

Matrix Algebra Review

Economics 5330

Tyler J. Brough

October 17, 2016

Introduction

These notes give a brief overview of matrix notation and operations. These tools from mathematics become useful for the linear regression model with multiple regressands.

These notes are based on the following sources:

- Chapter 2 of the Brooks textbook
- Chapter 8 of the book Mathematics & Statistics for Financial Risk Management by Michael B. Miller.

Matrix Notation

A *matrix* is a two-dimensional array of numbers or variables. We note the size of a matrix first by its number of rows and then by its number of columns. Thus, an example of a 3 x 2 (pronounced “three by two”) matrix is the following:

$$\mathbf{A} = \begin{bmatrix} 3 & 5 \\ -9 & 3 \\ 10 & 8 \end{bmatrix}$$

Matrices with a single column or row are called *vectors*. For example, the following is a 4 x 1 column vector:

$$\mathbf{b} = \begin{bmatrix} 43 \\ 17 \\ -56 \\ 64 \end{bmatrix}$$

We denote matrices with boldface capital letters, such as \mathbf{A} and vectors with boldface lowercase letters, such as \mathbf{b} . The elements of a matrix are ordinary real numbers and we refer to them as *scalars*.

We can denote the various elements with subscripts as follows:

$$\mathbf{C} = \begin{bmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{bmatrix}$$

where the first subscript i in $c_{i,j}$ denotes the i^{th} row and the second subscript j denotes the j^{th} column.

For a vector we can usually drop down to just one subscript without loss of information as follows:

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

Matrices such as C , for which there are the same number of rows as columns are called *Square matrices*. The main diagonal consists of the elements running down the diagonal from the top-left to the bottom-right. In other words, the elements $x_{i,j}$ for $i = j$.

If all of the entries above the main diagonal are zero it is called a lower triangular matrix. The following is a 3 x 3 lower triangular matrix:

$$\mathbf{L} = \begin{bmatrix} l_{1,1} & 0 & 0 \\ l_{2,1} & l_{2,2} & 0 \\ l_{3,1} & l_{3,2} & l_{3,3} \end{bmatrix}$$

Similarly, if a matrix has zero entries below the main diagonal, it is called a *upper triangular*:

$$\mathbf{U} = \begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} \\ 0 & u_{2,2} & u_{2,3} \\ 0 & 0 & u_{3,3} \end{bmatrix}$$

If a matrix has all zero entries except for the main diagonal it is called a *diagonal matrix*. The following are diagonal matrices:

$$\begin{bmatrix} 15 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 5.6 & 0 \\ 0 & 23.9 \end{bmatrix} \begin{bmatrix} 47 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Matrix Operations

Just as we can add, subtract and multiply scalars we can do similar operations for matrices. Instead of matrix division, we speak of matrix inversion, which is analogous. There are additional matrix operations, such as the transpose, that are unique to matrices.

Matrix Addition

To add matrices we simply add the corresponding elements in each matrix. But to do so, they have to be *conformable*. That is, each matrix must have the same size. For example, suppose we have the following two matrices, \mathbf{D} and \mathbf{E} :

$$\mathbf{D} = \begin{bmatrix} 32 & 51 \\ -10 & 0 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 25 & -21 \\ 3 & 14 \end{bmatrix}$$

We add them as follows:

$$\mathbf{D} + \mathbf{E} = \begin{bmatrix} (d_{1,1} + e_{1,1}) & (d_{1,2} + e_{1,2}) \\ (d_{2,1} + e_{2,1}) & (d_{2,2} + e_{2,2}) \end{bmatrix} = \begin{bmatrix} 57 & 30 \\ -7 & 14 \end{bmatrix}$$

You can check that matrix multiplication is commutative, i.e., that the order does not matter when we add them:

$$\mathbf{D} + \mathbf{E} = \mathbf{E} + \mathbf{D}$$

Matrix addition is also associative (i.e. the order in which we carry out addition is not important):

$$\mathbf{D} + (\mathbf{E} + \mathbf{F}) = (\mathbf{D} + \mathbf{E}) + \mathbf{F}$$

We can also multiply a matrix by a scalar. The result is a new matrix, the same size as the original, but with all the elements multiplied by the scalar value. For example, using the matrix \mathbf{A} as before and $s = 10$ for the scalar value:

$$s \times \mathbf{A} = 10 \begin{bmatrix} 3 & 5 \\ -9 & 3 \\ 10 & 8 \end{bmatrix} = \begin{bmatrix} 30 & 50 \\ -90 & 30 \\ 100 & 80 \end{bmatrix}$$

For matrix subtraction, we simply subtract the corresponding elements in each matrix. Again, the matrices must be conformable. Using \mathbf{D} and \mathbf{E} as before:

$$\mathbf{D} - \mathbf{E} = \begin{bmatrix} 32 & 51 \\ -10 & 0 \end{bmatrix} - \begin{bmatrix} 25 & -21 \\ 3 & 14 \end{bmatrix}$$

$$\mathbf{D} - \mathbf{E} = \begin{bmatrix} 32 - 25 & 51 + 21 \\ -10 - 3 & 0 - 14 \end{bmatrix}$$

$$\mathbf{D} - \mathbf{E} = \begin{bmatrix} 7 & 72 \\ -13 & -14 \end{bmatrix}$$

Matrix Multiplication

We can also multiply two matrices together. In order to carry out matrix multiplication the matrices must again be conformable. For multiplication this means that the number of columns of the first matrix must be equal to the number of rows of the second. The resulting matrix has the same number of rows as the first matrix and the same number of columns as the second matrix. For example, if we multiply a 3×2 matrix by a 2×5 matrix we obtain a 3×5 matrix.

Each entry in the resulting product matrix is the sum of the products of the corresponding elements. For example, the first element in the product matrix is the sum of the products of the corresponding elements of the first row of the first matrix and the first column of the second matrix. For example, given the matrices \mathbf{G} and \mathbf{H} we obtain their product as follows:

$$\mathbf{G} = \begin{bmatrix} 6 & 1 \\ 9 & 5 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 2 & 3 \\ 8 & 4 \end{bmatrix}$$

$$\mathbf{J} = \mathbf{G} \bullet \mathbf{H} = \begin{bmatrix} 6 & 1 \\ 9 & 5 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 8 & 4 \end{bmatrix} = \begin{bmatrix} 6 \bullet 2 + 1 \bullet 8 & 6 \bullet 3 + 1 \bullet 4 \\ 9 \bullet 2 + 5 \bullet 8 & 9 \bullet 3 + 5 \bullet 4 \end{bmatrix} = \begin{bmatrix} 20 & 22 \\ 58 & 47 \end{bmatrix}$$

More formally, for the entry in \mathbf{J} in the i^{th} row and the j^{th} column, $j_{i,j}$, we have:

$$j_{i,j} = \sum_{k=1}^2 g_{i,k} h_{k,j}$$

As with addition, matrix multiplication is associative:

$$\mathbf{G}(\mathbf{HJ}) = (\mathbf{GH})\mathbf{J}$$

However, matrix multiplication is **not** commutative:

$$\mathbf{GH} \neq \mathbf{HG}$$

Question: Given the following matrices \mathbf{M} and \mathbf{N} , find the products \mathbf{MN} and \mathbf{NM} and compare:

$$\mathbf{M} = \begin{bmatrix} 3 & 5 \\ 9 & 8 \end{bmatrix}$$

$$\mathbf{N} = \begin{bmatrix} 2 & 6 \\ 1 & 5 \end{bmatrix}$$

A square matrix is one that has the same number of rows and columns. Thus, it can always be multiplied by itself. Also, because the resulting matrix has the same size we can also multiply the resulting product matrix by the original, and so on and so on. We denote this with an exponent:

$$\mathbf{MM} = \mathbf{M}^2$$

$$\mathbf{MMM} = \mathbf{M}^3$$

There is a distributive law for matrix multiplication, too. Assuming the matrices are conformable, we have:

$$\mathbf{F}(\mathbf{D} + \mathbf{E}) = \mathbf{FD} + \mathbf{FE}$$

$$(\mathbf{M} + \mathbf{N})\mathbf{P} = \mathbf{MP} + \mathbf{NP}$$

Beware! Because matrix multiplication is involved order very much matters!!

One special matrix that comes up quite often is called the *identity matrix*. The identity matrix is a diagonal matrix with 1's along the main diagonal. An identity matrix with n rows and columns is typically denoted \mathbf{I}_n , or simply \mathbf{I} if the number of rows and columns is implied.

$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

When we multiply a matrix by a conformable identity matrix, the result is the original matrix. Say we have an **rx**c matrix \mathbf{A} , then:

$$\mathbf{AI}_c = \mathbf{I}_r\mathbf{A} = \mathbf{A}$$

This leads to the definition of the matrix inverse. We denote the inverse of matrix \mathbf{A} as \mathbf{A}^{-1} . If we multiply a matrix by its inverse, we get an identity matrix:

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

Question: the following matrices, \mathbf{A} and \mathbf{A}^{-1} , are inverses of each other. Prove this by showing that the products \mathbf{AA}^{-1} and $\mathbf{A}^{-1}\mathbf{A}$ are both equal to the identity matrix.

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 9 \end{bmatrix}, \quad \mathbf{A}^{-1} = \begin{bmatrix} 9 & -4 \\ -2 & 1 \end{bmatrix}$$

Not every matrix has inverse. Take the following for example:

$$\mathbf{U} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

There is no matrix we can multiply \mathbf{U} by to get an identity matrix. It's easy to see that here. It's not always so apparent, so just be aware of this fact!

The Zero Matrix

In addition to the identity matrix, we will often find it convenient to define a zero matrix, where all of the elements of the matrix are zero. We will note this as a bold $\mathbf{0}$.

The zero matrix has properties similar to its scalar equivalent:

$$\mathbf{0A} = \mathbf{A0} = \mathbf{0}$$

Also, if we add the zero matrix to any other matrix we get back the original matrix.

$$\mathbf{0} + \mathbf{A} = \mathbf{A} + \mathbf{0} = \mathbf{A}$$

The Transpose

The transpose of a matrix can be formed by swapping the columns and rows of the original matrix. For matrix \mathbf{A} , we denote its transpose by \mathbf{A}' (pronounced "A prime" or "A transpose"). We can easily determine each element of the transpose matrix by reversing the row index and column index of each element of the original matrix:

$$a'_{ij} = a_{ji}$$

The following are examples of matrices and their transposes:

$$\mathbf{A} = \begin{bmatrix} 3 & 5 \\ -9 & 3 \\ 10 & 8 \end{bmatrix}, \quad \mathbf{A}' = \begin{bmatrix} 3 & -9 & 10 \\ 5 & 3 & 8 \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} 3 & 5 \\ 9 & 8 \end{bmatrix}, \quad \mathbf{M}' = \begin{bmatrix} 3 & 9 \\ 5 & 8 \end{bmatrix}$$

Note that for a square matrix, taking the transpose can be thought of as rotating elements around the main diagonal.

The transpose of a transpose is the original matrix:

$$(\mathbf{A}')'\mathbf{A}$$

A square matrix that is equal to its transpose is called a *symmetric matrix*. The following are both symmetrical:

$$\mathbf{S}_1 = \begin{bmatrix} 6 & -7 \\ -7 & 6 \end{bmatrix}$$

$$\mathbf{S}_2 = \begin{bmatrix} 9 & 5 & 8 \\ 5 & 3 & -2 \\ 8 & -2 & 14 \end{bmatrix}$$