MATH 320, WEEK 7: Matrices, Matrix Operations

1 Matrices

We have introduced ourselves to the notion of the grid-like **coefficient matrix** as a short-hand coefficient place-keeper for performing **Gaussian elimination**. It turns out that **matrices** turn up in a wide range of applications not restricted to solving linear systems of equations. In order to see how they arise, however, we will have to first define general matrices, their basic properties, and the basic operations we can perform on them.

Definition 1.1. *The grid structure $A = [a_{ij}]$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, is called a **matrix** of **dimension** $m \times n$ (or simply an m-by-n matrix). Explicitly, we have*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

There are a few notes worth making:

- By convention, we will use capital letters from the early part of the alphabet to denote matrices (e.g. A, B, C, etc.).

- All of the matrices we will see in this course will have *real-valued* (or *integer-valued*) coefficients. In other words, we will have $a_{ij} \in \mathbb{R}$ (or $a_{ij} \in \mathbb{Z}$). There are, however, a significant number of applications which require *complex-valued* matrix coefficients (e.g. quantum physics!). Most of the concepts and operations we will introduce will have natural extensions to that setting.

- By convention, the first index of the dimension corresponds to the **number of rows**, and the second index corresponds to the **number of columns**. For example, a 5-by-3 matrix will a matrix with five rows and three columns, *not* three rows and five columns. We will also follow the convention that the entry a_{ij} will correspond to the element in the i^{th} row and j^{th} row. It will be important to remember this order!

- A matrix is called a **square matrix** if it has the same number of rows and columns (i.e. it has dimension $n \times n$, is an n-by-n matrix, etc.). Square matrices will have special properties we will investigate later in the week.

We will be interested in defining operations on matrices. In many ways, the operations we will define for **matrices** will be the same as those defined for **numbers**. That is to say, we are interested in such things are addition (e.g. $A+B$), subtraction (e.g. $A-B$), multiplication (e.g. $A \cdot B$), exponentiation (e.g. A^2), etc., of matrices. This is not a coincidence, but there are subtleties of which we will have to be aware.

The most basic operation we can perform is to take the *transpose*, which is defined in the following way.

Definition 1.2. *Let A be an m-by-n matrix with entries $A = [a_{ij}]$. Then the **transpose** of A is denoted A^T and has entries $A^T = [a_{ji}]$.*

In other words, we simply switch the indices of the entries. An entry in the 5^{th} row and 1^{st} column of A will be in the 1^{st} row and 5^{th} column of A^T. For example, for the matrix

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

has the transpose

$$A^T = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 5 & -1 \end{bmatrix}.$$

It should be obvious from both the definition and the example that, if A is an m-by-n matrix, then A^T will be a n-by-m matrix. Nevertheless, it is *very important* to remember this fact! When performing matrix operations, it is important to make sure we are considering matrices of the right dimension.

2 Matrix Addition

The first and most basic algebraic operation we can be define for matrices is *matrix addition.*

Definition 2.1. *Let A and B be two m-by-n matrices with entries $A = [a_{ij}]$, $B = [b_{ij}]$. Then the matrix $A + B$ is defined to be the m-by-n matrix with entries $A + B = [a_{ij} + b_{ij}]$.*

Example: Find the matrix $A + B$ for

$$A = \begin{bmatrix} 2 & -3 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 7 & -1 \end{bmatrix}.$$

Solution: It is just a matter of performing element-wise addition. We have

$$\begin{aligned} A + B &= \begin{bmatrix} 2+1 & (-3)+1 & (0)+1 \\ 1+(0) & (0)+7 & (-1)+(-1) \end{bmatrix} \\ &= \begin{bmatrix} 3 & -2 & 1 \\ 1 & 7 & -2 \end{bmatrix}. \end{aligned}$$

The only really important note to make about matrix addition is that the matrices being added *must* have *exactly* the same dimensions. We are not allowed, for instance, to add a 2-by-6 matrix to a 3-by-5 matrix, or even a 2-by-6 matrix to a 2-by-5 matrix. Passing this test, however, the procedure is exactly as easy as component-wise addition.

3 Matrix Multiplication

To motivate the discussion of matrix multiplication, let's consider first of all multiplying a matrix by a *scalar* value. That is to say, let's consider the case of multiply a matrix by a number, i.e. cA. We have the following definition.

Definition 3.1. *Let A be an m-by-n matrix and $c \in \mathbb{R}$ be a real number. Then the matrix cA is define to be the matrix with entries $cA = [ca_{ij}]$.*

That it to say, we simply multiply each entry in the matrix by the constant c. For instance, if we want to compute the matrix $2A$ or $(-1)A$ for

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

we can easily evaluate

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

and

$$(-1)A = (-1)\begin{bmatrix} 2 & -3 & 0 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -2 & 3 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

Now that we have seen matrix operations which amounted to performing operations component-wise on the matrices involved, we might wonder how far this intuition extends in expanding on system of matrix algebra. In particular, we might wonder if general matrix multiplication can be defined in this manner. That is to say, if we have two matrices, A and B, can we compute their product $A \cdot B$ by simply taking the product of each element component-wise?

The answer, unfortunately, is a pronounced **NO**. The definition of matrix multiplication is more complicated than the procedure we have defined for matrix addition and scalar multiplication. But we will see as the course progresses that the definition we introduce will have a significantly higher number of applications than the naive notion of multiplying matrices component-wise.

Definition 3.2. *Let A and B be matrices with dimension m-by-p and p-by-n respectively. Then the matrix $A \cdot B$ (or simply AB) is the matrix with entries*

$$A \cdot B = \left[\sum_{k=1}^{p} a_{ik} b_{kj} \right].$$

Note: The more common intuition of matrix multiplication is that we multiply the elements of the i^{th} row of A by the elements of the j^{th} column of B, add the results, and place the final product in the i^{th} row and j^{th} column of the new matrix.

For example, consider the matrices

$$A = \begin{bmatrix} -2 & 1 & 0 & 1 \\ 1 & 1 & -2 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & 5 \\ 2 & 0 \\ 1 & -3 \\ 1 & 1 \end{bmatrix}.$$

Then the matrix $A \cdot B$ is given by

$$\begin{aligned} A \cdot B &= \begin{bmatrix} -2 & 1 & 0 & 1 \\ 1 & 1 & -2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 2 & 0 \\ 1 & -3 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} (-2)(3)+(1)(2)+(0)(1)+(1)(1) & (-2)(5)+(1)(0)+(0)(-3)+(1)(1) \\ (1)(3)+(1)(2)+(-2)(1)+(1)(-1) & (1)(5)+(1)(0)+(-2)(-3)+(1)(-1) \end{bmatrix} \\ &= \begin{bmatrix} -3 & -9 \\ 2 & 10 \end{bmatrix}. \end{aligned}$$

There are a few notes worth making about this process:

- As with matrix addition, the dimension of the matrices involved will be very important in determining whether matrix multiplication can be applied to them. Since we are multiplying *across* the columns of A and the rows of B, we will need the number of columns of A and rows of B to match!

- We might notice that the dimension of the matrix which resulted from the multiplication operation is different than either of the original matrix. We can reason basically that the matrix which results from the multiplication of an m-by-p matrix and a p-by-n is necessarily an m-by-n matrix. For instance, if we multiply a 2-by-6 matrix and a 6-by-8 matrix, the resulting matrix will always be a 2-by-8 matrix, regardless of the particular entries. The trick is easy: remove the interior indices!

- An immediate result of these observations is that *matrix multiplication is not commutative*. That is to say, in general, we have that $A \cdot B \neq B \cdot A$. This is a significant difference between matrix multiplication and the intuition guiding multiplication of numbers (where we trivially have $a \cdot b = b \cdot a$). Nevertheless, we can clearly see that, for the previous example, we have

$$\begin{aligned} B \cdot A &= \begin{bmatrix} 3 & 5 \\ 2 & 0 \\ 1 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 & 1 \\ 1 & 1 & -2 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 8 & -10 & -2 \\ -4 & 2 & 0 & 2 \\ -6 & -2 & 6 & 4 \\ -1 & 2 & -2 & 0 \end{bmatrix}. \end{aligned}$$

  This matrix does not even have the same dimensions as the matrix obtained from multiplying in the other order! (Worse still, in general, matrix multiplication may not even have been defined in the other direction.)

- A result of the previous observation is that square matrices are particularly nice for multiplying. If we multiply square matrices (of the same dimension) we obtain a matrix of the same dimension. It will be important to remember, however, that square matrices are *still* not commutative (i.e. $A \cdot B \neq B \cdot A$ in general).

4 Vectors

Matrices may seem like a foreign concept at first glance, but if we think a little bit we may realize that we have already seen matrices in other contexts. Matrices with only a single relevant dimension (i.e. an n-by-1 or 1-by-n matrix) commonly go by another name: they are called **vectors**.

Definition 4.1. *A* 1*-by-*n *matrix is called an* n*-dimensional* ***row vector*** *and an* n*-by-*1 *matrix is called an* n*-dimensional* ***column vector****.*

Notes:

- Vectors are commonly denoted with boldface (e.g. $\mathbf{v}$, $\mathbf{w}$, etc.) or with a small arrow overtop (e.g. $\vec{v}$, $\vec{w}$, etc.). The text uses $\mathbf{v}$. In lecture, I will use $\vec{v}$ (due to the difficulty in writing bold-face with chalk!).

- Since vectors only have one relevant dimension, they will be indexed with a single index (i.e. $\vec{v} = [v_i]$).

- The distinction between row and column vectors is clear from explicitly writing out the form of the matrices. For a row vector, we have

$$\vec{v} = [v_1 \; v_2 \; \cdots \; v_n]$$

  while for a column vector we have

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

We will encounter vectors in a number of contexts throughout the remainder of this course (including when we revisit differential equations) but we will not study them in depth until a few weeks from now. It is worth noting now, however, that the matrix operations we have defined for matrices are also valid for vectors. In fact, we may have seen a few of them already in earlier algebra, computer science, or physics courses.

Definition 4.2. *Given a* n*-dimensional vectors* $\vec{v} = [v_1 \;\; v_2 \;\; \cdots \;\; v_n]$ *and* $\vec{w} = [v_1 \;\; v_2 \;\; \cdots \;\; v_n]$*, and a constant* $c \in \mathbb{R}$*, the following operations are defined:*

1. ***vector addition****:* $\vec{v} + \vec{w} = [v_i + w_i]$*;*

2. ***vector scalar multiplication:*** $c\vec{v} = [cv_i]$*; and*

3. ***dot product:*** $\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} v_i w_i$.

The key observation here is that these are exactly the same operations we have previous defined for matrices! The dot product may appear different than anything we have seen so far, but it can in fact always be written as a matrix multiplication operation by taking one or the other of $\vec{v}$ or $\vec{w}$ to be its transpose (i.e. writing $\vec{v}^T\vec{w}$ or $\vec{v}\vec{w}^T$). The only difference after that is that we have dropped the matrix notation for the end result. (Since multiplying a 1-by-n matrix by an n-by-1 matrix produces a 1-by-1 matrix, it is a pretty trivial matrix!)

For example, if we take $\vec{v} = [1\ 2\ -1]^T$, $\vec{w} = [0\ -1\ 3]^T$, and $c = 5$, then we have

$$\vec{v} + \vec{w} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix},$$

$$c\vec{v} = \begin{bmatrix} 5 \\ 10 \\ -5 \end{bmatrix}, \quad c\vec{w} = \begin{bmatrix} 0 \\ -5 \\ 15 \end{bmatrix},$$

and

$$\vec{v} \cdot \vec{w} = \vec{v}^T\vec{w} = \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix} = (1)(0) + (2)(-1) + (-1)(3) = -5.$$

5 Linear Systems

One immediate application of vectors and the matrix multiplication operation we have defined comes from reconsidering the linear systems we studied last week. If we define the m-by-n matrix $A = [a_{ij}]$ and the vectors $\vec{x} = [x_1\ x_2\ \cdots\ x_n]^T$ and $\vec{b} = [b_1\ b_2\ \cdots\ b_n]^T$, we have that $A\vec{x} = \vec{b}$ gives

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

If we multiply this out, we obtain

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

This is exactly the general form of a linear system of m equations in n unknowns! It should come as no surprise then that we will prefer the shorthand notation $A\vec{x} = \vec{b}$ to the explicit form above. We will revisit this matrix formulation later in the course (although we will need to define a few further matrix operation first!).

6 Special Matrices

There are several matrices which will turn out to have particularly nice properties. The first is a matrix which is particularly well-behaved under the matrix addition operator

Definition 6.1. *The **zero matrix** of dimension m-by-n is denoted $\mathbf{0}$ and is defined as the n-by-m matrix with zeroes in every entry.*

The key feature of the zero matrix is that it does not change a matrix upon matrix addition. That is to say $A + \mathbf{0} = A = \mathbf{0} + A$ for all matrices A where matrix addition is defined.

Next, we define a square matrix which is particularly well behaved under the multiplication operation.

Definition 6.2. *The **identity matrix** of dimension n is denoted $\mathbf{I}$ and is defined as the n-by-n matrix with entries $a_{ij} = 1$ if $i = j$ and $a_{ij} = 0$ if $i \neq j$. In other words, it has ones along the diagonal and zeroes elsewhere:*

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

The identity matrix has the property that it does not change a matrix upon matrix multiplication. For instance, if we have

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix} 2 & -1 & -2 \\ 0 & 3 & 1 \end{bmatrix}$$

then

$$
\begin{aligned}
I \cdot A &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & -2 \\ 0 & 3 & 1 \end{bmatrix} \\
&= \begin{bmatrix} (1)(2)+(0)(-1) & (1)(-1)+(0)(3) & (1)(-2)+(0)(1) \\ (0)(2)+(1)(0) & (0)(-1)+(1)(3) & (0)(-2)+(1)(1) \end{bmatrix} \\
&= \begin{bmatrix} 2 & -1 & -2 \\ 0 & 3 & 1 \end{bmatrix}.
\end{aligned}
$$

There are a few notes worth making:

1. It also holds that $A \cdot \mathbf{I} = A$ for all matrices where the multiplication operation is defined.
2. The zero and identity matrices are the matrix analogues of the real values 0 and 1 in the sense that $a + 0 = a = 0 + a$ and $a \cdot 1 = a = 1 \cdot a$ for all real values a just as $A + \mathbf{0} = A = \mathbf{0} + A$ and $A \cdot \mathbf{I} = A = \mathbf{I} \cdot A$. We will just have to be careful that we make sure the matrices have the correct dimensions for the operations involved!

Another square matrix which has particularly nice properties (which is not contained in the text) are *symmetric matrices.*

Definition 6.3. *An n-by-n matrix A is said to be* ***symmetric*** *if $a_{ij} = a_{ji}$ for all $i, j = 1, \ldots, n$. That is to say, a matrix is symmetric if $A = A^T$.*

We will not investigate in depth the properties of symmetric matrices for a few weeks yet. Nevertheless, we recognize that it is easy to identify symmetric matrices. We just need to remember what the transpose is. It is the matrix where all entries are reflected across the diagonal. For symmetric matrices, that reflection produces the same matrix! For example, the matrices

$$
A = \begin{bmatrix} -1 & 2 & 7 \\ 2 & 0 & 3 \\ 7 & 3 & 5 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}
$$

are symmetric while the matrix

$$
C = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}
$$

is not.