# 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 $p$x$n$ 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 $p$x$n$ matrix and if $\lambda$ 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 $i$th row of $A$ by the corresponding elements in the $j$th 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 $m$x$n$ matrix, $B = [b_{ij}]$ be an $n$x$p$ 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 $m$x$n$, and $B$ and $C$ are $n$x$p$. 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 $m$x$n$, and $B$ and $C$ are $n$x$p$. Then

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

$\square$

### 1.3 Special Matrices

The transpose of a matrix $A$, denoted by $A^T$, is obtained by converting all the rows of $A$ into the columns of $A^T$ while preserving the ordering of the rows/columns. More formally, if $A = [a_{ij}]$ is an $n$x$p$ matrix, then the transpose of A, denoted by $A^T = [a^T_{ij}]$, is a $p$x$n$ matrix where $a^T_{ij} = a_{ji}$.

**Theorem 4.** The following properties are true for any scalar $\lambda$ and any matrices for which the indicated additions and multiplications are defined.

(a) $(A^T)^T = A$

(b) $(\lambda A)^T = \lambda A^T$

(c) $(A+B)^T = A^T + B^T$

(d) $(AB)^T = B^T A^T$

*Proof.* (a) Let $A$ be an $m$x$n$ matrix $A = [a_{ij}]$. Then

$$\begin{aligned}(A^T)^T &= ([a_{ij}]^T)^T \\ &= ([a_{ji}])^T \\ &= [a_{ij}]\end{aligned}$$

(b)

$$\begin{aligned}(\lambda A)^T &= (\lambda[a_{ij}])^T \\ &= ([\lambda a_{ij}])^T \\ &= [\lambda a_{ji}] \\ &= \lambda[a_{ji}] \\ &= \lambda[a_{ij}]^T \\ &= \lambda A^T\end{aligned}$$

(c)

$$(A+B)^T = ([a_{ij}] + [b_{ij}])^T = ([a_{ij} + b_{ij}])^T = [a_{ji} + b_{ji}] = [a_{ji}] + [b_{ji}] = A^T + B^T$$

(d) Note: I'm pretty sure the following is wrong.

$$(AB)^T = (\sum_k a_{ik} b_{kj})^T = \sum_k a_{jk} b_{ki} = \sum_k a^T_{kj} b^T_{ik} = \sum_k b^T_{ik} a^T_{kj} = (B^T A^T)$$

$\square$

A matrix $A$ is *symmetric* if it equals its own transpose; that is, if $A = A^T$. A matrix $A$ is *skew-symmetric* if it equals the negative of its transpose; that is, if $A = -A^T$.

A *submatrix* of a matrix $A$ is a matrix obtained by removing any number of rows or columns from $A$. By removing no rows and no columns from $A$, it follows that $A$ is a submatrix of itself.

A matrix is *partitioned* if it is divided into submatrices by horizontal or vertical lines between rows and columns.

A *zero row* in a matrix is a row containing only zero elements, whereas a nonzero row is a row that contains at least one nonzero element.

**Definition 1.** A matrix is in *row-reduced form* if it satisfies the following four conditions:

1. All zero rows appear below nonzero rows when both types are present in the matrix.
2. The first nonzero element in any nonzero row is 1.
3. All elements directly below (that is, in the same column but in succeeding rows from) the first nonzero elements of a nonzero row are zero.
4. The first nonzero element of any nonzero row appears in a later column (further to the right) than the first nonzero element in any preceding row.

A *diagonal matrix* is a square matrix having only zeros as non-diagonal elements.

An *identity matrix*, denoted by $I$, is a diagonal matrix having all its diagonal elements equal to 1, and 0 otherwise.

If $A$ and $I$ are square matrices of the same order, then $AI = IA = A$.

A *block diagonal matrix* $A$ is one that can be partitioned into the form

$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & 0\dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & A_k \end{bmatrix}$, where $A_1, A_2, ..., A_k$ are square matrices.

A matrix $A = [a_{ij}]$ is *upper triangular* if $a_{ij} = 0 \; \forall i > j$; that is, if all elements below the main diagonal are zero.

If $a_{ij} = 0 \; \forall i < j$; that is, if all elements above the main diagonal are zero, then $A$ is *lower triangular*.

**Theorem 5.** The product of two lower (upper) triangular matrices of the same order is also lower (upper) triangular.

*Proof.* Let $A = [a_{ij}] \wedge B = [b_{ij}]$ both be $nxn$ lower triangular matrices. Then

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

Since $a_{ik} = 0 \ \forall i < j$

?

Since $b_{kj} = 0 \ \forall k < j$

??

$\square$

## 1.4 Linear Systems of Equations

A system of $m$-linear equations in $n$-variables $x_1, x_2, \ldots, x_n$ has the general form