Matrices 1

1. An $n \times m$ matrix A is a rectangular array of numbers with n rows and m columns. By $A = (a_{ij})$ we mean that a_{ij} is the entry in the ith row and the jth column. For example,

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

is a 2×3 matrix. We denote by $\mathbf{R}^{n\times m}$ the class of $n \times m$ matrices with real entries.

2. An $n \times 1$ matrix is called a column vector, and a $1 \times m$ matrix, a row vector. An $n \times n$ matrix is called *square*. An $n \times n$ matrix $A = (a_{ij})$ is called diagonal if $a_{ij} = 0$ for $i \neq j$. The main diagonal of A is the set of elements a_{ii}, $i = 1, \ldots, n$.

3. The transpose of the $n \times m$ matrix $A = (a_{ij})$ is the $m \times n$ matrix $A^T = (a_{ji})$. Thus you get A^T from A by transposing the rows and the columns. For example, thr tranpose of

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

is

$$A^T = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ -2 & 4 \end{bmatrix},$$

and the transpose of

$$x = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \tag{1}$$

is

$$x^T = [\,1 \quad -2 \quad 3\,]. \tag{2}$$

Note that $\left(A^T\right)^T = A$.

4. For reasons we'll discuss later, we denote points in $\mathbf{R}^n$ by column vectors. For example, we write

$$x = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \tag{3}$$

instead of the $(2, 1)$ you might be used to. To save space while observing the column vector convention, some authors will write x as $[\,2 \quad 1\,]^T$ or $(2, 1)^T$.

5. Matrix Addition: If $A = (a_{ij})$ and $B = (b_{ij})$ are $n \times m$ matrices, then $A + B$ is the $n \times m$ matrix with ijth entry $a_{ij} + b_{ij}$.

6. Scalar Multiplication: If $A = (a_{ij})$ is an $n \times m$ matrix and c is a scalar, then cA is the $n \times m$ matrix with ijth entry ca_{ij}.

7. We usually write $-A$ instead of $-1A$. By $A - B$ we mean $A + (-1)B$.

8. Matrix Multiplication: If $A = (a_{ij})$ is $n \times m$ and $B = (b_{ij})$ is $m \times k$, then we can form the matrix product AB. To be precise, AB is the $n \times k$ matrix whose ijth entry is $\sum_{l=1}^{m} a_{il} b_{lj}$. In other words, the ijth entry of AB is the dot product of the ith row of A with the jth column of B.

9. Note that matrix multiplication is not commutative. If A is $n \times m$ and B is $m \times k$, where $k \neq n$, then AB is defined, but BA is not. Even if $k = n$, it is not generally true that $AB = BA$.

10. Let A, B and C be matrices and k a scalar. Then,

$$A + (B + C) = (A + B) + C, \tag{4}$$

$$A(B + C) = (A + B)C, \tag{5}$$

$$(AB)C = A(BC), \tag{6}$$

and

$$k(AB) = (kA)B = A(kB), \tag{7}$$

whenever the operations are defined.

11. The $n \times n$ *identity* is the matrix $I \in \mathbf{R}^{n \times n}$, with 1's one the main diagonal and 0's elsewhere:

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{8}$$

Let I be the $n \times n$ identity. If A is $n \times m$, then $IA = A$. If A is $m \times n$, then $AI = A$. In particular, if A is $n \times n$ and x is $n \times 1$, then

$$AI = IA = A, \tag{9}$$

and

$$Ix = x. \tag{10}$$

12. Consider the system of n equations in m unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n \end{aligned} \tag{11}$$

Let $A = (a_{ij})$, $x = [\, x_1 \cdots x_m \,]^T$ and $b = [\, b_1 \cdots b_n \,]^T$. Then the above system of equations can be written in matrix form as

$$Ax = b.$$

When $b = 0$ (that is, the zero vector in $\mathbf{R}^n$), the system is called homogeneous. In these notes, we'll only be concerned with the case of $m = n$.

13. Let I be the $n \times n$ identity. An matrix $A \in \mathbf{R}^{n\times n}$, is called *invertible* or *nonsinuglar* if there is a matrix A^{-1} such that

$$A^{-1}A = I. \tag{12}$$

The matrix A^{-1} is called the *inverse* of A. Note that A^{-1} must also be $n \times n$. If A has no inverse, it is called *singular*.

14. If $A \in \mathbf{R}^{n\times n}$, is nonsingular, then $\left(A^{-1}\right)^{-1} = A$, and $AA^{-1} = I$. If $B \in \mathbf{R}^{n\times n}$, is also nonsingular, then AB is nonsingular and $(AB)^{-1} = B^{-1}A^{-1}$

15. **Proposition:** Consider the n-dimensional system

$$Ax = b. \tag{13}$$

a. If A is invertible, then (13) has the unique solution $x = A^{-1}b$.

b. It follows from (a), that if A is invertible and $b = 0$, then (13) has the unique solution $x = 0$.

c. If A is singular and $b \neq 0$, then (13) has either no solution or infinitely many solutions.

d. If A is singular and $b = 0$, then (13) has infinitely many solutions.

16. **Proposition:** $A \in \mathbf{R}^{n\times n}$, is invertible if and only if $\det A \neq 0$.

17. The *null space* or *kernel* of $A \in \mathbf{R}^{n\times m}$, is

$$\mathcal{N}(A) = \{x \in \mathbf{R}^m \mid Ax = 0 \in \mathbf{R}^n\}.$$

If $m = n$, then by the previous two paragraphs,

$$\mathcal{N}(A) = \{0\} \Longleftrightarrow \det A \neq 0 \Longleftrightarrow A \text{ is nonsingular}. \tag{14}$$

18. Vectors $v_1, \ldots, v_m$ are *linearly indepedent* or simply *independent* if no one of them is a linear combination of the others. It isn't hard to show that $v_1, \ldots, v_m$ are independent if

$$c_1 v_1 + \cdots + c_m v_m = 0,$$

implies that

$$c_1 = c_2 = \cdots = c_m = 0.$$

In other words, the v_i are linearly independent if the only linear combination of the v_i that equals zero has coefficients that are all zero.

19. **Proposition:** For $j = 1, \ldots, n$, let

$$a_j = \begin{bmatrix} a_{1j} \\ \vdots \\ \alpha_{nj} \end{bmatrix}.$$

Let A be the $n \times n$ matrix with columns $a_1, \ldots, a_n$: $A = (a_{ij})$. Then $\det A \neq 0$ if and only if the column vectors $a_1, \ldots, a_n$ are linearly independent. Thus, for a square matrix A,

$$\text{The columns of } A \text{ are independent} \Longleftrightarrow \det A \neq 0 \Longleftrightarrow A \text{ is nonsingular.} \qquad (15)$$

20. Let A be an $n \times m$ matrix. If x is in $\mathbf{R}^m$, then Ax is in $\mathbf{R}^n$. Thus,

$$A : \mathbf{R}^m \mapsto \mathbf{R}^n.$$

Moreover,

$$A(x + y) = Ax + Ay,$$

and

$$A(cx) = cAx.$$

You can thus think of an $n \times m$ matrix A as a linear operator (or mapping, or transformation) taking $\mathbf{R}^m$ to $\mathbf{R}^n$. If B is $k \times n$, then BA takes $x \in \mathbf{R}^m$ to $Ax \in \mathbf{R}^n$ and then to $BAx \in \mathbf{R}^k$. In a nutshell,

$$BA : \mathbf{R}^m \mapsto \mathbf{R}^k, \textit{ linearly.}$$

You can thus think of matrix multiplication as composition of linear operators.