MATRIX ALGEBRA

Simultaneous Equations

Consider a system of m linear equations in n unknowns:

$$(1)\qquad \begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ &\vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n. \end{aligned}$$

There are three sorts of elements here:

The constants: $\{y_i; i = 1, \ldots, m\}$,

The unknowns: $\{x_j; j = 1, \ldots, n\}$,

The coefficients: $\{a_{ij}; i = 1, \ldots, m, j = 1, \ldots, n\}$;

and they can be gathered into three arrays:

$$(2)\qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{13} \\ a_{21} & a_{22} & \cdots & a_{23} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

The arrays y and x are column vectors of order m and n respectively whilst the array A is a matrix of order $m \times n$, which is to say that it has m rows and n columns. A summary notation for the equations under (1) is then

$$(3)\qquad y = Ax.$$

There are two objects on our initial agenda. The first is to show, in detail, how the summary matrix representation corresponds to the explicit form of the equation under (1). For this purpose we need to define, at least, the operation of matrix multiplication.

The second object is to describe a method for finding the values of the unknown elements. Each of the m equations is a statement about a linear relationship amongst the n unknowns. The unknowns can be determined if and only if there can be found, amongst the m equations, a subset of n equations which are mutually independent in the sense that none of the corresponding statements can be deduced from the others.

Example. Consider the system

$$(4)\qquad \begin{aligned} 9 &= x_1 + 3x_2 + 2x_3, \\ 8 &= 4x_1 + 5x_2 - 6x_3, \\ 8 &= 3x_1 + 2x_2 + x_3. \end{aligned}$$

The corresponding arrays are

$$(5) \qquad y = \begin{bmatrix} 9 \\ 8 \\ 8 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 5 & -6 \\ 3 & 2 & 1 \end{bmatrix}, \qquad x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Here we have placed the solution for x_1, x_2 and x_3 within the vector x. The correctness of these values may be confirmed by substituting them into the equations of (4).

Elementary Operations with Matrices

It is often useful to display the generic element of a matrix together with the symbol for the matrix in the summary notation. Thus, to denote the $m \times n$ matrix of (2), we write $A = [a_{ij}]$. Likewise, we can write $y = [y_i]$ and $x = [x_j]$ for the vectors. In fact, the vectors y and x may be regarded as degenerate matrices of orders $m \times 1$ and $n \times 1$ respectively. The purpose of this is to avoid having to enunciate rules of vector algebra alongside those of matrix algebra.

Matrix Addition. *If $A = [a_{ij}]$ and $B = [b_{ij}]$ are two matrices of order $m \times n$, then their sum is the matrix $C = [c_{ij}]$ whose generic element is $c_{ij} = a_{ij} + b_{ij}$.*

The sum of A and B is defined only if the two matrices have the same order which is $m \times n$; in which case they are said to be conformable with respect to addition. Notice that the notation reveals that the matrices are conformable by giving the same indices $i = 1, \ldots, m$ and $j = 1, \ldots, n$ to their generic elements.

The operation of matrix addition is commutative such that $A + B = B + A$ and associative such that $A + (B + C) = (A + B) + C$. These results are, of course, trivial since they amount to nothing but the assertions that $a_{ij} + b_{ij} = b_{ij} + a_{ij}$ and that $a_{ij} + (b_{ij} + c_{ij}) = (b_{ij} + a_{ij}) + c_{ij}$ for all i, j.

Scalar Multiplication of Matrices. *The product of the matrix $A = [a_{ij}]$ with an arbitrary scalar, or number, λ is the matrix $\lambda A = [\lambda a_{ij}]$.*

Matrix Multiplication. *The product of the matrices $A = [a_{ij}]$ and $B = [b_{jk}]$ of orders $m \times n$ and $n \times p$ respectively is the matrix $AB = C = [c_{ik}]$ of order $m \times p$ whose generic element is $c_{ik} = \sum_j a_{ij} b_{jk} = a_{i1} b_{1k} + a_{i2} b_{2k} + \cdots + a_{in} b_{nk}$.*

The product of AB is defined only if B has a number n of rows equal to the number of columns of A, in which case A and B are said to be conformable with respect to multiplication. Notice that the notation reveals that the matrices are conformable by the fact that j is, at the same time, the column index of $A = [a_{ij}]$ and the row index of $B = [b_{jk}]$. The manner in which the product $C = AB$ inherits its orders from its factors is revealed in the following display:

$$(6) \qquad (C : m \times p) = (A : m \times n)(B : n \times p).$$

The operation of matrix multiplication is not commutative in general. Thus, whereas the product of $A = [a_{ij}]$ and $B = [b_{jk}]$ is well-defined by virtue of the common index j, the product BA is not defined unless the indices $i = 1, \ldots, m$

and $k = 1, \ldots, p$ have the same range, which is to say that we must have $m = p$. Even if BA is defined, there is no expectation that $AB = BA$, although this is a possibility when A and B are conformable square matrices with equal numbers of rows and columns.

The rule for matrix multiplication which we have stated is sufficient for deriving the explicit expression for m equation in n unknowns found under (1) from the equation under (3) and the definitions under (2).

Example. If

$$(7) \qquad A = \begin{bmatrix} 1 & -4 & 2 \\ 2 & 3 & 6 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 4 & 2 \\ 1 & 2 \\ -3 & 1 \end{bmatrix}$$

Then

$$(8) \qquad \begin{aligned} AB &= \begin{bmatrix} 1(4) - 4(1) - 2(3) & 1(2) - 4(2) + 2(1) \\ 2(4) + 3(1) - 6(3) & 2(2) + 3(2) + 6(1) \end{bmatrix} = \begin{bmatrix} -6 & -4 \\ -7 & 16 \end{bmatrix} \quad \text{and} \\ BA &= \begin{bmatrix} 4(1) + 2(2) & -4(4) + 2(3) & 4(2) + 2(6) \\ 1(1) + 2(2) & -1(4) + 2(3) & 1(2) + 2(6) \\ -3(1) + 1(2) & 3(4) + 1(3) & -3(2) + 1(6) \end{bmatrix} = \begin{bmatrix} 8 & -10 & 20 \\ 5 & 2 & 14 \\ -1 & 15 & 0 \end{bmatrix}. \end{aligned}$$

Transposition. *The transpose of the matrix $A = [a_{ij}]$ of order $m \times n$ is the matrix $A' = [a_{ji}]$ of order $n \times m$ which has the rows of A for its columns and the columns of A for its rows.*

Thus the element of A from the ith row and jth column becomes the element of the jth row and ith column of A'. The symbol $\{'\}$ is a prime, and we refer to A' equally as A–*prime* or A–*transpose*.

Example. Let D be the sub-matrix formed by taking the first two columns of the matrix A of (5). Then

$$(9) \qquad D = \begin{bmatrix} 1 & 3 \\ 4 & 5 \\ 3 & 2 \end{bmatrix} \qquad \text{and} \qquad D' = \begin{bmatrix} 1 & 4 & 3 \\ 3 & 5 & 2 \end{bmatrix}.$$

The basic rules of transposition are as follows

$$(10) \qquad \begin{aligned} &\text{(i)} && \text{The transpose of } A' \text{ is } A, \text{ that is } (A')' = A, \\ &\text{(ii)} && \text{If } C = A + B \text{ then } C' = A' + B', \\ &\text{(iii)} && \text{If } C = AB \text{ then } C' = B'A'. \end{aligned}$$

Of these, only (iii), which is called the reversal rule, requires explanation. For a start, when the product is written as

$$(11) \qquad (C' : p \times m) = (B' : p \times n)(A' : n \times m)$$

and when the latter is compared with the expression under (6), it becomes clear that the reversal rule ensures the correct orders for the product. More explicitly,

$$
(12)\qquad \begin{aligned}
&\text{if}\quad A=[a_{ij}],\quad B=[b_{jk}]\quad\text{and}\quad AB=C=[c_{ik}],\\
&\qquad\text{where}\quad c_{ik}=\sum_{j=1}^{n}a_{ij}b_{jk},\\
&\text{then}\qquad A'=[a_{ji}],\quad B'=[b_{kj}]\quad\text{and}\quad B'A'=C'=[c_{ki}]\\
&\qquad\text{where}\quad c_{ki}=\sum_{j=1}^{n}b_{kj}a_{ji}.
\end{aligned}
$$

Matrix Inversion. *If $A=[a_{ij}]$ is a square matrix of order $n\times n$, then its inverse, if it exists, is a uniquely defined matrix A^{-1} of order $n\times n$ which satisfies the condition $AA^{-1}=A^{-1}A=I$, where $I=[\delta_{ij}]$ is the identity matrix of order n which has units on its principal diagonal and zeros elsewhere.*

In this notation, δ_{ij} is Kronecker's delta defined by

$$
(13)\qquad \delta_{ij}=\begin{cases}0, & \text{if } i\neq j;\\ 1, & \text{if } i=j.\end{cases}
$$

Usually, we may rely upon the computer to perform the inversion of a numerical matrix of order 3 or more. Also, for orders of three or more, the symbolic expressions for the individual elements of the inverse matrix become intractable.

In order to derive the explicit expression for the inverse of a 2×2 matrix A, we may consider the following equation $BA=I$:

$$
(14)\qquad \begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix}\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}.
$$

The elements of the inverse matrix $A^{-1}=B$ are obtained by solving the following equations in pairs:

$$
(15)\qquad \begin{aligned}
&\text{(i)}\quad b_{11}a_{11}+b_{12}a_{21}=1, &\qquad &\text{(ii)}\quad b_{11}a_{12}+b_{12}a_{22}=0,\\
&\text{(iii)}\quad b_{21}a_{11}+b_{22}a_{21}=0, &\qquad &\text{(iv)}\quad b_{21}a_{12}+b_{22}a_{22}=1.
\end{aligned}
$$

From (i) and (ii) we get

$$
(16)\qquad b_{11}=\frac{a_{22}}{a_{11}a_{22}-a_{12}a_{21}}\quad\text{and}\quad b_{12}=\frac{-a_{12}}{a_{11}a_{22}-a_{12}a_{21}},
$$

whereas, from (iii) and (iv), we get

$$
(17)\qquad b_{21}=\frac{-a_{21}}{a_{11}a_{22}-a_{12}a_{21}}\quad\text{and}\quad b_{22}=\frac{a_{11}}{a_{11}a_{22}-a_{12}a_{21}}.
$$

The common denominator in these expressions is the so-called determinant of the original matrix A.