad} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{d} \end{pmatrix}. \qquad (2)$$

As a quick check that (2) is correct we observe that

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

Notice that (2) generalizes to any diagonal n x n matrix. Namely, assume that $A = (a_{ij})$ is any n x n diagonal matrix with $a_{11}, a_{22}, \ldots,$ and a_{nn} all unequal to 0. Then $a_{11}^{-1}, a_{22}^{-1}, \ldots$ and a_{nn}^{-1} are real numbers with

$$a_{11}a_{11}^{-1} = \ldots = a_{nn}a_{nn}^{-1} = 1.$$

4.3.3 continued

Now we already know that to multiply two n x n diagonal matrices we simply multiply the diagonal entries term by term. Hence

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & 0_{nn} \end{pmatrix} \begin{pmatrix} a_{11}^{-1} & 0 & \dots & 0 \\ 0 & a_{22}^{-1} & \dots & 0 \\ 0 & 0 & \dots & a_{nn}^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}a_{11}^{-1} & 0 & \dots & 0 \\ 0 & a_{22}a_{22}^{-1} & \dots & 0 \\ 0 & 0 & \dots & a_{nn}a_{nn}^{-1} \end{pmatrix}$$

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

$$= I_n .$$

4.3.4(L)

(a) The main aim of this part of the exercise is to emphasize the arithmetic structure of matrices. Certainly, one could elect to solve this problem by brute force as follows. We let

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

and we then compute A^{-1} and B^{-1}, after which we compute $B^{-1}A^{-1}$. We then compute AB, after which we compute its inverse $(AB)^{-1}$. Finally we compare $B^{-1}A^{-1}$ with $(AB)^{-1}$, and if we haven't made any computational errors, we find that the two matrices are equal. The fact that we are dealing with 2 x 2 matrices makes the computation a bit easier since we may then use the "recipe"

4.3.4(L) continued

derived in Exercise 4.3.2 for finding inverse matrices.

The point is that the above procedure, even in the 2 x 2 case, is cumbersome, and in the general n x n case the computations quickly tend to get out of hand.

Yet from a structural point of view, we can use our "rules of the game" to obtain the result asked for in part (a) in an efficient manner which also happens to apply in the more general case of n x n matrices.

More specifically, we look at

$$(AB)(B^{-1}A^{-1}) \tag{1}$$

where A and B are both n x n matrices, and we are assuming that both A^{-1} and B^{-1} exist.

Since multiplication of matrices is associative, our expression (1) may be rewritten as

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} \ . \tag{2}$$

By definition, $BB^{-1} = I_n$ (the n x n identity matrix), so that (2) may be rewritten as

$$(AB)(B^{-1}A^{-1}) = A(I_n)A^{-1}. \tag{3}$$

Again, by associativity, $A(I_n)A^{-1}$ is equal to $(AI_n)A^{-1}$, and by the definition of I_n, $AI_n = A$, so that (3) may be rewritten in the form

$$(AB)(B^{-1}A^{-1}) = (AI_n)A^{-1} = AA^{-1}. \tag{4}$$

Finally, from the definition of A^{-1}, $AA^{-1} = I_n$, so that (4) becomes

$$(AB)(B^{-1}A^{-1}) = I_n \ . \tag{5}$$

4.3.4(L) continued

A similar sequence of steps shows that

$$(B^{-1}A^{-1})(AB) = I_n \tag{6}$$

and equations (5) and (6) show that by the definition of

$(AB)^{-1}$ [i.e., $(AB)(AB)^{-1} = (AB)^{-1}(AB) = I_n$], $(AB)^{-1} = B^{-1}A^{-1}$.

As a final note to our computation, notice that just as in ordinary algebra, we may abbreviate the above sequence of steps simply by writing

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = (AI_n)A^{-1} = AA^{-1} = I_n .$$

The beauty of the structural proof, aside from the fact that we don't get bogged down in computational details, is that we see how the result follows from the basic properties of matrices independently of the dimension of the matrix.

It should also be noted that it might seem more natural that $(AB)^{-1} = A^{-1}B^{-1}$. This result is not usually true.

(b) In this part of the exercise we simply want to check the result of part (a) in a particular example. We have

$A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$ while $B = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}$.

Consequently, as seen in Exercise 4.3.1,

$AB = \begin{pmatrix} 5 & 7 \\ 12 & 17 \end{pmatrix}$.

We may now compute A^{-1}, B^{-1}, and $(AB)^{-1}$ by the process derived in Exercise 4.3.2, noting that we have "cleverly" chosen A and B so that det A, det B, and det AB are all equal to 1*.

*While we do not need this result for our present work, it is a fact (theorem) that if A and B are n x n matrices then det AB = det A x det B. In particular, if det A = det B = 1, then det AB = 1. In a similar way, if either det A or det B is zero then det AB = 0, but we shall not explore such results in more detail here.

4.3.4(L) continued

Thus from the result of Exercise 4.3.2 we have:

$$A^{-1} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \tag{7}$$

$$B^{-1} = \begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix} \tag{8}$$

and

$$(AB)^{-1} = \begin{pmatrix} 17 & -7 \\ -12 & 5 \end{pmatrix}. \tag{9}$$

From (7) and (8) we see that

$$A^{-1}B^{-1} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 11 & -15 \\ -8 & 11 \end{pmatrix} \tag{10}$$

while

$$B^{-1}A^{-1} = \begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 17 & -7 \\ -12 & 5 \end{pmatrix} \tag{11}$$

Comparing expression (9) and (11) we see that $(AB)^{-1} = B^{-1}A^{-1}$, but if we look at (10) we also see that $(AB)^{-1}$ is not the same as $A^{-1}B^{-1}$.

In fact, if we look at A and B as denoting the first and second factors respectively rather than specific matrices, the result of part (a) indicates that $A^{-1}B^{-1}$ <u>should be the inverse of BA</u>. That is, part (a) says that to invert the product of two matrices we take the product of the inverses with the order commuted. As a check, we have,

$$BA = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 11 & 15 \\ 8 & 11 \end{pmatrix}.$$

Again, det BA = 1, so the recipe of Exercise 4.3.2 tells us that

4.3.4(L) continued

$$(BA)^{-1} = \begin{pmatrix} 11 & -15 \\ -8 & 11 \end{pmatrix}$$

and this result checks with the result of equation (10). That is

$(BA)^{-1} = A^{-1}B^{-1}.$

Thus, once again we see that it is important to keep matrix factors in their given order. In other words, since, in general, $AB \neq BA$, $(AB)^{-1}$ will not be the same as $(BA)^{-1}$.

4.3.5

a. We have

$$\begin{bmatrix} 1 & 3 & 5 & 1 & 0 & 0 \\ 2 & 7 & 9 & 0 & 1 & 0 \\ 3 & 9 & 7 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 5 & 1 & 0 & 0 \\ 0 & 1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -8 & -3 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 8 & 7 & -3 & 0 \\ 0 & 1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -8 & -3 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 8 & 7 & -3 & 0 \\ 0 & -8 & 8 & 16 & -8 & 0 \\ 0 & 0 & -8 & -3 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 4 & -3 & 1 \\ 0 & -8 & 0 & 13 & -8 & 1 \\ 0 & 0 & -8 & -3 & 0 & 1 \end{bmatrix}$$

4.3.5 continued

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & -3 & 1 \\ 0 & 1 & 0 & -\frac{13}{8} & 1 & -\frac{1}{8} \\ 0 & 0 & 1 & \frac{3}{8} & 0 & -\frac{1}{8} \end{array}\right].$$

Therefore,

$$A^{-1} = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 7 & 9 \\ 3 & 9 & 7 \end{pmatrix}^{-1} = \begin{pmatrix} 4 & -3 & 1 \\ -\frac{13}{8} & 1 & -\frac{1}{8} \\ \frac{3}{8} & 0 & -\frac{1}{8} \end{pmatrix}.$$

As a check we have

$$\begin{pmatrix} 1 & 3 & 5 \\ 2 & 7 & 9 \\ 3 & 9 & 7 \end{pmatrix} \begin{pmatrix} 4 & -3 & 1 \\ -\frac{13}{8} & 1 & -\frac{1}{8} \\ \frac{3}{8} & 0 & -\frac{1}{8} \end{pmatrix}$$

$$= \begin{pmatrix} 4 - \frac{39}{8} + \frac{15}{8} & -3 + 3 + 0 & 1 - \frac{3}{8} - \frac{5}{8} \\ 8 - \frac{91}{8} + \frac{27}{8} & -6 + 7 + 0 & 2 - \frac{7}{8} - \frac{9}{8} \\ 12 - \frac{117}{8} + \frac{21}{8} & -9 + 9 + 0 & 3 - \frac{9}{8} - \frac{7}{8} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3 .$$

Similarly,

$$\begin{pmatrix} 4 & -3 & 1 \\ -\frac{13}{8} & 1 & -\frac{1}{8} \\ \frac{3}{8} & 0 & -\frac{1}{8} \end{pmatrix} \begin{pmatrix} 1 & 3 & 5 \\ 2 & 7 & 9 \\ 3 & 9 & 7 \end{pmatrix} = \begin{pmatrix} 4 - 6 + 3 & 12 - 21 + 9 & 20 - 27 + 7 \\ -\frac{13}{8} + 2 - \frac{3}{8} & -\frac{39}{8} + 7 - \frac{9}{8} & -\frac{65}{8} + 9 - \frac{7}{8} \\ \frac{3}{8} + 0 - \frac{3}{8} & \frac{9}{8} + 0 - \frac{9}{8} & \frac{15}{8} + 0 - \frac{7}{8} \end{pmatrix}$$

4.3.5 continued

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3 .$$

[Note: The fact that AX = I is not enough to guarantee that XA = I (see for example Exercise 4.3.6 (d) and (e). This is why we must show that both AX and XA equal I before we conclude that $X = A^{-1}$.]

(b) Solving part (b) is a corollory of our solution to part (a).

Namely, if we are given the system

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ y_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{cases} \quad (1)$$

and if we let $A = (a_{ij})$ then if A^{-1} exists it follows that if $A^{-1} = (b_{ij})$, equations (1) may be solved for x_1, x_2, and x_3 in terms of y_1, y_2, and y_3 by

$$\begin{cases} x_1 = b_{11}y_1 + b_{12}y_2 + b_{13}y_3 \\ x_2 = b_{21}y_1 + b_{22}y_2 + b_{23}y_3 \\ x_3 = b_{31}y_1 + b_{31}y_2 + b_{33}y_3 \end{cases} . \quad (2)$$

This follows from coding equations (1) in the augmented matrix form

$$\begin{array}{ccc|ccc} x_1 & x_2 & x_3 & y_1 & y_2 & y_3 \\ \hline a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 1 \end{array}$$

and using the procedure of part (a).

Therefore

4.3.5 continued

Therefore

$$\left.\begin{aligned} y_1 &= x_1 + 3x_2 + 5x_3 \\ y_2 &= 2x_1 + 7x_2 + 9x_3 \\ y_3 &= 3x_1 + 9x_2 + 7x_3 \end{aligned}\right\} \qquad (3)$$

implies from our result of part (a) that

$$\left.\begin{aligned} x_1 &= 4y_1 - 3y_2 + y_3 \\ x_2 &= -\tfrac{13}{8} y_1 + y_2 - \tfrac{1}{8} y_3 \\ x_3 &= \tfrac{3}{8} y_1 - \tfrac{1}{8} y_3 \end{aligned}\right\} \qquad (4)$$

and one can now check directly by replacing x_1, x_2 and x_3 in (3) by their values in (4).

c. This is part (b) with $y_1 = 8$, $y_2 = 16$ and $y_3 = 32$. Putting these values into equations (4) we have

$$\begin{aligned} x_1 &= 4(8) - 3(16) + 32 = 32 - 48 + 32 = 16 \\ x_2 &= -\tfrac{13}{8}(8) + 16 - \tfrac{1}{8}(32) = -13 + 16 - 4 = -1 \\ x_3 &= \tfrac{3}{8}(8) - \tfrac{1}{8}(32) = 3 - 4 = -1 . \end{aligned}$$

As a check

$$\begin{aligned} 16 + 3(-1) + 5(-1) &= 8 \\ 2(16) + 7(-1) + 9(-1) &= 16 \\ 3(16) + 9(-1) + 7(-1) &= 32 . \end{aligned}$$

4.3.6

With $A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 7 & 9 \\ 3 & 9 & 7 \end{pmatrix}$ we know from the previous exercise that

$$A^{-1} = \begin{pmatrix} 4 & -3 & 1 \\ -\frac{13}{8} & 1 & -\frac{1}{8} \\ \frac{3}{8} & 0 & -\frac{1}{8} \end{pmatrix} .$$

4.3.6 continued

Therefore

(a) AX = C means, first of all, since A is 3 x 3 and C is 3 x 2 that X must be 3 x 2. That is, AX must be 3 x 2 and the number of rows in X must equal the number of columns in A.

At any rate,

AX = C

implies

$X = A^{-1}C$

provided, of course that $A^{-1}C$ makes sense. Since A^{-1} is 3 x 3 and C is 3 x 2, the product is well-defined.

Hence,

$$X = \begin{pmatrix} 4 & -3 & 1 \\ -\frac{13}{8} & 1 & -\frac{1}{8} \\ -\frac{3}{8} & 0 & -\frac{1}{8} \end{pmatrix} \begin{pmatrix} 0 & 8 \\ 4 & 2 \\ 8 & 16 \end{pmatrix}$$

$$= \begin{pmatrix} 0 - 12 + 8 & 32 - 6 + 16 \\ 0 + 4 - 1 & -13 + 2 - 2 \\ 0 + 0 - 1 & 3 + 0 - 2 \end{pmatrix}$$

$$= \begin{pmatrix} -4 & 42 \\ 3 & -13 \\ -1 & 1 \end{pmatrix}.$$

Check

$$AX = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 7 & 9 \\ 3 & 9 & 7 \end{pmatrix} \begin{pmatrix} -4 & 42 \\ 3 & -13 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -4 + 9 - 5 & 42 - 39 + 5 \\ -8 + 21 - 9 & 84 - 91 + 9 \\ -12 + 27 - 7 & 126 - 117 + 7 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 8 \\ 4 & 2 \\ 8 & 16 \end{pmatrix} = C.$$

4.3.6 continued

b. Mechanically, we have that $YA = B$ means $Y = BA^{-1}$ [<u>not</u> $A^{-1}B$ since $YA = B$ implies $(YA)A^{-1} = BA^{-1}$]

Therefore,

$$Y = \begin{pmatrix} 0 & 4 & 8 \\ 8 & 2 & 16 \end{pmatrix} \begin{pmatrix} 4 & -3 & 1 \\ -\frac{13}{8} & 1 & -\frac{1}{8} \\ \frac{3}{8} & 0 & -\frac{1}{8} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -\frac{13}{2} + 3 & 0 + 4 + 0 & 0 - \frac{4}{8} - 1 \\ 32 & -\frac{13}{4} + 6 & -24 + 2 + 0 & 8 - \frac{2}{8} - 2 \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{7}{2} & 4 & -\frac{3}{2} \\ \frac{139}{4} & -22 & \frac{23}{4} \end{pmatrix}.$$

c. Notice that B and C are related by $B = C^T$ (where as in the exercises of the previous unit C^T denotes the transpose of C).

The point is that from the equation

$$AX = C \tag{1}$$

it follows that

$$(AX)^T = C^T.$$

Since $(AX)^T = X^T A^T$, this means that

$$X^T A^T = C^T . \tag{2}$$

But in this part A is a <u>symmetric</u> matrix (where by symmetric we mean that $A = A^T$ which in turn means that A remains the same when we interchange its row and columns) which was not the case when A was chosen as in parts (a) and (b).

Since $A^T = A$, equation (2) becomes

4.3.6 continued

$X^T A = C^T$

and since $C^T = B$, we have

$$X^T A = B . \qquad (3)$$

Comparing (2) with (3) we see that if A is symmetric then $X^T A = B$ as soon as $AX = B^T$.

To check this in our particular exercise, we have

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 7 & 0 & 1 & 0 \\ 3 & 7 & 8 & 0 & 0 & 1 \end{array}\right] \sim \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & -3 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 1 & 5 & -2 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & -2 & -1 & -1 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & 0 & 2 & 10 & -4 & 0 \\ 0 & 2 & 2 & -4 & 2 & 0 \\ 0 & 0 & -2 & -1 & -1 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & 0 & 0 & 9 & -5 & 1 \\ 0 & 2 & 0 & -5 & 1 & 1 \\ 0 & 0 & -2 & -1 & -1 & 1 \end{bmatrix} .$$

4.3.6 continued

Therefore

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 7 & 0 & 1 & 0 \\ 3 & 7 & 8 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cccccc} 1 & 0 & 0 & \frac{9}{2} & -\frac{5}{2} & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{5}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{array}\right].$$

This shows us that if A is the symmetric matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 8 \end{pmatrix}$$

then A^{-1} is the symmetric matrix

$$\begin{pmatrix} \frac{9}{2} & -\frac{5}{2} & \frac{1}{2} \\ -\frac{5}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

At any rate, solving AX = C yields

$$X = A^{-1}C$$

or

$$X = \begin{pmatrix} \frac{9}{2} & -\frac{5}{2} & \frac{1}{2} \\ -\frac{5}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 8 \\ 4 & 2 \\ 8 & 16 \end{pmatrix}$$

$$= \begin{pmatrix} -10 + 4 & 36 - 5 + 8 \\ 2 + 4 & -20 + 1 + 8 \\ 2 - 4 & 4 + 1 - 8 \end{pmatrix}$$

4.3.6 continued

$$\begin{pmatrix} -6 & 39 \\ 6 & -11 \\ -2 & -3 \end{pmatrix}. \qquad (4)$$

Similarly YA = B implies $Y = BA^{-1}$, or

$$Y = \begin{pmatrix} 0 & 4 & 8 \\ 8 & 2 & 16 \end{pmatrix} \begin{pmatrix} \frac{9}{2} & -\frac{5}{2} & \frac{1}{2} \\ -\frac{5}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

$$\begin{pmatrix} 0 - 10 + 4 & 0 + 2 + 4 & 0 + 2 - 4 \\ 36 - 5 + 8 & -20 + 1 + 8 & 4 + 1 - 8 \end{pmatrix}$$

$$= \begin{pmatrix} -6 & 6 & -2 \\ 39 & -11 & -3 \end{pmatrix}. \qquad (5)$$

Comparing (4) and (5) we see that

$$Y = X^T$$

in this case.

d. We have

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore

$$\begin{pmatrix} x_{11} + x_{31} & x_{12} + x_{32} \\ x_{21} + x_{31} & x_{22} + x_{32} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

or

$$x_{11} + x_{31} = 1 \qquad (6)$$
$$x_{21} + x_{31} = 0 \qquad (7)$$
$$x_{12} + x_{32} = 0 \qquad (8)$$
$$x_{22} + x_{32} = 1 \qquad (9)$$

4.3.6 continued

From (7) we have $x_{21} = -x_{31}$ and from (6), $x_{11} = 1 - x_{31}$. Similarly from (8) $x_{12} = -x_{32}$ while from (9), $x_{22} = 1 - x_{32}$.

Thus, we may pick x_{31} and x_{32}, say, at random after which we must have $x_{21} = -x_{31}$, $x_{11} = 1 - x_{31}$, $x_{12} = - x_{32}$ and $x_{22} = 1 - x_{32}$.

Hence our matrix X is

$$\begin{pmatrix} 1 - x_{31} & -x_{32} \\ - x_{31} & 1 - x_{32} \\ x_{31} & x_{32} \end{pmatrix}$$

and in this matrix we are free to choose x_{31} and x_{32} at random.

As a check

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 - x_{31} & - x_{32} \\ - x_{31} & 1 - x_{32} \\ x_{31} & x_{32} \end{pmatrix} = \begin{pmatrix} 1 - x_{31} + x_{31} & - x_{32} + x_{32} \\ - x_{31} + x_{31} & 1 - x_{32} + x_{32} \end{pmatrix}$$

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

e. On the other hand, if we look at

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{24} & x_{22} \\ x_{31} & x_{32} \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

we have

$$\begin{pmatrix} x_{11} & x_{12} & x_{11} + x_{12} \\ x_{21} & x_{22} & x_{21} + x_{22} \\ x_{31} & x_{32} & x_{31} + x_{32} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Comparing the first rows of these last two matrices we have $x_{11} = 1$, $x_{12} = 0$, and $x_{11} + x_{12} = 0$ which is impossible since $x_{11} + x_{12} = 1 + 0 = 1$.

4.3.6 continued

Hence, there is no matrix X such that $XA = I_3$.

4.3.7

The main aim of the exercise is to point out the subtlety that we must not make too many changes in our augmented matrix at one time.

The point is that if we start with

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

and replace the first row by the sum of the first and the second, we obtain the new matrix

$$B = \begin{pmatrix} 4 & 6 \\ 3 & 4 \end{pmatrix} .$$

When we now replace the 2nd row by the sum of the first and the second, we are referring to B, not A. That is, we form the new matrix C where

$$C = \begin{pmatrix} 4 & 6 \\ 7 & 10 \end{pmatrix} .$$

In augmented matrix form

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 4 & 6 & 1 & 1 \\ 3 & 4 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 4 & 6 & 1 & 1 \\ 7 & 10 & 1 & 2 \end{bmatrix} .$$

The point is that wherever we operated on more than one row of a matrix at a time, we never touched a row once it was changed. For example, when we subtract the first row from the second, then the first row from the third , etc., none of the rows are being changed more than once in the same augmented matrix and, moreover, no changed row is ever used to operate on an unchanged row. The point of this exercise is to show you that serious errors may occur if we take too many liberties in replacing rows too quickly. When in doubt, stop and begin with the next matrix.