Matrices and Matrix Algebra

Unless otherwise stated, a matrix in this section is assumed to be a true matrix as oppose to an augmented matrix.

1. Preliminaries

1.1. **Notation.** A matrix with m rows and n columns is referred to as an $\mathbf{m \times n}$ **matrix**. A generic $m \times n$ matrix A can be written as:

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

However, if we know that the row i, column j entry is given by a_{ij}, then we will often use the abbreviated notation:

$$A = (a_{ij})$$

1.2. **Equality.** Suppose that $A = (a_{ij})$ and $B = (b_{ij})$. We say that A is **equal** to B, and write $A = B$ if A and B are the same size **and** $a_{ij} = b_{ij}$ for every i and j.

1.3. **Vectors.** An $m \times 1$ matrix is called a **column vector** or simply a **vector**. A $1 \times n$ matrix is referred to as a **row vector**. The **dimension** of a column or row vector is the number of rows or columns, respectively, in the vector.

2. Matrix Algebra

In this section we develop much of the algebra of matrices. Depending on what types of numbers we are working with, the word **scalar** can refer to a rational number, a real number, a a complex number, or possibly some other type of number. As we work mostly with **real matrices**, that is, matrices whose entries are real numbers, we will usually assume that our scalars are real numbers as well.

2.1. **Scalar Multiplication.** Let $A = (a_{ij})$ be an $m \times n$ matrix and let α be a scalar. The **scalar product** of α and A is defined to be the matrix:

$$\alpha A = (\alpha a_{ij})$$

That is, every entry of A is multiplied by the scalar α.

Example:

$$3 \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 3a & 3b & 3c \\ 3d & 3e & 3f \\ 3g & 3h & 3i \end{pmatrix}$$

2.2. **Matrix Addition.** Let $A = (a_{ij})$ and $B = (b_{ij})$ both be matrices of size $m \times n$. Then the sum of A and B is the $m \times n$ matrix given by:

$$A + B = (a_{ij} + b_{ij})$$

That is, the entries of A and B are added pointwise.

For each m and n there is an unique matrix $\mathbf{0}$ called the **$\mathbf{m} \times \mathbf{n}$ zero matrix**. It is defined by $\mathbf{0} = (z_{ij})$ where $z_{ij} = 0$ for all i and j. It is the unique matrix such that

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

for every $m \times n$ matrix A.

Example:

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} a+1 & b+2 & c+3 \\ d+4 & e+5 & f+6 \\ g+7 & h+8 & i+9 \end{pmatrix}$$

2.3. **Matrix Multiplication.** Let $A = (a_{ij})$ be a matrix of size $m_1 \times n$ matrix and $B = (b_{ij})$ be a matrix of size $n \times m_2$. Then the product of A and B is the $m_1 \times m_2$ matrix given by:

$$AB = (c_{ij}) \text{ where } c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

It is important to note that matrix multiplication is not **commutative**. That is, AB is generally not equal to BA. As we will see late, matrix multiplication can be thought of as function composition.

For each m there is an unique matrix I_m called the **$\mathbf{m} \times \mathbf{m}$ identity matrix**. It is defined by $I_m = (\delta_{ij})$ where $\delta_{ij} = 1$ when $i = j$ and $\delta_{ij} = 0$ when $i \neq j$. It is the unique matrix such that

$$I_m A = A \text{ and } B I_m = B$$

for every $m \times n$ matrix A and $n \times m$ matrix B.

Example:

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1a + 3b + 5c & 2a + 4b + 6c \\ 1d + 3e + 5f & 2d + 4e + 6f \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} 1a + 2d & 1b + 2e & 1c + 2f \\ 3a + 4d & 3b + 4e & 3c + 4f \\ 5a + 6d & 5b + 6e & 5c + 6f \end{pmatrix}$$

2.4. **Properties of Matrix Algebra.** The reason the various algebraic operation are defined as they are is so that we may perform "algebra" on matrices. In particular, we have the following properties of matrix algebra. Note that they resemble the properties of the operations on real numbers. We assume below that A, B and C are matrices of compatible sizes (so that the given operations are defined) and α and β are scalars:

1. $A + B = B + A$ (matrix addition is commutative)
2. $(A + B) + C = A + (B + C)$ (matrix addition is associative)
3. $(AB)C = A(BC)$ (matrix multiplication is associative)
4. $\alpha(AB) = \quad (\alpha A)B = A(\alpha B)$
5. $\alpha(A + B) = \alpha A + \alpha B$ (scalar multiplication distributes over matrix addition)
6. $(\alpha + \beta)A = \alpha A + \beta A$
7. $(\alpha\beta)A = \alpha(\beta A)$
8. $A(B + C) = AB + AC$ (matrix multiplication distributes over matrix addition)
9. $(A + B)C = AC + BC$ (matrix multiplication distributes over matrix addition)

2.5. **The Inverse of a Matrix.** Consider the equation involving real numbers: $ax = b$. This equation can be solved for x by dividing both sides by a, unless $a = 0$ in which case there is no solution if $b \neq 0$. We can do a similar thing with matrix multiplication, and we run across the same problem as well. First a definition:

Definition 2.5.1. *Suppose that A is an $n \times n$ matrix and that there is a matrix B such that $AB = BA = I_n$. Then we say that B is the* **inverse** *of A, A is* **invertible** *or* **non-singular** *and we write $B = A^{-1}$.*

In subsequent notes, we will see that a square matrix does not necessarily have an inverse. This is ultimately a consequence of there being functions which are not one-to-one and hence have no inverse.

Example:

2.6. **The Transpose and Conjugate Transpose of a Matrix.** Let $A = (a_{ij})$ be a matrix of of size $m \times n$. We define the **transpose** of A to be the $n \times m$ matrix $A^T = (a_{ji})$. That is, the rows of A become the columns of A^T. When the entries of A are complex numbers, we can also define the **conjugate transpose** of A. This is the matrix $A^H = (\overline{a}_{ji})$, where $\overline{a}_{ji}$ is the complex conjugate of a_{ji}. Note that $A^H = A^T$ if the entries of A are all real.

Example:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \text{ and } A^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

and

$$A = \begin{pmatrix} 1+i & 2-i \\ 3+2i & 4i \\ 5 & 6 \end{pmatrix} \text{ and } A^H = \begin{pmatrix} 1-i & 3-2i & 5 \\ 2+i & -4i & 6 \end{pmatrix}$$

3. Matrix and Vector Equations

3.1. **Matrix Equations.** Suppose that $A = (a_{ij})$ is a matrix of of size $m \times n$. A matrix equation is an equation of the form:

$$A\mathbf{x} = \mathbf{b}$$

where $\mathbf{x}$ is a vector with n rows and $\mathbf{b}$ is a vector with m rows.

3.2. **Vector Equations.** Suppose that $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is a collection of vectors with m rows, and $x_1, \ldots, x_n$ are scalars. A vector equation is an equation of the form:

$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

where $\mathbf{b}$ is some vector with m rows.

3.3. **Relation of Linear Systems, Matrix Equations and Vector Equations.** Suppose that:

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

is a linear system. Considering the notion of equality for matrices, we have the equivalent vector equation:

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

and the equivalent matrix equation:

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