Chapter 8: Matrix Notation and Arithmetic

1. (a) Let the 10 rows represent the 10 respondents, let the columns respectively represent each respondents' union membership, marital status, and religious service attendance. Let 1 indicate that the respondent belongs to a union or is married, and 0 represent that the respondent is not in a union or is not married. Let the last column be represented by the number of services per week the individual attends. The dataset can be written in a table as

Obs.	Union	Married	Religious Services
1	1	0	0
2	0	1	1
3	0	1	1
4	1	1	3
5	1	0	0
6	0	1	5
7	0	0	1
8	1	0	0
9	0	1	1
10	0	1	0

This table can be written as a (10×3) matrix X:

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \\ 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Note that it is no longer necessary to include labels for the rows and columns, as it is with the table. For the matrix, the row and column numbers describe the meaning of each particular datapoint.

(b) These data consist of the distance between each pair of cities within a group of these four: Washington, Boston, New York, and Chicago. Let's create a matrix with 4 rows and 4 columns in each the first row and first column each refer to Washington, the second row and column refer to Boston, the third row and column refer to New York, and the fourth row and column refer to Chicago. The elements of the matrix will be the distance between the city represented by the row and the city represented by the column. Note that there will be 0s on the diagonal because a city is 0 miles away from itself. Also note that the matrix will be symmetric – that is, every element will be equal to the element with the row and column numbers switched – because the distance from New York to Chicago is equal to the distance from Chicago to New York. The (4×4) matrix is

$$X = \begin{bmatrix} 0 & 440 & 229 & 695 \\ 440 & 0 & 213 & 983 \\ 229 & 213 & 0 & 840 \\ 695 & 983 & 840 & 0 \end{bmatrix}.$$

(c) We have data on three countries and 8 years, so let's create a table with 8 rows to represent the years and 3 columns to represent the countries. We only have one economic indicator to consider. For the UK in 2008 we start with 4.8. From the problem we know that the next 7 values are (4.6, 4.4, 4.2, 4.5, 4.8, 5.1, 5.4). For France in 2008 we start with 4.1 and the next seven values are (3.6, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9). For Germany we start with 5.1, and the next seven values are (4.8, 4.5, 4.2, 4.7, 5.2, 5.7, 6.2). Arranging these values in the (8×3) matrix gives us

$$X = \begin{bmatrix} 4.8 & 4.1 & 5.1 \\ 4.6 & 3.6 & 4.8 \\ 4.4 & 3.4 & 4.5 \\ 4.2 & 3.5 & 4.2 \\ 4.5 & 3.6 & 4.7 \\ 4.8 & 3.7 & 5.2 \\ 5.1 & 3.8 & 5.7 \\ 5.4 & 3.9 & 6.2 \end{bmatrix}.$$

(d) There are three patients, and three variables: the times at which each patient is observed, the amount of medication each patient receives, and whether the patient leaves the study or not. Note that patient 3 hasn't yet left the study even at week 7. There are a few ways to arrange these data. But one way is to create a table that has two nested ID variables, a patient identifier and a time identifier. For each combination of patient and time, we can record the dosage and whether or not the patient leaves the study (1 if the patient leaves, 0 if not). We can create a table:

Patient	Week	Dosage	Left Study
1	1	1	0
1	2	1.5	0
1	3	2	1
2	1	1	0
2	2	1.5	0
2	3	2	0
2	4	2.5	0
2	5	3	1
3	1	1	0
3	2	1.5	0
3	3	2	0
3	4	2.5	0
3	5	3	0
3	6	3.5	0
3	7	4	0

We can place this table in a (15×3) matrix:

$$X = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1.5 & 0 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \\ 2 & 1.5 & 0 \\ 3 & 2 & 0 \\ 4 & 2.5 & 0 \\ 5 & 3 & 1 \\ 1 & 1 & 0 \\ 2 & 1.5 & 0 \\ 3 & 2 & 0 \\ 4 & 2.5 & 0 \\ 5 & 3 & 0 \\ 6 & 3.5 & 0 \\ 7 & 4 & 0 \end{bmatrix}.$$

These data are a basic example of a dataset that can be analyzed using survival models, also known as duration or event history models. These models can be used to understand medical prognoses, regime survival, the length of marriages, the duration of wars, and other social phenomena.

(e) There are two voters each making a choice between 3 parties. There are two variables measured: the voters' favorability rating of each party and the voters' choice among the parties. Let's create a data table with 6 rows – one for each unique combination of voter and party – and two columns for the favorability and vote variables. Let's denote favorability on a 1 to 5 scale, where 1 means "very unfavorable," 2 means "slightly unfavorable," 3 means "neutral," 4 means "slightly favorable," and 5 means "very favorable." Let's denote the vote as a binary outcome where 1 indicates a vote for the party and 0 indicates that the voter did not cote for the party. The table is

Voter	Party	Favorability	Vote
1	Labour	5	1
1	Conservative	1	0
1	Liberal	2	0
2	Labour	2	0
2	Conservative	5	1
2	Liberal	4	0

We can place this table in a (6×2) matrix:

$$X = \begin{bmatrix} 5 & 1 \\ 1 & 0 \\ 2 & 0 \\ 2 & 0 \\ 5 & 1 \\ 4 & 0 \end{bmatrix}.$$

2. (a) 12 is just a number. But in matrix terms, we can think of it as a matrix with only one row and one column. That makes 12 a scalar.

   (b) This is a matrix with 3 rows and 1 column. That makes this object a column vector.

(c) This is a matrix with 1 row and 3 columns. That makes this object a row vector.

(d) This matrix has two rows and two columns. It is a matrix, and that its elements are not numbers does not change that fact. Remember that a matrix is a two-dimensional array of objects that are often, *but not necessarily*, numbers. Since this matrix has the same number of rows and columns, it is a square matrix.

(e) This matrix has three rows and three columns. It is a matrix, and as with part (d) that its elements are not numbers does not change that fact. Since this matrix has the same number of rows and columns, it is a square matrix. Also, remember that a matrix is a two-dimensional array of objects in which the row and column positions have meaning. In this case the first row and column each represent zombies, the second row and column represent ninjas, and the third row and column represent pirates. Each element is the combination of the categories represented by the row and column. Since the combination of zombies and ninjas produce the same result as the combination of ninjas and zombies (the dreaded "zombie ninjas"), each element is equal to the one in which the row and column positions are switched. Therefore this matrix is a symmetric matrix.

(f) This matrix has the same number of rows and columns, so it is square. The dots indicate that the upper triangular elements are equal to the elements with the row and column numbers switched. Therefore this matrix is symmetric.

(g) This matrix has the same number of rows and columns, so it is square. All of the elements above the diagonal are zero, so this is a lower-triangular matrix.

(h) This matrix has the same number of rows and columns, so it is square. Each element is equal to the one in which the row and column positions are switched, so this matrix is a symmetric matrix. All of the elements above and below the diagonal are zero, so this is a diagonal matrix. Furthermore, all of the elements on the diagonal are equal, so this is a scalar matrix. Finally, since all of the diagonal elements are equal to 1, this is an identity matrix. Since there are 9 rows and columns, the name of this identity matrix is I_9.

(i) This matrix has the same number of rows and columns, so it is square, but it is not symmetric or an upper or lower triangular matrix. Instead, observe that the matrix can be broken into partitions as follows:

$$\left[\begin{array}{rrr|rr} 8 & 7 & -9 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & -7 & 8 \\ 0 & 0 & 0 & 12 & 2 \\ 0 & 0 & 0 & -9 & 7 \end{array}\right].$$

So this matrix is a partitioned matrix. We can rewrite this matrix as

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where

$$A = \begin{bmatrix} 8 & 7 & -9 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 6 \\ -7 & 8 \\ 12 & 2 \\ -9 & 7 \end{bmatrix}.$$

Since matrices B and C are off-diagonal in the partitioned matrix and since they consist of all zeroes, this matrix is a block-diagonal matrix.

3. (a) First, we apply scalar multiplication to find $3A$ and $2B$:

$$3A = 3\begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \times -1 \\ 3 \times 3 \\ 3 \times 4 \end{bmatrix} = \begin{bmatrix} -3 \\ 9 \\ 12 \end{bmatrix}.$$

$$2B = 2\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \times 3 \\ 2 \times -2 \\ 2 \times 1 \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 2 \end{bmatrix}.$$

Addition and subtraction of two matrices or vectors is conformable (that is, possible) only if the dimensions of the two matrices/vectors are exactly the same. In this case, $3A$ and $2B$ are both (3×1) vectors, so subtraction is conformable, and the difference is a (3×1) matrix whose elements are differences of the corresponding elements of the two vectors:

$$3A - 2B = \begin{bmatrix} -3 \\ 9 \\ 12 \end{bmatrix} - \begin{bmatrix} 6 \\ -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -3-6 \\ 9-(-4) \\ 12-2 \end{bmatrix} = \begin{bmatrix} -9 \\ 13 \\ 10 \end{bmatrix}.$$

(b) $A \cdot B$ is the inner-product (or dot-product) of two vectors, and to find it we multiply the corresponding elements of each vector together, and sum these products:

$$A \cdot B = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = (-1 \times 3) + (3 \times -2) + (4 \times 1) = -3 - 6 + 4 = -5.$$

(c) $A \times B$ is the outer-product (or cross-product) of two vectors, and to find it we multiply each element of A by each element of B, and we align these products in a matrix whose row number corresponds to the row of the element from A and whose column number corresponds to the row of the element from B:

$$A \times B = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} \times \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \times 3 & -1 \times -2 & -1 \times 1 \\ 3 \times 3 & 3 \times -2 & 3 \times 1 \\ 4 \times 3 & 4 \times -2 & 4 \times 1 \end{bmatrix} = \begin{bmatrix} -3 & 2 & -1 \\ 9 & -6 & 3 \\ 12 & -8 & 4 \end{bmatrix}.$$

(d) CA is an example of matrix multiplication, which is conformable only if the number of columns of the left factor equal the number of rows in the right factor. If multiplication is conformable, then the product is a matrix with the same number of rows as the left factor and the same number of columns as the right factor. In this case, C is a (2×3) matrix and A is a (3×1) vector. Therefore multiplication is conformable, and the product is a (2×1) vector.

In general, the (i, j) element of the product is the inner-product of the ith row of the left factor and the jth column of the right factor. in this case the product is

$$CA = \begin{bmatrix} 3 & 2 & -4 \\ -8 & 0 & 6 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} (3 \times -1) + (2 \times 3) + (-4 \times 4) \\ (-8 \times -1) + (0 \times 3) + (6 \times 4) \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \end{bmatrix}.$$

(e) B is a (3×1) vector and D is a (3×2) matrix. In this case the number of columns of the left factor does not equal the number of columns of the right factor. So matrix multiplication is not conformable.

(f) $B \otimes D$ is the Kronecker product of matrices B and D. To find the Kronecker product, we scalar multiply the entire right factor by every element of the left factor, and expand the number of rows and columns of the product as necessary. In this case the Kronecker product is

$$B \otimes D = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix}$$

$$= \begin{bmatrix} 3 \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix} \\ \\ -2 \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix} \\ \\ 1 \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 18 & -6 \\ -3 & 9 \\ -9 & 24 \\ -12 & 4 \\ 2 & -6 \\ 6 & -16 \\ 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix}.$$

(g) C is a (2×3) matrix and D is a (3×2) matrix. Since the number of columns of the left factor equal the number of rows of the right factor, matrix multiplication is conformable and the product is (2×2).

Each element of the product is the inner-product of the corresponding row in the left factor and the corresponding column in the right factor. In this case the product is

$$CD = \begin{bmatrix} 3 & 2 & -4 \\ -8 & 0 & 6 \end{bmatrix} \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix}$$

$$= \begin{bmatrix} (3 \times 6) + (2 \times -1) + (-4 \times -3) & (3 \times -2) + (2 \times 3) + (-4 \times 8) \\ \\ (-8 \times 6) + (0 \times -1) + (6 \times -3) & (-8 \times -2) + (0 \times 3) + (6 \times 8) \end{bmatrix} = \begin{bmatrix} 28 & -32 \\ -66 & 64 \end{bmatrix}.$$

(h) D is a (3×2) matrix and C is a (2×3) matrix. Since the number of columns of the left factor equal the number of rows of the right factor, matrix multiplication is conformable and the product is (3×3).

Each element of the product is the inner-product of the corresponding row in the left factor and the corresponding column in the right factor. In this case the product is

$$DC = \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix} \begin{bmatrix} 3 & 2 & -4 \\ -8 & 0 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} (6\times 3)+(-2\times -8) & (6\times 2)+(-2\times 0) & (6\times -4)+(-2\times 6) \\ (-1\times 3)+(3\times -8) & (-1\times 2)+(3\times 0) & (-1\times -4)+(3\times 6) \\ (-3\times 3)+(8\times -8) & (-3\times 2)+(8\times 0) & (-3\times -4)+(8\times 6) \end{bmatrix} = \begin{bmatrix} 34 & 12 & -36 \\ -27 & -2 & 22 \\ -73 & -6 & 60 \end{bmatrix}.$$

The fact that our calculations for CD and DC differ demonstrates that the order of matrix multiplication can change the answer.

(i) The transpose of matrix C contains the elements on C, but the row and column position of each element is switched:

$$C' = \begin{bmatrix} 3 & -8 \\ 2 & 0 \\ -4 & 6 \end{bmatrix}.$$

C' is a (3×2) matrix and C is a (2×3) matrix. Since the number of columns of the left factor equal the number of rows of the right factor, matrix multiplication is conformable and the product is (3×3).

Each element of the product is the inner-product of the corresponding row in the left factor and the corresponding column in the right factor. In this case the product is

$$C'C = \begin{bmatrix} 3 & -8 \\ 2 & 0 \\ -4 & 6 \end{bmatrix} \begin{bmatrix} 3 & 2 & -4 \\ -8 & 0 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} (3\times 3)+(-8\times -8) & (3\times 2)+(-8\times 0) & (3\times -4)+(-8\times 6) \\ (2\times 3)+(0\times -8) & (2\times 2)+(0\times 0) & (2\times -4)+(0\times 6) \\ (-4\times 3)+(6\times -8) & (-4\times 2)+(6\times 0) & (-4\times -4)+(6\times 6) \end{bmatrix} = \begin{bmatrix} 73 & 6 & -60 \\ 6 & 4 & -8 \\ -60 & -8 & 52 \end{bmatrix}.$$

4. A symmetric matrix is a square matrix that has the property that any element is equal to the element with switched row and column numbers. The answer in question 3, part (i) is symmetric because it is square and the (1,2), (1,3), and (2,3) elements are equal to the (2,1), (3,1), and (3,2) elements respectively.

   Any matrix that is left-multiplied by its transpose is symmetric because the ith row in the transpose is equal to the ith column in the original matrix. Consider a general element in the (i, j) position, where $i \neq j$. This element is equal to the inner-product of the ith row of the transpose and the jth column of the original matrix. The symmetric element in the (j, i) position is equal to the inner-product of the jth row of the transpose and the ith column of the original matrix. But since we are multiplying by the transpose, these two inner-products must be exactly equal. Therefore every (i, j) element is equal to the (j, i) element and the matrix is symmetric.

   The same argument applies when a matrix is right-multiplied by its transpose. So CC' must also be a symmetric matrix.

5. The transpose of the X matrix, X', switches the row and column position of every element in X and is

therefore a (2×10) matrix:

$$X' = \begin{bmatrix} 2 & 4 & -12 & -5 & 9 & 6 & -7 & 7 & -13 & 2 \\ -7 & 3 & 9 & -8 & 0 & -1 & 5 & 10 & -2 & -4 \end{bmatrix},$$

The product $X'X$ multiplies a (2×10) by a (10×2) matrix, which is conformable to matrix multiplication and results in a (2×2) matrix. The (1,1) element is the dot product of the first row of X' and the first column of X:

$$(2\times 2)+(4\times 4)+(-12\times -12)+(-5\times -5)+(9\times 9)+(6\times 6)+(-7\times -7)+(7\times 7)+(-13\times -13)+(2\times 2) = 577.$$

The (1,2) element is the dot product of the first row of X' and the second column of X:

$$(2\times -7)+(4\times 3)+(-12\times 9)+(-5\times -8)+(9\times 0)+(6\times -1)+(-7\times 5)+(7\times 10)+(-13\times -2)+(2\times -4) = -23.$$

Note that the calculation for the (2,1) element – the dot product of the second row of X' and the first column of X – results in exactly the same calculation we encountered for the (1,2) element:

$$(-7\times 2)+(3\times 4)+(9\times -12)+(-8\times -5)+(0\times 9)+(-1\times 6)+(5\times -7)+(10\times 7)+(-2\times -13)+(-4\times 2) = -23.$$

The (2,2) element is the dot product of the second row of X' and the second column of X:

$$(-7\times -7)+(3\times 3)+(9\times 9)+(-8\times -8)+(0\times 0)+(-1\times -1)+(5\times 5)+(10\times 10)+(-2\times -2)+(-4\times -4) = 349.$$

Put together, the whole product is

$$X'X = \begin{bmatrix} 577 & -23 \\ -23 & 349 \end{bmatrix}.$$

Multiplying this matrix by $\frac{1}{10}$ is scalar multiplication, which involves multiplying every element of the matrix by the scalar. Applying this multiplication, we find that

$$\frac{1}{10}X'X = \begin{bmatrix} 57.7 & -2.3 \\ -2.3 & 34.9 \end{bmatrix}.$$

6. (a) First let's compute the product AB. This matrix is the product of a (3×2) and (2×3) matrix, so it is conformable and (3×3). The product is

$$AB = \begin{bmatrix} (10\times -4)+(-2\times 2) & (10\times 9)+(-2\times 1) & (10\times -12)+(-2\times 6) \\ (-1\times -4)+(6\times 2) & (-1\times 9)+(6\times 1) & (-1\times -12)+(6\times 6) \\ (8\times -4)+(3\times 2) & (8\times 9)+(3\times 1) & (8\times -12)+(3\times 6) \end{bmatrix} = \begin{bmatrix} -44 & 88 & -132 \\ 16 & -3 & 48 \\ -26 & 75 & -78 \end{bmatrix}.$$

The trace of AB is the sum of the diagonal elements:

$$tr(AB) = -44 - 3 - 78 = -125.$$

Next let's compute the product BA. This matrix is the product of a (2×3) and (3×2) matrix, so it is conformable and (2×2). The product is

$$BA = \begin{bmatrix} (-4\times 10)+(9\times -1)+(-12\times 8) & (-4\times -2)+(9\times 6)+(-12\times 3) \\ (2\times 10)+(1\times -1)+(6\times 8) & (2\times -2)+(1\times 6)+(6\times 3) \end{bmatrix} = \begin{bmatrix} -145 & 26 \\ 67 & 20 \end{bmatrix}.$$

The trace of BA is the sum of the diagonal elements:

$$tr(BA) = -145 + 20 = -125.$$

We have confirmed that $tr(AB) = tr(BA)$ in this case. This property is true in general beyond this special case.

(b) First let's compute the product CD. This matrix is the product of a (2×3) and (3×2) matrix, so it is conformable and (2×2). The product is

$$CD = \begin{bmatrix} (6 \times -1) + (1 \times 4) + (-2 \times 5) & (6 \times 3) + (1 \times -3) + (-2 \times 2) \\ (4 \times -1) + (3 \times 4) + (7 \times 5) & (4 \times 3) + (3 \times -3) + (7 \times 2) \end{bmatrix} = \begin{bmatrix} -12 & 11 \\ 43 & 17 \end{bmatrix}.$$

The transpose of CD switches the row and column number of every element of CD:

$$(CD)' = \begin{bmatrix} -12 & 43 \\ 11 & 17 \end{bmatrix}.$$

Likewise, the transposes of C and D individually switch the row and column numbers of C and D,

$$C' = \begin{bmatrix} 6 & 4 \\ 1 & 3 \\ -2 & 7 \end{bmatrix}, \qquad D' = \begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix}$$

and the product $D'C'$ is the product of a (2×3) and (3×2) matrix, so it is conformable and (2×2). The product is

$$D'C' = \begin{bmatrix} (-1 \times 6) + (4 \times 1) + (5 \times -2) & (-1 \times 4) + (4 \times 3) + (5 \times 7) \\ (3 \times 6) + (-3 \times 1) + (2 \times -2) & (3 \times 4) + (-3 \times 3) + (2 \times 7) \end{bmatrix} = \begin{bmatrix} -12 & 43 \\ 11 & 17 \end{bmatrix}.$$

We've confirmed that $(CD)' = D'C'$. This property is also true in general situations.

(c) A and D are both (3×2) matrices, and since they have the same dimensions they are conformable to addition. The sum of these matrices is

$$A + D = \begin{bmatrix} 10 - 1 & -2 + 3 \\ -1 + 4 & 6 - 3 \\ 8 + 5 & 3 + 2 \end{bmatrix} = \begin{bmatrix} 9 & 1 \\ 3 & 3 \\ 13 & 5 \end{bmatrix}.$$

The transpose of $A + D$ switches the row and column numbers of every element of $A + D$:

$$(A + D)' = \begin{bmatrix} 9 & 3 & 13 \\ 1 & 3 & 5 \end{bmatrix}.$$

Likewise the transposes of A and D individually are

$$A' = \begin{bmatrix} 10 & -1 & 8 \\ -2 & 6 & 3 \end{bmatrix}, \qquad D' = \begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix},$$

and the sum of A' and D' is

$$A' + D' = \begin{bmatrix} 10 - 1 & -1 + 4 & 8 + 5 \\ -2 + 3 & 6 - 3 & 3 + 2 \end{bmatrix} = \begin{bmatrix} 9 & 3 & 13 \\ 1 & 3 & 5 \end{bmatrix}.$$

We've demonstrated that $(A + D)' = A' + D'$ in this case, and this property is true in general as well.

(d) The scalar product $5B$ is equal to

$$5B = 5 \begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} = \begin{bmatrix} -20 & 45 & -60 \\ 10 & 5 & 30 \end{bmatrix}.$$

The transpose of $5B$ is

$$(5B)' = \begin{bmatrix} -20 & 10 \\ 45 & 5 \\ -60 & 30 \end{bmatrix}.$$

The scalar product $5B'$ is equal to

$$5B' = 5 \begin{bmatrix} -4 & 2 \\ 9 & 1 \\ -12 & 6 \end{bmatrix} = \begin{bmatrix} -20 & 10 \\ 45 & 5 \\ -60 & 30 \end{bmatrix}.$$

We've demonstrated that $(5B)' = 5B'$ in this case, and this property is true in general as well.

(e) The sum $B + C$ equals

$$B + C = \begin{bmatrix} -4+6 & 9+1 & -12-2 \\ 2+4 & 1+3 & 6+7 \end{bmatrix} = \begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix}.$$

The Kronecker product $A \otimes (B + C)$ equals

$$A \otimes (B + C) = \begin{bmatrix} 10 & -2 \\ -1 & 6 \\ 8 & 3 \end{bmatrix} \otimes \begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix}$$

$$= \begin{bmatrix} 10\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} & -2\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} \\ -1\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} & 6\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} \\ 8\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} & 3\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 20 & 100 & -140 & -4 & -20 & 28 \\ 60 & 40 & 130 & -12 & -8 & -26 \\ -2 & -10 & 14 & 12 & 60 & -84 \\ -6 & -4 & -13 & 36 & 24 & 78 \\ 16 & 80 & -112 & 6 & 30 & -42 \\ 48 & 32 & 104 & 18 & 12 & 39 \end{bmatrix}.$$

The Kronecker product $A \otimes B$ is

$$A \otimes B = \begin{bmatrix} 10 & -2 \\ -1 & 6 \\ 8 & 3 \end{bmatrix} \otimes \begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} 10\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} & -2\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} \\ -1\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} & 6\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} \\ 8\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} & 3\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} -40 & 90 & -120 & 8 & -18 & 24 \\ 20 & 10 & 60 & -4 & -2 & -12 \\ 4 & -9 & 12 & -24 & 54 & -72 \\ -2 & -1 & -6 & 12 & 6 & 36 \\ -32 & 78 & -106 & -12 & 27 & -36 \\ 16 & 8 & 48 & 6 & 3 & 18 \end{bmatrix},$$

the Kronecker product $A \otimes C$ is

$$A \otimes B = \begin{bmatrix} 10 & -2 \\ -1 & 6 \\ 8 & 3 \end{bmatrix} \otimes \begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix}$$

$$= \begin{bmatrix} 10\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} & -2\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} \\ -1\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} & 6\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} \\ 8\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} & 3\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 60 & 10 & -20 & -12 & -2 & 4 \\ 40 & 30 & 70 & -8 & -6 & -14 \\ -6 & -1 & 2 & 36 & 6 & -12 \\ -4 & -3 & -7 & 24 & 18 & 42 \\ 48 & 8 & -16 & 18 & 3 & -6 \\ 32 & 24 & 56 & 12 & 9 & 21 \end{bmatrix},$$

and the sum of these two Kronecker products $(A \otimes B) + (A \otimes C)$ is

$$\begin{bmatrix} -40 & 90 & -120 & 8 & -18 & 24 \\ 20 & 10 & 60 & -4 & -2 & -12 \\ 4 & -9 & 12 & -24 & 54 & -72 \\ -2 & -1 & -6 & 12 & 6 & 36 \\ -32 & 78 & -106 & -12 & 27 & -36 \\ 16 & 8 & 48 & 6 & 3 & 18 \end{bmatrix} + \begin{bmatrix} 60 & 10 & -20 & -12 & -2 & 4 \\ 40 & 30 & 70 & -8 & -6 & -14 \\ -6 & -1 & 2 & 36 & 6 & -12 \\ -4 & -3 & -7 & 24 & 18 & 42 \\ 48 & 8 & -16 & 18 & 3 & -6 \\ 32 & 24 & 56 & 12 & 9 & 21 \end{bmatrix}$$

$$= \begin{bmatrix} 20 & 100 & -140 & -4 & -20 & 28 \\ 60 & 40 & 130 & -12 & -8 & -26 \\ -2 & -10 & 14 & 12 & 60 & -84 \\ -6 & -4 & -13 & 36 & 24 & 78 \\ 16 & 80 & -112 & 6 & 30 & -42 \\ 48 & 32 & 104 & 18 & 12 & 39 \end{bmatrix}.$$

We've demonstrated that $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ in this case, and this property is true in general as well.

(f) The Kronecker product $C \otimes D$ is

$$C \otimes D = \begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} \otimes \begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 6\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} & 1\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} & -2\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} \\ 4\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} & 3\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} & 7\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} -6 & 18 & -1 & 3 & 2 & -6 \\ 24 & -18 & 4 & -3 & -8 & 6 \\ 30 & 12 & 5 & 2 & -10 & -4 \\ -4 & 12 & -3 & 9 & -7 & 21 \\ 16 & -12 & 12 & -9 & 28 & -21 \\ 20 & 8 & 15 & 6 & 35 & 14 \end{bmatrix},$$

and the transpose of this matrix is

$$(C \otimes D)' = \begin{bmatrix} -6 & 24 & 30 & -4 & 16 & 20 \\ 18 & -18 & 12 & 12 & -12 & 8 \\ -1 & 4 & 5 & -3 & 12 & 15 \\ 3 & -3 & 2 & 9 & -9 & 6 \\ 2 & -8 & -10 & -7 & 28 & 35 \\ -6 & 6 & -4 & 21 & -21 & 14 \end{bmatrix}.$$

Individually, the transposes of matrices C and D are

$$C' = \begin{bmatrix} 6 & 4 \\ 1 & 3 \\ -2 & 7 \end{bmatrix}, \qquad D' = \begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix},$$

and the Kronecker product $C' \otimes D'$ is

$$C' \otimes D' = \begin{bmatrix} 6 & 4 \\ 1 & 3 \\ -2 & 7 \end{bmatrix} \otimes \begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 6\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} & 4\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} \\ 1\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} & 3\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} \\ -2\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} & 7\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} -6 & 24 & 30 & -4 & 16 & 20 \\ 18 & -18 & 12 & 12 & -12 & 8 \\ -1 & 4 & 5 & -3 & 12 & 15 \\ 3 & -3 & 2 & 9 & -9 & 6 \\ 2 & -8 & -10 & -7 & 28 & 35 \\ -6 & 6 & -4 & 21 & -21 & 14 \end{bmatrix}.$$

We've demonstrated that $(C \otimes D)' = C' \otimes D'$ in this case, and this property is true in general as well.

(g) We previously calculated $A + D$ in part (c):

$$A + D = \begin{bmatrix} 9 & 1 \\ 3 & 3 \\ 13 & 5 \end{bmatrix}.$$

The product $B(A+D)$ multiplies a (2×3) matrix by a (3×2) matrix. The product, which is conformable and (2×2), is

$$B(A+D) = \begin{bmatrix} (-4\times9)+(9\times3)+(-12\times13) & (-4\times1)+(9\times3)+(-12\times5) \\ (2\times9)+(1\times3)+(6\times13) & (2\times1)+(1\times3)+(6\times5) \end{bmatrix} = \begin{bmatrix} -165 & -37 \\ 99 & 35 \end{bmatrix}.$$

We previously calculated the product BA in part (a):

$$BA = \begin{bmatrix} -145 & 26 \\ 67 & 20 \end{bmatrix},$$

and the product BD multiplies a (2×3) matrix by a (3×2) matrix. The product, which is conformable and (2×2), is

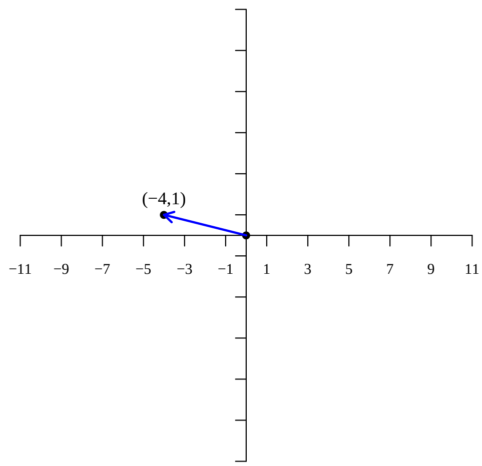
$$BD = \begin{bmatrix} (-4\times-1)+(9\times4)+(-12\times5) & (-4\times3)+(9\times-3)+(-12\times2) \\ (2\times-1)+(1\times4)+(6\times5) & (2\times3)+(1\times-3)+(6\times2) \end{bmatrix} = \begin{bmatrix} -20 & -63 \\ 32 & 15 \end{bmatrix}.$$

Finally, the sum of BA and BD is

$$BA + BD = \begin{bmatrix} -145 & 26 \\ 67 & 20 \end{bmatrix} + \begin{bmatrix} -20 & -63 \\ 32 & 15 \end{bmatrix} = \begin{bmatrix} -165 & -37 \\ 99 & 35 \end{bmatrix}.$$

We've demonstrated that $B(A+D) = BA+BD$ in this case, and this property is true in general as well.

7. (a) A plot of this vector is



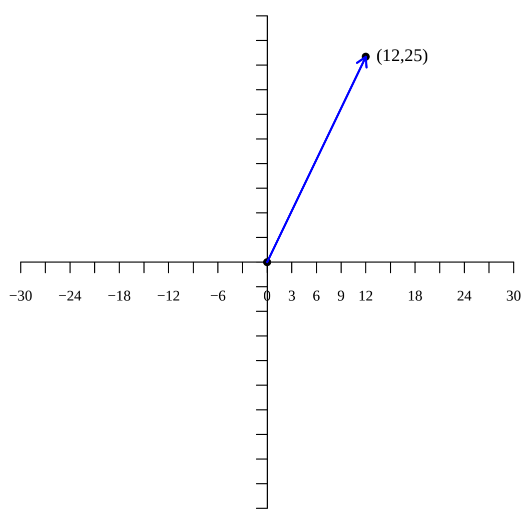
The magnitude of the vector is

$$\left\| \begin{bmatrix} -4 \\ 1 \end{bmatrix} \right\| = \sqrt{(-4)^2 + (1)^2} = \sqrt{17} = 4.12.$$

A unit vector in the same direction is

$$\frac{1}{4.12} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.97 \\ 0.24 \end{bmatrix}.$$

(b) A plot of this vector is



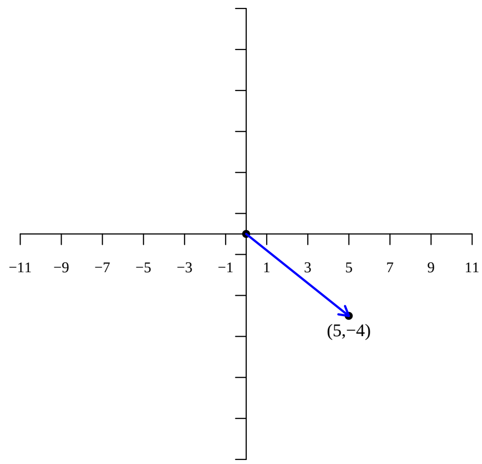
The magnitude of the vector is

$$\left\| \begin{bmatrix} 12 \\ 25 \end{bmatrix} \right\| = \sqrt{(12)^2 + (25)^2} = \sqrt{769} = 27.73.$$

A unit vector in the same direction is

$$\frac{1}{27.73} \begin{bmatrix} 12 \\ 25 \end{bmatrix} = \begin{bmatrix} 0.43 \\ 0.90 \end{bmatrix}.$$

(c) A plot of this vector is



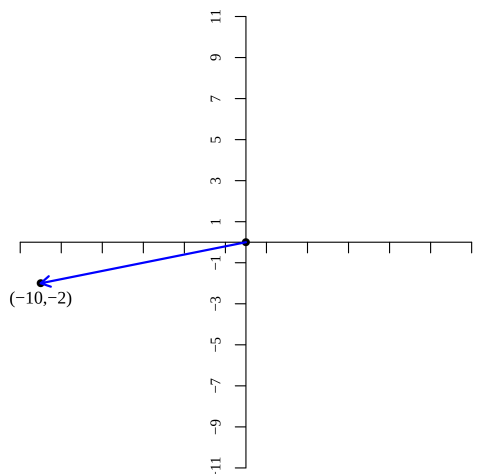
The magnitude of the vector is

$$\left|\left|\begin{bmatrix} 5 \\ -4 \end{bmatrix}\right|\right| = \sqrt{(5)^2 + (-4)^2} = \sqrt{41} = 6.4.$$

A unit vector in the same direction is

$$\frac{1}{6.4}\begin{bmatrix} 5 \\ -4 \end{bmatrix} = \begin{bmatrix} 0.78 \\ -0.62 \end{bmatrix}.$$

(d) A plot of this vector is



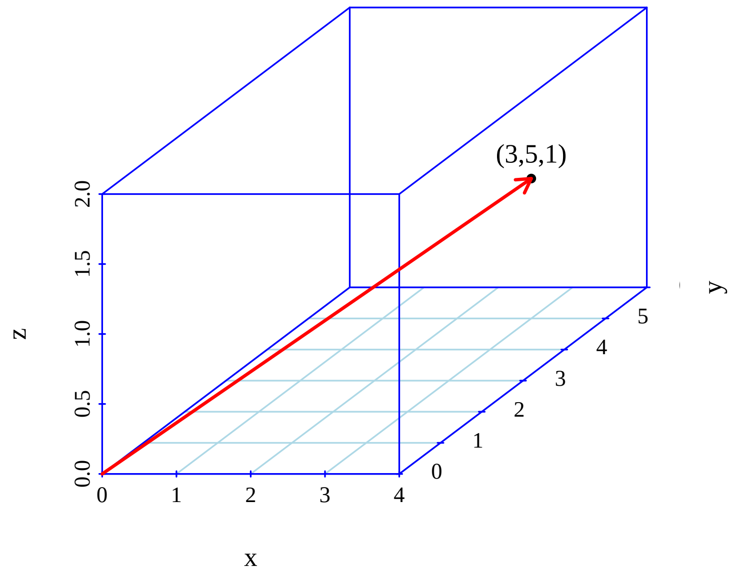
The magnitude of the vector is

$$\left|\left|\begin{bmatrix} -10 \\ -2 \end{bmatrix}\right|\right| = \sqrt{(-10)^2 + (-2)^2} = \sqrt{104} = 6.4.$$

A unit vector in the same direction is

$$\frac{1}{6.4}\begin{bmatrix} -10 \\ -2 \end{bmatrix} = \begin{bmatrix} -0.98 \\ -0.20 \end{bmatrix}.$$

(e) A plot of this vector is



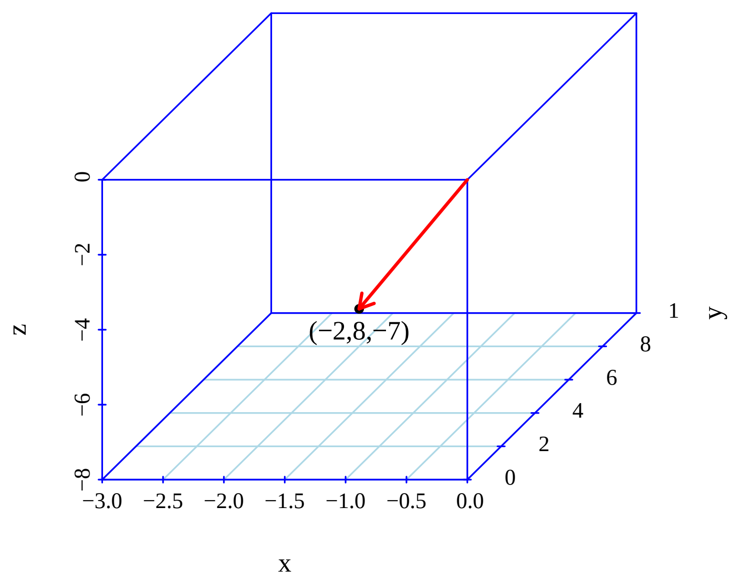
The magnitude of the vector is

$$\left| \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} \right| = \sqrt{(3)^2 + (5)^2 + (1)^2} = \sqrt{35} = 5.92.$$

A unit vector in the same direction is

$$\frac{1}{5.92} \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.51 \\ 0.84 \\ 0.17 \end{bmatrix}.$$

(f) A plot of this vector is



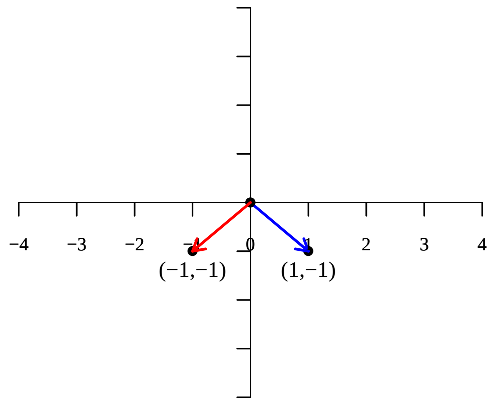
The magnitude of the vector is

$$\left| \begin{bmatrix} -2 \\ 8 \\ -7 \end{bmatrix} \right| = \sqrt{(-2)^2 + (8)^2 + (-7)^2} = \sqrt{117} = 10.82.$$

A unit vector in the same direction is

$$\frac{1}{10.82} \begin{bmatrix} -2 \\ 8 \\ -7 \end{bmatrix} = \begin{bmatrix} -0.18 \\ 0.74 \\ -0.65 \end{bmatrix}.$$

8. The following plot indicates that the vector that is turned 90 degrees clockwise is $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$:



The trick is finding a matrix that when left-multiplied by the vector yields a product of $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$. Consider a general (2×2) matrix as a left-factor:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix},$$

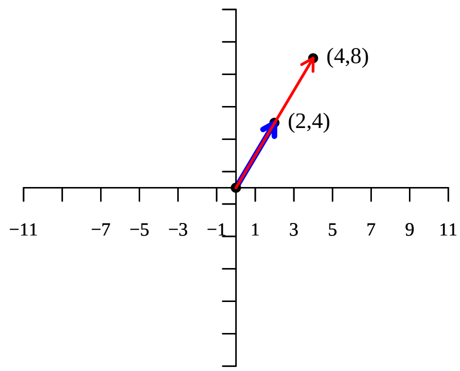
which will be true only if

$$\begin{cases} a - b = -1, \\ c - d = -1. \end{cases}$$

There are many matrices that perform this transformation, but one example is the (2×2) matrix where $a = 2$, $b = 3$, $c = 4$, and $d = 5$. To see that this works, we compute the product:

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

9. Any multiple of $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ is in the same direction. Consider, for example, $\begin{bmatrix} 4 \\ 8 \end{bmatrix}$. Both vectors are graphed below:



The trick is finding a matrix that when left-multiplied by the vector yields a product of $\begin{bmatrix} 4 \\ 8 \end{bmatrix}$. Consider a general (2×2) matrix as a left-factor:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix},$$

which will be true only if

$$\begin{cases} 2a + 4b = 4, \\ 2c + 4d = 8. \end{cases}$$

There are many matrices that perform this transformation, but one example is the (2×2) matrix where $a = 0$, $b = 1$, $c = 2$, and $d = 1$. To see that this works, we compute the product:

$$\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 + 4 \\ 4 + 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}.$$

10. (a) We begin with

$$\begin{bmatrix} 2 & 1 & 6 \\ 8 & 4 & 3 \\ 12 & 1 & 4 \end{bmatrix}.$$

We multiply the first row by -4 and add it to the second row,

$$\begin{bmatrix} 2 & 1 & 6 \\ 0 & 0 & -21 \\ 12 & 1 & 4 \end{bmatrix},$$

then multiply the first row by -6 and add it to the third row,

$$\begin{bmatrix} 2 & 1 & 6 \\ 0 & 0 & -21 \\ 0 & -5 & -32 \end{bmatrix}.$$

We interchange second and third rows,

$$\begin{bmatrix} 2 & 1 & 6 \\ 0 & -5 & -32 \\ 0 & 0 & -21 \end{bmatrix},$$

divide the third row by -21,

$$\begin{bmatrix} 2 & 1 & 6 \\ 0 & -5 & -32 \\ 0 & 0 & 1 \end{bmatrix},$$

multiply the third row by 32 and add it to the second row,

$$\begin{bmatrix} 2 & 1 & 6 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and multiply it by -6 and add it to the first row,

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

We divide the second row by -5,

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

multiply it by -1, and add it to the first row,

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

Finally we divide the first row by 2,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(b) We begin with

$$\begin{bmatrix} 1 & -3 & 4 \\ 3 & 7 & -4 \\ 5 & 9 & 12 \end{bmatrix}.$$

We multiply the first row by -3 and add it to the second row,

$$\begin{bmatrix} 1 & -3 & 4 \\ 0 & 16 & -16 \\ 5 & 9 & 12 \end{bmatrix},$$

and multiply it by -5 and add it to the third row,

$$\begin{bmatrix} 1 & -3 & 4 \\ 0 & 16 & -16 \\ 0 & 24 & -8 \end{bmatrix}.$$

We divide the second row by 16 and the third row by 8,

$$\begin{bmatrix} 1 & -3 & 4 \\ 0 & 1 & -1 \\ 0 & 3 & -1 \end{bmatrix}.$$

We then multiply the second row by -3 and add it to the third row,

$$\begin{bmatrix} 1 & -3 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix},$$

and divide the third row by 2,

$$\begin{bmatrix} 1 & -3 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now we add the third row to the second,

$$\begin{bmatrix} 1 & -3 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

we multiply the third row by -4 and add it to the first row,

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

and we multiply the second row by 3 and add it to the first row,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(c) We begin with

$$\begin{bmatrix} 3 & 4 & 7 \\ 9 & -3 & 1 \\ -3 & 11 & 13 \end{bmatrix}.$$

We multiply the first row by -3 and add it to the second row,

$$\begin{bmatrix} 3 & 4 & 7 \\ 0 & -15 & -20 \\ -3 & 11 & 13 \end{bmatrix},$$

and add the first row to the third row,

$$\begin{bmatrix} 3 & 4 & 7 \\ 0 & -15 & -20 \\ 0 & 15 & 20 \end{bmatrix}.$$

Next we add the second row to the third,

$$\begin{bmatrix} 3 & 4 & 7 \\ 0 & -15 & -20 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since we now have a row of all zeroes, it will not be possible to reduce this matrix to an identity matrix. Let's continue reducing the matrix as much as possible. Next we divide the second row by -5,

$$\begin{bmatrix} 3 & 4 & 7 \\ 0 & 3 & 4 \\ 0 & 0 & 0 \end{bmatrix},$$

and in order to avoid fractions, we multiply the first row by 3 and the second row by 4,

$$\begin{bmatrix} 9 & 12 & 21 \\ 0 & 12 & 16 \\ 0 & 0 & 0 \end{bmatrix}.$$

We multiply the second row by -1 and add it to the first,

$$\begin{bmatrix} 9 & 0 & 5 \\ 0 & 12 & 16 \\ 0 & 0 & 0 \end{bmatrix}.$$

Finally – there's no avoiding fractions any longer – we divide the first row by 9 and the second row by 12,

$$\begin{bmatrix} 1 & 0 & \frac{5}{9} \\ 0 & 1 & \frac{3}{4} \\ 0 & 0 & 0 \end{bmatrix}.$$

This matrix is reduced as much as possible because the first two rows and columns contain the I_2 identity matrix. There are two rows that do not consist entirely of zeroes.

(d) We begin with

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 6 & 6 & -8 & 3 \\ -4 & 0 & 2 & 9 \\ 10 & 0 & 6 & -2 \end{bmatrix}.$$

We multiply the first row by -3 and add it to the second,

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ -4 & 0 & 2 & 9 \\ 10 & 0 & 6 & -2 \end{bmatrix},$$

then multiply it by 2 and add it to the third row (we don't worry about erasing the zero in the (3,2) spot even though eventually we want this element to be 0 – we will come back to fix this element eventually),

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 6 & 4 & -5 \\ 10 & 0 & 6 & -2 \end{bmatrix},$$

and then multiply it by -5 and add it to the fourth row (we will also fix the (4,2) element later),

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 6 & 4 & -5 \\ 0 & -15 & 1 & 33 \end{bmatrix}.$$

Next we multiply the second row by 2 and add it to the third row (see?),

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 0 & -18 & 43 \\ 0 & -15 & 1 & 33 \end{bmatrix},$$

and we multiply it by 5 and add it to the fourth row,

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 0 & -18 & 43 \\ 0 & 0 & -54 & 177 \end{bmatrix}.$$

We multiply the third row by -3 and add it to the fourth row,

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 0 & -18 & 43 \\ 0 & 0 & 0 & 48 \end{bmatrix},$$

and divide the fourth row by 48,

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 0 & -18 & 43 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

We can take a series of easy and useful steps now. We multiply the fourth row by -43 and add it to the third row, then we multiply it by -24 and add it to the second row, and multiply it by 7 and add it to the first row,

$$\begin{bmatrix} 2 & 3 & 1 & 0 \\ 0 & -3 & -11 & 0 \\ 0 & 0 & -18 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Next we divide the third row by -18,

$$\begin{bmatrix} 2 & 3 & 1 & 0 \\ 0 & -3 & -11 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

then multiply it by 11 and add it to the second row, and multiply it by -1 and add it to the first row,

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

We add the second row to the first,

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

and divide the first row by 2 and the second row by -3,

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(e) We begin with

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 3 & 6 & 2 & -2 \\ 4 & 8 & -6 & 1 \\ -2 & -4 & -10 & 5 \end{bmatrix}.$$

We multiply the first row by -3 and add it to the second row,

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 0 & 0 & 26 & -11 \\ 4 & 8 & -6 & 1 \\ -2 & -4 & -10 & 5 \end{bmatrix},$$

then we multiply it by -4 and add it to the third row,

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 0 & 0 & 26 & -11 \\ 0 & 0 & 26 & -11 \\ -2 & -4 & -10 & 5 \end{bmatrix},$$

and then we multiply it by 2 and add it to the fourth row,

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 0 & 0 & 26 & -11 \\ 0 & 0 & 26 & -11 \\ 0 & 0 & -26 & 11 \end{bmatrix}.$$

Next we multiply the second row by -1 and add it to the third row,

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 0 & 0 & 26 & -11 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -26 & 11 \end{bmatrix},$$

and add it to the fourth row as is,

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 0 & 0 & 26 & -11 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We now have two rows that consist entirely of zeroes, so it will not be possible to reduce this matrix to an identity matrix. Let's reduce it as much as possible. Here's where it gets ugly. To avoid fractions, multiply the first row by 13 and the second row by 4,

$$\begin{bmatrix} 13 & 26 & -104 & 36 \\ 0 & 0 & 104 & -44 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and add the second row to the first,

$$\begin{bmatrix} 13 & 26 & 0 & -8 \\ 0 & 0 & 104 & -44 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Finally, divide the first row by 13 and the second row by 104,

$$\begin{bmatrix} 1 & 2 & 0 & -\frac{8}{13} \\ 0 & 0 & 1 & -\frac{11}{26} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We cannot reduce the fraction any further, and there are two rows remaining that do not consist entirely of zeroes.

11. Let's create a (10×1) column vector named Y that contains the values of the dependent variable. Likewise we can create a (10×1) column vector of regression errors named ϵ and a (3×1) column vector called B that contains the constant and the two β coefficients:

$$Y = \begin{bmatrix} 9 \\ 5 \\ 4 \\ 7 \\ 4 \\ 6 \\ 5 \\ 7 \\ 2 \\ 5 \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \end{bmatrix}, \quad B = \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{bmatrix}.$$

The trick is finding a matrix for the X variables that works within a matrix equation that is equivalent to the regression model. The best approach is to set X to be the following (10×3) matrix,

$$X = \begin{bmatrix} 1 & 0 & 8 \\ 1 & 0 & 6 \\ 1 & 1 & 7 \\ 1 & 1 & 8 \\ 1 & 1 & 3 \\ 1 & 0 & 4 \\ 1 & 1 & 6 \\ 1 & 0 & 3 \\ 1 & 1 & 5 \\ 1 & 0 & 4 \end{bmatrix},$$

where the column of all 1s is designed to work with the constant, as we will see below. How were you supposed to know that? Hopefully you gave this problem a lot of thought, and realized that accounting for the constant in the regression equation is the tricky part. If you tried several different approaches, then you are thinking about this problem correctly. Moreover, you are thinking about designing matrices for specific data applications, which is the very best way for a social scientist to think about matrices.

Now that we have defined vectors and matrices for every component of the regression, we can write the regression model as

$$Y = XB + \epsilon.$$

To check whether this equation is conformable, note that X is (10×3) and B is (3×1), so their product is conformable and (10×1). This product is itself conformable to add to the (10×1) vector of errors, and the sum is (10×1), which corresponds to the vector Y which is also (10×1). If we replace these terms with the actual data, the equation becomes

$$\begin{bmatrix} 9 \\ 5 \\ 4 \\ 7 \\ 4 \\ 6 \\ 5 \\ 7 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 8 \\ 1 & 0 & 6 \\ 1 & 1 & 7 \\ 1 & 1 & 8 \\ 1 & 1 & 3 \\ 1 & 0 & 4 \\ 1 & 1 & 6 \\ 1 & 0 & 3 \\ 1 & 1 & 5 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \end{bmatrix},$$

which when multiplied out is the following system of equations:

$$\begin{cases} 9 = 1\alpha + 0\beta_1 + 8\beta_2 + \varepsilon_1, \\ 5 = 1\alpha + 0\beta_1 + 6\beta_2 + \varepsilon_2, \\ 4 = 1\alpha + 1\beta_1 + 7\beta_2 + \varepsilon_3, \\ 7 = 1\alpha + 1\beta_1 + 8\beta_2 + \varepsilon_4, \\ 4 = 1\alpha + 1\beta_1 + 3\beta_2 + \varepsilon_5, \\ 6 = 1\alpha + 0\beta_1 + 4\beta_2 + \varepsilon_6, \\ 5 = 1\alpha + 1\beta_1 + 6\beta_2 + \varepsilon_7, \\ 7 = 1\alpha + 0\beta_1 + 3\beta_2 + \varepsilon_8, \\ 2 = 1\alpha + 1\beta_1 + 5\beta_2 + \varepsilon_9, \\ 5 = 1\alpha + 0\beta_1 + 4\beta_2 + \varepsilon_{10}. \end{cases}$$

Each one of these equations is another realization of the familiar linear regression equation. When we use matrices instead, we can work with all of them simultaneously with the simple equation listed above.