G31.1101-04: Applied Statistics and Econometrics I (Lab)

Matthias Kredler

September 9th, 2004

Basics in Matrix Algebra

1 Definition of a Matrix

A matrix is a collection of numbers in a table-like structure. Each of the elements has a row and a column index. A matrix that has 2 rows and 3 columns –for example– is said to be "2 by 3" or "2×3". A typical matrix would look like this:

To refer to the element in the i-th row and the j-th column of a matrix $\mathbf{A}$, one writes a_{ij}. For example, $a_{12} = 0$ and $a_{22} = 5$.

Matrices that have only one column are called *column vectors* – they are N by 1. 1-by-N matrices are called *row vectors*; they consist only of one row. For column vectors, it is common to omit the index for the column (which would always be one if we bothered to write it), for example:

2 Where Are Matrices Useful?

- In areas outside Economics: Solving systems of linear equations, computer graphics, genetics, cryptography, electrical networks and many more.
- Good programs to work with matrices: *Matlab* and *Gauss* (are installed on all workstations in the computer lab!)

3 Matrix Addition

Two matrices of the same dimension can be added. Addition goes element by element. The resulting matrix is of the same dimension as the added matrices:

Subtraction of two matrices is defined equivalently, i.e. it is also carried out element-wise. Since all operations are element by element, **commutativity** and **associativity** carry over from normal calculus:

Here, **A**, **B** and **C** are matrices of the same dimension.

4 Scalar Multiplication

In matrix algebra, we refer to "normal numbers" as *scalars*. If we multiply every component of a matrix by a scalar, we perform an operation called *scalar multiplication*. An example:

In this example, λ is a scalar and **A** is the matrix defined before. Again, since operations are performed element-wise, we can use a law we already know from normal calculus: **distributivity**.

5 Matrix Multiplication

Another form of multiplication widely used in matrix algebra is *matrix multiplication*. As the term suggests, it is used for multiplying two matrices with each other. A matrix **B** can only be multiplied by a matrix **C** if **B** has just as many columns as **C** has rows. If this is not the case, the product $\mathbf{B} * \mathbf{C}$ is not defined. In general, if **B** is n by m and **C** is m by p, then the resulting matrix $\mathbf{B} * \mathbf{C}$ (or also written simply as $\mathbf{BC}$) will be a n-by-p matrix.

The rules for multiplying two matrices are most easily understood by looking at some examples:

...or if you prefer to work with numbers:

In general, the element in the i-th row and the j-th column of $\mathbf{D} = \mathbf{B} * \mathbf{C}$ is calculated by taking the i-th row of $\mathbf{B}$ and the j-th column of $\mathbf{C}$ and summing over the products of the single terms (multiplying the first entry in the i-th row of $\mathbf{B}$ with the first entry of the j-th column of $\mathbf{C}$, the second entry with the second etc.). Mathematically, this amounts to

where m is the number of columns in $\mathbf{B}$ (and the number of rows in $\mathbf{C}$).

Matrix multiplication is often used when representing systems of linear equations. Let us define the matrix $\mathbf{G}$ and the vector $\mathbf{h}$ by

and see what happens if we create the following equation using the vector $\mathbf{x}$ from before:

If we consider x_1 and x_2 as unknowns, this is equivalent to the following system of two equations with two unknowns:

If you look at the first line in the matrix representation of the system ($\mathbf{Gx} = \mathbf{h}$), you can see why we might be interested in matrix algebra when working with systems of linear equations; if there was some way of dividing by $\mathbf{G}$ then we could write the solution of the system very conveniently. Later in this course, when we talk about inverting matrices, we will see what this kind of "division" means in the context of matrix algebra.

As for rules of calculus, **associativity** and **distributivity** hold again for matrix multiplication. Convince yourself with some examples that the following rules hold:

It is *very important* to notice that **commutativity does not hold for matrix multiplication**. First, notice that the product $\mathbf{B} * \mathbf{A}$ may not be defined while $\mathbf{A} * \mathbf{B}$ makes perfect sense. This is the case, for example, if $\mathbf{A}$ is 1 by 2 and $\mathbf{B}$ is 2 by 3.

Even if the product is defined both ways, commutativity does not hold as a rule. To see this, first go through an example where $\mathbf{A} * \mathbf{B}$ is of a different dimension than $\mathbf{B} * \mathbf{A}$ (e.g. where $\mathbf{A}$ is 1 by 2 and $\mathbf{B}$ is 2 by 1). Then, find a numeric example where the resulting matrices $\mathbf{A} * \mathbf{B}$ and $\mathbf{B} * \mathbf{A}$ are of the same dimension but have different entries; notice that the matrices $\mathbf{A}$ and $\mathbf{B}$ have to be square (i.e. n-by-n matrices) for this to happen.

6 Matrix Transpose

The transpose operation interchanges the rows and the columns of a matrix. It is denoted by $\mathbf{A}'$ (you will also see the notation $\mathbf{A}^T$ in some books). If $\mathbf{A}$ is n by m, then $\mathbf{A}'$ m by n. An example

Go through some examples and convince yourself that the following rules hold:

If a column vector is transposed, we obtain a row vector. An interesting thing happens when we multiply this transpose by the original column vector. If we try it with our vector $\mathbf{x}$ that we dealt with before we get

In this course, we will deal a lot with structures like this one. In the general case, where there is a matrix multiplied between the two vectors ($\mathbf{x}'\mathbf{A}\mathbf{x}$), these structures are called *quadratic* *forms*. They play a very important role in matrix algebra. For our purposes, it will be very important to deal with the case where a sum of two vectors is plugged into a structure like this. If we apply the laws for transposes and the rule of distributivity for matrix multiplication (twice!), we get the following result that should remind you of a basic rule in normal calculus:

The last step follows from the fact that the transpose of a scalar is the scalar itself again, and hence

7 The Identity Matrix

The identity matrix of dimension n is an n-by-n matrix whose elements are 1 along the diagonal and 0 otherwise. It is denoted by $\mathbf{I}_n$. Example:

If it is clear from the context, the subscript n is omitted. Convince yourself with an example of the following rules for a n-by-m matrix $\mathbf{A}$: