Chapter 1

1 Matrices

1.1 Basic Concepts

A *matrix* is a rectangular array arranged in vertical columns.

The matrix $L = \begin{bmatrix} 1 & 3 \\ 5 & 2 \\ 0 & -1 \end{bmatrix}$ is said to have *order* 3 x 2.

The entries of a matrix are called *elements*. l_{12} refers to the element in the first row and second column of the matrix L.

In general, a matrix A of order pxn has the form $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n|} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pn} \end{bmatrix}$

Any element having its row index equal to its column index is a *diagonal element.*

A matrix is *square* if it has the same number of rows as columns. In a square matrix, the elements $a_{11}, a_{22}, a_{33}, \dots$ form the *main* or *principal* diagonal.

The elements of a matrix need not be numbers; they can be functions or matrices themselves.

A *row matrix* is a matrix having a single row; a *column matrix* is a matrix having a single column. The elements of such a matrix are commonly called its *components*, and the number of components its *dimension.*

The term *n-tuple* refers to either a row matrix or a column matrix having dimension n.

Two matrices A and B are *equal* if they have the same order and if their correspondng elements are equal.

The sum of two matrices of the same order is a matrix obtained by adding together corresponding elements of the original matrices. Addition is not defined for matrices of different orders.

Theorem 1. If matrices A, B, C all have the same order, then

(a) the commutative law of addition holds; that is, $A + B = B + A$

(b) the associative law of addition holds; that is, $A + (B + C) = (A + B) + C$

Proof. (a) Let $A = [a_{ij}] \wedge B = [b_{ij}]$. Then

$$\begin{aligned} A + B &= [a_{ij}] + [b_{ij}] && \text{by defs. of A, B} \\ &= [a_{ij}] + [b_{ij}] && \text{by def. of matrix addition} \\ &= [b_{ij} + a_{ij}] && \text{by commutative property of addition} \\ &= [b_{ij}] + [a_{ij}] && \text{by def. of matrix addition} \\ &= B + A \end{aligned}$$

(b) Let $A = [a_{ij}], B = [b_{ij}]$, and $C = [c_{ij}]$. Then

$$\begin{aligned} A + (B + C) &= [a_{ij}] + ([b_{ij}] + [c_{ij}]) \\ &= [a_{ij}] + [b_{ij} + c_{ij}] && \text{by def. of matrix addition} \\ &= [a_{ij} + (b_{ij} + c_{ij})] && \text{by def. of matrix addition} \\ &= [(a_{ij} + b_{ij}) + c_{ij}] && \text{by associative property of addition} \\ &= [(a_{ij} + b_{ij})] + [c_{ij}] && \text{by def. of matrix addition} = (A + B) + C \end{aligned}$$

□

We define the matrix 0 to be a matrix consisting of only zero elements. When a zero matrix has the same order as another matrix A, we have the additional property $A + 0 = A$.

Subtraction of matrices is defined analogously to addition. The difference $B - A$ of two matrices of the same order is the matrix obtained by subtracting from the elements of A the corresponding elements of B.

A matrix A can always be added to itself, forming the sum $A+A$. We would like to write $A+A = 2A$.

The right side of the equation is a number times a matrix, a product known as *scalar multiplication*.

If $A = [a_{ij}]$ is a pxn matrix and if λ is a real number, then $\lambda A = [\lambda a_{ij}], (i = 1, 2, \ldots, p; j = 1, 2, \ldots, n)$.

Theorem 2. If $A \wedge B$ are matrices of the same order and if $\lambda_1 \wedge \lambda_2$ denote scalars, then the following distributive laws hold:

(a) $\lambda_1(A + B) = \lambda_1 A + \lambda_1 B$.

(b) $(\lambda_1 + \lambda_2)A = \lambda_1 A + \lambda_2 A$

(c) $(\lambda_1 \lambda_2)A = \lambda_1(\lambda_2 A)$

Proof. (a) Let $A = [a_{ij}] \wedge B = [b_{ij}]$. Then

$$\begin{aligned}
\lambda_1(A+B) &= \lambda_1([a_{ij}] + [b_{ij}]) \\
&= \lambda_1[(a_{ij} + b_{ij})] && \text{def. of matrix addition} \\
&= [\lambda_1(a_{ij} + b_{ij})] && \text{def. of scalar multiplication} \\
&= [(\lambda_1 a_{ij} + \lambda_1 b_{ij})] && \text{distributive property of scalars} \\
&= [\lambda_1 a_{ij}] + [\lambda_1 b_{ij}] && \text{def. of matrix addition} \\
&= \lambda_1[a_{ij}] + \lambda_1[b_{ij}] && \text{def. of scalar multiplication} = \lambda_1 A + \lambda_1 B
\end{aligned}$$

(b) Let $A = [a_{ij}]$. Then

$$\begin{aligned}
(\lambda_1 + \lambda_2)A &= (\lambda_1 + \lambda_2)[a_{ij}] \\
&= [(\lambda_1 + \lambda_2)a_{ij}] && \text{def. of scalar multiplication} \\
&= [\lambda_1 a_{ij} + \lambda_2 a_{ij}] && \text{distributive property of multiplication} \\
&= [\lambda_1 a_{ij}] + [\lambda_2 a_{ij}] && \text{def. of matrix addition} \\
&= \lambda_1[a_{ij}] + \lambda_2[a_{ij}] && \text{def. of scalar multiplication} \\
&= \lambda_1 A + \lambda_2 A
\end{aligned}$$

(c) Let $A = [a_{ij}]$. Then

$$\begin{aligned}
(\lambda_1\lambda_2)A &= (\lambda_1\lambda_2)[a_{ij}] \\
&= [(\lambda_1\lambda_2)a_{ij}] && \text{def. of scalar multiplication} \\
&= [\lambda_1(\lambda_2 a_{ij})] && \text{associative property of multiplication} \\
&= \lambda_1[\lambda_2 a_{ij}] && \text{def. of scalar multiplication} \\
&= \lambda_1(\lambda_2 A)
\end{aligned}$$

□

1.2 Matrix Multiplication

A single system of two linear equations in two unknowns is

$$\begin{aligned}
2x + 3y &= 10 \\
4x + 5y &= 20
\end{aligned}$$

Combining all the coefficients of the variables on the left of each equation into a *coefficient matrix*, all the variables into a column matrix of variables; and the constants on the right of each equation into another column matrix, we separate the matrix system

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 10 \\ 20 \end{bmatrix}$$

We want to define matrix multiplication so that

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (2x+3y) \\ (4x+5y) \end{bmatrix} = \begin{bmatrix} 10 \\ 20 \end{bmatrix}$$

We shall define the product AB of two matrices A and B when the number of columns of A is equal to the number of rows of B, and the result will be a matrix having the same number of rows as A and the same number of columns as B.

When the product AB is considered, A is said to *premultiply* B, while B is said to *postmultiply* A.

To calculate the $i-j$ element of AB, when the multiplication is defined, multiply the elements in the ith row of A by the corresponding elements in the jth column of B and sum the results.

Let $AB = C = [c_{ij}]$. Then $[c_{ij}] = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj} = \sum_{k=1}^{r} a_{ik}b_{kj}$.

Although matrix multiplication is not commutative, some matrix products are. Also, matrices exist for which $AB = 0$ without either A or B being 0. The cancellation law also does not apply. In general, $AB = AC$ does not imply $B = C$.

Theorem 3. If $A, B, \wedge C$ have appropriate orders so that the following additions and multiplications are defined, then

(a) $A(BC) = (AB)C$ (associative law of multiplication)

(b) $A(B+C) = AB + AC$ (left distributive law)

(c) $(B+C)A = BA + CA$ (right distributive law)

Proof. (a) Let $A = [a_{ij}]$ be an mxn matrix, $B = [b_{ij}]$ be an nxp matrix, and $C = [c_{ij}]$ be a p$x$$q$ matrix. Then

$$\begin{aligned} (AB)C &= ([a_{ij}][b_{jk}])[c] \\ &= (\sum_{k=1}^{r} a_{ik}b_{kj})[c] \\ &= \sum_{l=1}^{p}(\sum_{k=1}^{r} a_{ik}b_{kj})c \\ &= \sum_{l}([ab]_{il}[c]) \\ &= [ab]_{il}[c]_{lj} \\ &= (AB)C \end{aligned}$$

Note 1. I got lost in the sea of subscripts, but labeled appropriately (and arbitrarily) the logic still holds.

(b) Let $A = [a_{ij}] \wedge B = [b_{ij}] \wedge C = [c_{ij}]$, where A is mxn, and B and C are nxp. Then

$$\begin{aligned} A(B+C) &= [a_{ij}]([b_{ij}] + [c_{ij}]) \\ &= [a_{ij}][(b_{ij} + c_{ij})] \\ &= \sum_k a_{ik}(b_{kj} + c_{kj}) \\ &= \sum_k a_{ik}b_{kj} + a_{ik}c_{kj} \\ &= \sum_k a_{ik}b_{kj} + \sum_k a_{ik}c_{kj} \\ &= [a_{ij}][b_{ij}] + [a_{ij}][c_{ij}] \\ &= AB + AC \end{aligned}$$

(c) Let $A = [a_{ij}] \wedge B = [b_{ij}] \wedge C = [c_{ij}]$, where A is mxn, and B and C are nxp. Then

$$\begin{aligned} (B+C)A &= ([b_{ij}] + [c_{ij}])[a_{ij}] \\ &= \sum_k (b_{ik} + c_{ik})a_{kj} \\ &= \sum_k (b_{ik}a_{kj} + c_{ik})a_{kj} \\ &= \sum_k (b_{ik}a_{kj} + \sum_k c_{ik})a_{kj} \\ &= BA + BC \end{aligned}$$

□