Appendix A

Introduction to Matrix Algebra I

Today we will begin the course with a discussion of matrix algebra. Why are we studying this?

- We will use matrix algebra to derive the linear regression model – the main topic of POL 603.
- Matrices are an intuitive way to think about data. We have a set of observations (perhaps individuals) on the row, and observe many different characteristics (such as race, gender, PID, etc.).
- Matrices are useful for solving systems of equations, including ones that we will see in class.

A.1 Definition of Matrices and Vectors

A matrix is simply an arrangement of numbers in rectangular form. Generally, a $(j \times k)$ matrix $\mathbf{A}$ can be written as follows:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jk} \end{bmatrix}$$

Note that there are j rows and k columns. Also note that the elements are double subscripted, with the row number first, and the column number second. When reading the text, and doing your assignments, you should always keep the dimensionality of matrices in mind. The dimensionality of the matrix is also called the *order* of a matrix. In general terms, the $\mathbf{A}$ above is of order (j, k). Let us look at a couple of examples:

$$\mathbf{W} = \begin{bmatrix} 1 & 3 \\ 2 & -6 \end{bmatrix}$$

is of order $(2,2)$. This is also called a square matrix, because the row dimension equals the column dimension ($j = k$). There are also rectangular matrices ($j \neq k$), such as:

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & 4 \\ 1 & 3 \\ 1 & -2 \\ 0 & 3 \end{bmatrix}$$

which is of order $(4,2)$. In the text, and on the board, we will denote matrices as capital, bold-faced Roman or Greek letters. Roman is typically data, and Greek is typically parameters. This is *not* a universal convention, so be aware. Of course, I cannot do boldface on the board. Thus, if you forget the dimensionality (or question whether I am talking about a matrix or a vector), stop me and ask.

There exists a special kind of matrix called a vector. Vectors are matrices that have either one row or one column. Of course, a matrix with one row *and* one column is the same as a scalar – a regular number.

Row vectors are those that have a single row and multiple columns. For example, an order $(1,k)$ row vector looks like this:

$$\alpha = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_k \end{bmatrix}$$

Similarly, column vectors are those that have a single column and multiple rows. For example, an order $(k,1)$ column vector looks like this:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}$$

Again, the convention for vectors is just like that for matrices, except the letters are lower-case. Thus, vectors are represented as lower-case, bold-faced Roman or Greek letters. Again, Roman letters typically represent data, and Greek letters represent parameters. Here, the elements are typically given a single subscript.

A.2 Matrix Addition and Subtraction

Now that we have defined matrices, vectors, and scalars, we can start to consider the operations we can perform on them. Given a matrix of numbers, one can extend regular scalar algebra in a straight forward way. Scalar addition is simply:

$$m + n = 2 + 5 = 7$$

Addition is similarly defined for matrices. If matrices or vectors are of the same order, then they can be added. One performs the addition element by element.

Thus, for a pair of order $(2,2)$ matrices, addition proceeds as follows for the problem $\mathbf{A}+\mathbf{B} = \mathbf{C}$:

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

Subtraction similarly follows. It is important to keep in mind that matrix addition and subtraction is only defined in the matrices are the same order, or, in other words, share the same dimensionality. If they do, they are said to be *conformable* for addition. If not, they are *nonconformable*.

Here is another example:

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

There are two important properties of matrix addition that are worth noting:

- $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$. In other words, matrix addition is commutative.
- $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$. In other words, matrix addition is associative.

Another operation that is often useful is transposition. In this operation, the order subscripts are exchanged for each element of the matrix $\mathbf{A}$. Thus, an order (j,k) matrix becomes an order (k,j) matrix. Transposition is denoted by placing a prime after a matrix or by placing a superscript T. Here is an example:

$$\mathbf{Q} = \begin{bmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \\ q_{3,1} & q_{3,2} \end{bmatrix} \quad \mathbf{Q}' = \begin{bmatrix} q_{1,1} & q_{2,1} & q_{3,1} \\ q_{1,2} & q_{2,2} & q_{3,2} \end{bmatrix}$$

Note that the subscripts in the transpose remain the same, they are just exchanged. Transposition makes more sense when using numbers. Here is an example for a row vector:

$$\omega = \begin{bmatrix} 1 & 3 & 2 & -5 \end{bmatrix} \quad \omega' = \begin{bmatrix} 1 \\ 3 \\ 2 \\ -5 \end{bmatrix}$$

Note that transposing a row vector turns it into a column vector, and vice versa.

There are a couple of results regarding transposition that are important to remember:

- An order (j,j) matrix $\mathbf{A}$ is said to be symmetric $\iff \mathbf{A} = \mathbf{A}'$. This implies, of course, that all symmetric matrices are square. Here is an example:

$$\mathbf{W} = \begin{bmatrix} 1 & .2 & -.5 \\ .2 & 1 & .4 \\ -.5 & .4 & 1 \end{bmatrix} \quad \mathbf{W}' = \begin{bmatrix} 1 & .2 & -.5 \\ .2 & 1 & .4 \\ -.5 & .4 & 1 \end{bmatrix}$$

- $(\mathbf{A}')' = \mathbf{A}$. In words, the transpose of the transpose is the original matrix.
- For a scalar k, $(k\mathbf{A})' = k\mathbf{A}'$.
- For two matrices of the same order, it can be shown that the transpose of the sum is equal to the sum of the transposes. Symbolically: $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$ Transposition is also commutative.

A.3 Matrices and Multiplication

So far we have defined addition and subtraction, as well as transposition. Now we turn our attention to multiplication. The first type of multiplication is a scalar times a matrix. In words, a scalar α times a matrix $\mathbf{A}$ equals the scalar times each element of $\mathbf{A}$. Thus,

$$\mathbf{A} = \alpha \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} = \begin{bmatrix} \alpha a_{1,1} & \alpha a_{1,2} \\ \alpha a_{2,1} & \alpha a_{2,2} \end{bmatrix}$$

So, for:

$$\mathbf{A} = \begin{bmatrix} 4 & 8 & 2 \\ 6 & 8 & 10 \end{bmatrix} \quad \frac{1}{2}\mathbf{A} = \begin{bmatrix} 2 & 4 & 1 \\ 3 & 4 & 5 \end{bmatrix}$$

Now we will discuss the process of multiplying two matrices. We apply the following definition of matrix multiplication. Given $\mathbf{A}$ of order (m, n) and $\mathbf{B}$ of order (n, r), then the product $\mathbf{AB} = \mathbf{C}$ is the order (m, r) matrix whose entries are defined by:

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

where $i = 1, \ldots, m$ and $j = 1, \ldots, r$. Note that for matrices to be multiplication conformable, the number of columns in the first matrix n must equal the number of rows in the second matrix n.

It is easier to see this by looking at a few examples. Let

$$\mathbf{A} = \begin{bmatrix} -2 & 1 & 3 \\ 4 & 1 & 6 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{bmatrix}$$

We can now define their product. Here we would say that $\mathbf{B}$ is pre-multiplied by $\mathbf{A}$, or that $\mathbf{A}$ is post-multiplied by $\mathbf{B}$:

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} -2 \cdot 3 + 1 \cdot 2 + 3 \cdot 1 & -2 \cdot (-2) + 1 \cdot 4 + 3 \cdot (-3) \\ 4 \cdot 3 + 1 \cdot 2 + 6 \cdot 1 & 4 \cdot (-2) + 1 \cdot 4 + 6 \cdot (-3) \end{bmatrix} \\ &= \begin{bmatrix} -1 & -1 \\ 20 & -22 \end{bmatrix} \end{aligned}$$

Note that $\mathbf{A}$ is of order $(2, 3)$, and $\mathbf{B}$ is of order $(3, 2)$. Thus, the product $\mathbf{AB}$ is of order $(2, 2)$.

We can similary compute the product $\mathbf{BA}$ which will be of order $(3, 3)$. You can verify that this product is:

$$\mathbf{BA} = \begin{bmatrix} -14 & 1 & -3 \\ 12 & 6 & 30 \\ -14 & -2 & -15 \end{bmatrix}$$

This shows that the multiplication of matrices is not commutative. In other words: $\mathbf{AB} \neq \mathbf{BA}$.

Given this notation, there are a couple of identities worth noting:

- We have alread shown the matrix multiplication is not commutative: $\mathbf{AB} \neq \mathbf{BA}$.
- Matrix multiplication is associative. In other words:

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$

- Matrix multiplication is distributive. In other words:

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

- Scalar multiplication commutative, associative, and distributive.
- The transpose of a product takes an intersting form, that can easily be proven:

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

Just as with scalar algebra, we use the exponentiation operator to denote repeated multiplication. For a square matrix (Why? $\rightarrow$ *Because it is the only type of matrix conformable with itself.*), we use the notation

$$\mathbf{A^4} = \mathbf{A} \cdot \mathbf{A} \cdot \mathbf{A} \cdot \mathbf{A}$$

to denote exponentiation.

A.4 Vectors and Multiplication

We of course use the same formula for vector multiplication as we do for matrix multiplication. There are a couple of examples that are worth looking let. Let us define the column vector $\mathbf{e}$. By definition, the order of $\mathbf{e}$ is $(N, 1)$. We can take the *inner product* of $\mathbf{e}$, which is simply:

$$\mathbf{e}'\mathbf{e} = \begin{bmatrix} e_1 & e_2 & \cdots & e_N \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}$$

This can be simplified:

$$\mathbf{e}'\mathbf{e} = e_1e_1 + e_2e_2 + \cdots + e_Ne_N = \sum_{i=1}^{N} e_i^2$$

Question → The inner product of a column vector with itself is simply equal to the sum of the square values of the vector, which is used quite often in the regression model. Geometrically, the square root of the inner product is the length of the vector. One can similarly define the outer product for column vector $\mathbf{e}$, denoted $\mathbf{ee}'$, which yields an order (N, N) matrix.

There are couple of other vector products that are interesting to note. Let $\mathbf{i}$ denote an order $(N, 1)$ vector of ones, and $\mathbf{x}$ denote an order $(N, 1)$ vector of data. The following is an interesting quantity:

$$\frac{1}{N}\mathbf{i}'\mathbf{x} = \frac{1}{N}(x_1 + x_2 + \cdots x_N) = \frac{1}{N}\sum x_i = \bar{x}$$

From this, it follows that:

$$\mathbf{i}'\mathbf{x} = \sum x_i$$

Similarly, let $\mathbf{y}$ denote another $(N, 1)$ vector of data. The following is also interesting:

$$\mathbf{x}'\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_Ny_N = \sum_{i=1}^{n} x_iy_i$$

Note that the following have nothing to do with mean deviation form (we actually will not be using mean deviation form much more in this course – it is most useful in scalar from). The lower case letters represent elements of a vector.

One can employ matrix multiplication using vectors. Let $\mathbf{A}$ be an order $(2, 3)$ matrix that looks like this:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Define the row vector $\mathbf{a_1'} = [a_{11}\ a_{12}\ a_{13}]$ and the row vector $\mathbf{a_2'} = [a_{21}\ a_{22}\ a_{23}]$. We can now express $\mathbf{A}$ as follows:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a_1'} \\ \mathbf{a_2'} \end{bmatrix}$$

It is most common to represent all vectors as column vectors, so to write a row vector you use the transposition operator. This is a good habit to get into, and one I will try to use throughout the term. We are similarly given a matrix $\mathbf{B}$, which is of order $(3, 2)$. It can be written:

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} \mathbf{b_1} & \mathbf{b_2} \end{bmatrix}$$

where $\mathbf{b_1}$ and $\mathbf{b_2}$ represent the columns of $\mathbf{B}$. The product of the matrices, then, can be expressed in terms of four inner products.

$$\mathbf{AB} = \mathbf{C} = \begin{bmatrix} \mathbf{a_1'b_1} & \mathbf{a_1'b_2} \\ \mathbf{a_2'b_1} & \mathbf{a_2'b_2} \end{bmatrix}$$

This is the same as the summation definition of multiplication. This too will come in useful later in the semester.

A.5 Special Matrices and Their Properties

When performing scalar algebra, we know that $x \cdot 1 = x$, which is known as the identity relationship. There is a similar relationship in matrix algebra: $\mathbf{AI} = \mathbf{A}$. What is $\mathbf{I}$? It can be shown that $\mathbf{I}$ is a diagonal, square matrix with ones on the main diagonal, and zeros on the off diagonal. For example, the order three identity matrix is:

$$\mathbf{I_3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Notice that $\mathbf{I}$ is oftentimes subscripted to denote its dimensionality. Here is an example of the use of an identity matrix:

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

One of the nice properties of the identity matrix is that it is commutative, associative, and distributive with respect to multiplication. That is,

$$\mathbf{AIB} = \mathbf{IAB} = \mathbf{ABI} = \mathbf{AB}$$

We similary are presented with the identity in scalar algebra that $x + 0 = x$. This generalizes to matrix algebra, with the definition of the null matrix, which is simply a matrix of zeros, denoted $\mathbf{0_{j,k}}$. Here is an example:

$$\mathbf{A} + \mathbf{0_{2,2}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \mathbf{A}$$

There are two other matrices worth mentioning. The first is a diagonal matrix, which takes values on the main diagonal, and zeros on the off diagonal. Formally, matrix $\mathbf{A}$ is diagonal if $a_{i,j} = 0 \ \forall i \neq j$. The identity matrix is an example, so is the matrix $\mathbf{\Omega}$:

$$\mathbf{\Omega} = \begin{bmatrix} 2.5667 & 0 & 0 \\ 0 & 3.4126 & 0 \\ 0 & 0 & 7.4329 \end{bmatrix}$$

The two other types of special matrices that we will be used often are square and symmetric matrices, which I defined earlier.

Appendix B

Introduction to Matrix Algebra II

B.1 Computing Determinants

So far we have defined addition, subtraction, and multiplication, along with a few related operators. We have not, however, yet defined division. Remember this simple algebraic problem:

$$\begin{aligned} 2x &= 6 \\ \frac{1}{2}2x &= \frac{1}{2}6 \\ x &= 3 \end{aligned}$$

The quantity $\frac{1}{2} = 2^{-1}$ can be called the inverse of 2. This is exactly what we are doing when we divide in scalar algebra.

Now, let us define a matrix which the inverse of $\mathbf{A}$. Let us call that matrix $\mathbf{A^{-1}}$. In scalar algebra, a number times its inverse equals one. In matrix algebra, then, we must find the matrix $\mathbf{A^{-1}}$ where $\mathbf{AA^{-1}} = \mathbf{A^{-1}A} = \mathbf{I}$.

Given this matrix, we can then do the following, which is what one needs to do when solving systems of equations:

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} \\ \mathbf{A^{-1}Ax} &= \mathbf{A^{-1}b} \\ \mathbf{x} &= \mathbf{A^{-1}b} \end{aligned}$$

because $\mathbf{A^{-1}A} = \mathbf{I}$. The question remains as to how to compute $\mathbf{A^{-1}}$.

Today we will solve this problem, first by computing determinants of square matrices using the cofactor expansion. Then, we will use these determinants to form the matrix $\mathbf{A^{-1}}$.

Determinants are defined only for square matrices, and are scalars. They are denoted $\det(\mathbf{A}) = |\mathbf{A}|$. Determinants take a very important role in determining whether a matrix is invertable, and what the inverse is.

For an order $(2,2)$ matrix, the determinant is defined as follows:

$$\det(\mathbf{A}) = |\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Here are two examples:

$$\mathbf{G} = \begin{bmatrix} 2 & 4 \\ 6 & 3 \end{bmatrix} \qquad |\mathbf{G}| = 2 \cdot 3 - 4 \cdot 6 = -18$$

$$\boldsymbol{\Gamma} = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} \qquad |\boldsymbol{\Gamma}| = 2 \cdot 2 - 4 \cdot 1 = 0$$

How do we go about computing determinants for large matrices? To do so, we need to define a few other quantities. For an order (n,n) square matrix $\mathbf{A}$, we can define the cofactor $\theta_{r,s}$ for each element of $\mathbf{A}$: $a_{r,s}$. The cofactor of $a_{r,s}$ is denoted:

$$\theta_{r,s} = (-1)^{(r+s)}|\mathbf{A_{r,s}}|$$

where $\mathbf{A_{r,s}}$ is the matrix formed after deleting row r and column s of the matrix (sometimes called the minor of $\mathbf{A}$). Thus, each element of the matrix $\mathbf{A}$ has its own cofactor. We therefore can compute the matrix of cofactors for a matrix.

Here is an example:

$$\mathbf{B} = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 5 & 6 \\ 8 & 7 & 9 \end{bmatrix}$$

We can compute the matrix of cofactors:

$$\boldsymbol{\Theta} = \begin{bmatrix} \begin{vmatrix} 5 & 6 \\ 7 & 9 \end{vmatrix} & -\begin{vmatrix} 4 & 6 \\ 8 & 9 \end{vmatrix} & \begin{vmatrix} 4 & 5 \\ 8 & 7 \end{vmatrix} \\ -\begin{vmatrix} 3 & 2 \\ 7 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 8 & 9 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 8 & 7 \end{vmatrix} \\ \begin{vmatrix} 3 & 2 \\ 5 & 6 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 4 & 6 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 4 & 5 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 3 & 12 & -12 \\ -13 & -7 & 17 \\ 8 & 2 & -7 \end{bmatrix}$$

We can use any row or column of the matrix of cofactors $\boldsymbol{\Theta}$ to compute the determinant of a matrix $\mathbf{A}$. For any row i,

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{ij}\theta_{ij}$$

Or, for any column j,

$$|\mathbf{A}| = \sum_{i=1}^{n} a_{ij}\theta_{ij}$$

These are handy formulas, because they allow the determinant of an order n matrix to be decreased to order $(n-1)$. Note that one can use any row or column to do the expansion, and compute the determinant. This process is called *cofactor expansion*. It can be repeated on very large matrices many times to get down to an order 2 matrix.

Let us return to our example, we will do the cofactor expansion on the first row, and the second column.

$$\begin{aligned}
|\mathbf{B}| &= \sum_{j=1}^{n} b_{1j}\theta_{1j} \\
&= 1 \cdot 3 + 3 \cdot 12 + 2 \cdot (-12) = 15 \\
|\mathbf{B}| &= \sum_{i=1}^{n} b_{i2}\theta_{i2} \\
&= 3 \cdot 12 + 5 \cdot (-7) + 7 \cdot 2 = 15
\end{aligned}$$

You can do this for the other two rows, or the other two columns, and get the same result.

Now, for any given matrix, you can perform the cofactor expansion and compute the determinant. Of course, as the order n increases, this gets terribly difficult, because to compute each element of the cofactor matrix, you have to do another expansion. Computers, however, do this quite easily.

B.2 Matrix Inversion

Now, let us define a matrix which is the inverse of $\mathbf{A}$. Let us call that matrix $\mathbf{A^{-1}}$. In scalar algebra, a number times its inverse equals one. In matrix algebra, then, we must find the matrix $\mathbf{A^{-1}}$ where $\mathbf{AA^{-1}} = \mathbf{A^{-1}A} = \mathbf{I}$.

Last time we defined two important quantities that one can use to compute inverses: the determinant and the matrix of cofactors. The determinant of a matrix $\mathbf{A}$ is denoted $|\mathbf{A}|$, and the matrix of cofactors we denoted $\mathbf{\Theta_A}$. There is one more quantity that we need to define, the adjoint.

The adjoint of a matrix $\mathbf{A}$ is denoted $\text{adj}(\mathbf{A})$. The adjoint is simply the matrix of cofactors transposed. Thus,

$$\text{adj}(\mathbf{A}) = \mathbf{\Theta'} = [\theta_{r,s}]' = \left[(-1)^{(r+s)}|\mathbf{A_{r,s}}|\right]'$$

where $\mathbf{A_{r,s}}$ is the matrix formed after deleting row r and column s of the matrix (sometimes called the minor of $\mathbf{A}$).

We now know all we need to know to compute inverses. It can be shown that the inverse of a matrix $\mathbf{A}$ is defined as follows:

$$\mathbf{A^{-1}} = \frac{1}{|\mathbf{A}|}\text{adj}(\mathbf{A})$$

Thus, for any matrix $\mathbf{A}$ that is invertable, we can compute the inverse. This is trival for order $(2,2)$ matrices, and only takes a few minutes for order $(3,3)$ matrices. It gets much more difficult after that, and we can use computers to compute inverses, both numerically (using Stata, Gauss, or some other package), or analytically (using Mathematica, Maple, etc.).

B.3 Examples of Inversion

We will do two examples. First, we will find the inverse of an order $(2,2)$ matrix:

$$\mathbf{B} = \begin{bmatrix} 2 & 4 \\ 6 & 3 \end{bmatrix}$$

We first must calculate the determinant of $\mathbf{B}$:

$$|\mathbf{B}| = 2 \cdot 3 - 6 \cdot 4 = -18$$

We can write down the matrix of cofactors for $\mathbf{B}$, which we then transpose to get the adjoint:

$$\text{adj}(\mathbf{B}) = \mathbf{\Theta_B}' = \begin{bmatrix} 3 & -6 \\ -4 & 2 \end{bmatrix}' = \begin{bmatrix} 3 & -4 \\ -6 & 2 \end{bmatrix}$$

Given the determinant and the adjoint, we can now write down the inverse of $\mathbf{B}$:

$$\mathbf{B^{-1}} = -\frac{1}{18}\begin{bmatrix} 3 & -4 \\ -6 & 2 \end{bmatrix} = \begin{bmatrix} \frac{-3}{18} & \frac{4}{18} \\ \frac{6}{18} & \frac{-2}{18} \end{bmatrix}$$

To make sure, let us check the product $\mathbf{B^{-1}B}$ to make sure it equals the identity matrix $\mathbf{I_2}$:

$$\mathbf{B^{-1}B} = \begin{bmatrix} \frac{-3}{18} & \frac{4}{18} \\ \frac{6}{18} & \frac{-2}{18} \end{bmatrix}\begin{bmatrix} 2 & 4 \\ 6 & 3 \end{bmatrix} = \begin{bmatrix} \frac{18}{18} & \frac{0}{18} \\ \frac{0}{18} & \frac{18}{18} \end{bmatrix} = \mathbf{I_2} \blacksquare$$

Here is another example. If you remember earlier, we were working on an order $(3,3)$ matrix also called $\mathbf{B}$. The matrix was as follows:

$$\mathbf{B} = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 5 & 6 \\ 8 & 7 & 9 \end{bmatrix}$$

We took the time to compute the matrix of cofactors for this matrix:

$$\mathbf{\Theta_B} = \begin{bmatrix} 3 & 12 & -12 \\ -13 & -7 & 17 \\ 8 & 2 & -7 \end{bmatrix}$$

We also showed that its determinant $|\mathbf{B}| = 15$. Given this information, we can write down the inverse of $\mathbf{B}$:

$$\mathbf{B^{-1}} = \frac{1}{15}\mathbf{\Theta_B}' = \frac{1}{15}\begin{bmatrix} 3 & -13 & 8 \\ 12 & -7 & 2 \\ -12 & 17 & 7 \end{bmatrix}$$

Again, let us check to make sure this is indeed the inverse of **B**.

$$\mathbf{B^{-1}B} = \frac{1}{|\mathbf{B}|}\mathrm{adj}(\mathbf{B})\mathbf{B} = \frac{1}{15}\begin{bmatrix} 3 & -13 & 8 \\ 12 & -7 & 2 \\ -12 & 17 & 7 \end{bmatrix}\begin{bmatrix} 1 & 3 & 2 \\ 4 & 5 & 6 \\ 8 & 7 & 9 \end{bmatrix}$$

$$= \frac{1}{15}\begin{bmatrix} 15 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 15 \end{bmatrix} = \mathbf{I_3} \ \blacksquare$$

B.4 Diagonal Matrices

There is one type of matrix for which the computation of the inverse is nearly trivial – diagonal matrices. Let **A** denote a diagonal matrix of order (k, k). Remember that diagonal matrices have to be square:

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{kk} \end{bmatrix}$$

It can be shown that the inverse of **A** is the following:

$$\mathbf{A}^{-1} = \begin{bmatrix} a_{11}^{-1} & 0 & 0 & \cdots & 0 \\ 0 & a_{22}^{-1} & 0 & \cdots & 0 \\ 0 & 0 & a_{33}^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{kk}^{-1} \end{bmatrix}$$

To show this is the case, simply multiply $\mathbf{AA^{-1}}$ and you will get $\mathbf{I_k}$.

B.5 Conditions for Singularity

Earlier I talked about the simple alebraic problem of solving for x, and show that one need to premultiply by x^{-1} to solve a system of equations. This is the same operation as division. In scalar algebra, there is one number for which the inverse is not defined: 0. The quantity $\frac{1}{0}$ is not defined.

Similarly, in matrix algebra there are a set of matrices for which an inverse does not exist. Remember our formula for the inverse of our matrix **A**:

$$\mathbf{A^{-1}} = \frac{1}{|\mathbf{A}|}\mathrm{adj}(\mathbf{A})$$

Question → Under what conditions will this not be defined? *When the determinant of* $\mathbf{A}$ *is not defined, of course.* This tells us, then, that if the determinant of $\mathbf{A}$ equals zero, then $\mathbf{A^{-1}}$ is not defined.

For what sorts of matrices is this a problem? It can be shown that matrices that have rows or columns that are *linearly dependent* on other rows or columns have determinants that are equal to zero. For these matrices, the determinant is undefined. We are given an order (k, k) matrix $\mathbf{A}$, that I will denote using column vectors:

$$\mathbf{A} = \left[\begin{array}{cccc} \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_k} \end{array} \right]$$

Each of the vectors $\mathbf{a_i}$ is of order $(k, 1)$. A column $\mathbf{a_i}$ of $\mathbf{A}$ is said to be *linearly independent* of the others if there exists no set of scalars α_j such that:

$$\mathbf{a_i} = \sum_{j \neq i} \alpha_j \mathbf{a_j}$$

Thus, given the rest of the columns, if we cannot find a weighted sum to get the column we are interested in, we say the it is linearly independent.

We can define the term *rank*. The *rank* of a matrix is defined as the number of linearly independent columns (or rows) of a matrix. If all of the columns are independent, we say that the matrix is of *full rank*. We denote the rank of a matrix as $r(\mathbf{A})$. By definition, $r(\mathbf{A})$ is a integer that can take values from 1 to k. This is something that can be computed by software packages, such as Mathematica, Maple, or Stata.

Here are some examples:

$$\left| \begin{array}{cc} 2 & -4 \\ 3 & -6 \end{array} \right| = 2(-6) - 3(-4) = -12 + 12 = 0$$

Notice that the second column is -2 times the first column. The rank of this matrix is 1 – it is not of full rank. Here is another example:

$$\left| \begin{array}{cc} 2 & -7 \\ 4 & -14 \end{array} \right| = 2(-14) - 4(-7) = -28 + 28 = 0$$

Notice that the second row is 2 times the first row. Again, the rank of this matrix is 1 – it is not of full rank. Here is a final example, for an order $(3, 3)$ matrix:

$$\left| \begin{array}{ccc} 1 & 2 & 4 \\ 3 & 0 & 6 \\ 5 & 3 & 13 \end{array} \right| = 0 \text{ using Mathematica.}$$

Notice that the first column times 2 plus the third column equals the third column. The rank of this matrix is 2 – it is not of full rank.

Here are a series of important statements about inverses and nomenclature:

- $\mathbf{A}$ must be square. It is a necessary, but not sufficient, condition that $\mathbf{A}$ is square for $\mathbf{A^{-1}}$ to exist. In other words, sometimes the inverse of a matrix does not exist. If the inverse of a matrix does not exist, we say that it is *singular*.

- The following statements are equivalent: full rank $\iff$ nonsingular $\iff$ invertable. All of these imply that $\mathbf{A^{-1}}$ exists.

- If the determinant of $\mathbf{A}$ equals zero, then $\mathbf{A}$ is said to be singular, or not invertable. More generally,
$$|\mathbf{A}| = \mathbf{0} \iff \mathbf{A} \text{ singular.}$$

- If the determinant of $\mathbf{A}$ is non-zero, then $\mathbf{A}$ is said to be nonsingular, or invertable. In other words, the inverse exists. More generally,
$$|\mathbf{A}| \neq \mathbf{0} \iff \mathbf{A} \text{ nonsingular.}$$

- If a matrix $\mathbf{A}$ is not of full rank, it is not invertable; i.e., it is singular.

B.6 Some Important Properties of Inverses

Here are some important identities that relate to matrix inversion:

- $\mathbf{AA^{-1}} = \mathbf{I} \quad \mathbf{A^{-1}A} = \mathbf{I} \quad \mathbf{A^{-1}}$ is unique.

- $\mathbf{A}$ must be square. It is a necessary, but not sufficient, condition that $\mathbf{A}$ is square for $\mathbf{A^{-1}}$ to exist. In other words, sometimes the inverse of a matrix does not exist – it is singular.

- $(\mathbf{A^{-1}})^{\mathbf{-1}} = \mathbf{A}$. In words, the inverse of an inverse is the original matrix.

- Just as with transposition, it can be shown that
$$(\mathbf{AB})^{\mathbf{-1}} = \mathbf{B^{-1}A^{-1}}$$

- One can also show that the inverse of the transpose is the transpose of the inverse. Symbolically,
$$(\mathbf{A}')^{\mathbf{-1}} = (\mathbf{A^{-1}})'$$

B.7 Solving Systems of Equations Using Matrices

Matrices are particularly useful when solving systems of equations, which, if you remeber, is what we did when we when solved for the least squares estimators. You covered this material in your high school algebra class. Here is an example, with three equations and three unknowns:

$$\begin{aligned} x + 2y + z &= 3 \\ 3x - y - 3z &= -1 \\ 2x + 3y + z &= 4 \end{aligned}$$

How would one go about solving this? There are various techniques, including substitution, and multiplying equations by constants and adding them to get single variables to cancel.

There is an easier way, however, and that is to use a matrix. Note that this system of equations can be represented as follows:

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & -1 & -3 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix} \iff \mathbf{Ax} = \mathbf{b}$$

We can solve the problem $\mathbf{Ax} = \mathbf{b}$ by pre-multiplying both sides by $\mathbf{A^{-1}}$ and simplifying. This yields the following:

$$\mathbf{Ax} = \mathbf{b} \rightarrow \mathbf{A^{-1}Ax} = \mathbf{A^{-1}b} \rightarrow \mathbf{x} = \mathbf{A^{-1}b}$$

We can therefore solve a system of equations by computing the inverse of $\mathbf{A}$, and multiplying it by $\mathbf{b}$. Here our matrix $\mathbf{A}$ and its inverse is as follows (using Mathematica to perform the calculation):

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & -1 & -3 \\ 2 & 3 & 1 \end{bmatrix} \qquad \mathbf{A^{-1}} = \begin{bmatrix} 8 & 1 & -5 \\ -9 & -1 & 6 \\ 11 & 1 & -7 \end{bmatrix}$$

We can now solve this system of equations:

$$\mathbf{x} = \mathbf{A^{-1}b} = \begin{bmatrix} 8 & 1 & -5 \\ -9 & -1 & 6 \\ 11 & 1 & -7 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 4 \end{bmatrix} \quad \blacksquare$$

If we plug these back into original equations, we see that they in fact fulfill the identities. Computationally, this is a much easier way to solve systems of equations – we just need to compute an inverse, and perform a single matrix multiplication.

This approach only works, however, if the matrix $\mathbf{A}$ is nonsingular. If it is not invertable, then this will not work. In fact, if a row or a column of the matrix $\mathbf{A}$ is a linear combination of the others, there are *no* solutions to the system of equations, or *many* solutions to the system of equations. In either case, the system is said to be under-determined. We can compute the determinant of a matrix to see if it in fact is underdetermined.