MATRICES, VECTORS, DETERMINANTS, AND LINEAR ALGEBRA

Tadao ODA
Tohoku University, Japan

Keywords: matrix, determinant, linear equation, Cramer's rule, eigenvalue, Jordan canonical form, symmetric matrix, vector space, linear map

Contents

Summary

A down-to-earth introduction of matrices and their basic operations will be followed by

basic results on determinants, systems of linear equations, eigenvalues, real symmetric matrices and complex Hermitian symmetric matrices.

Abstract vector spaces and linear maps will then be introduced. The power and merit of seemingly useless abstraction will make earlier results on matrices more transparent and easily understandable.

Matrices and linear algebra play important roles in applications. Unfortunately, however, space limitation prevents description of algorithmic and computational aspects of linear algebra indispensable to applications. The readers are referred to the references listed at the end.

1. Matrices, Vectors and their Basic Operations

1.1. Matrices

A matrix is a rectangular array

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$

of *entries* $a_{11},\ldots,a_{mn}$, which are numbers or symbols. Very often, such a matrix will be denoted by a single letter such as $\mathbf{A}$, thus

$$\mathbf{A} := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}.$$

The notation $\mathbf{A} = (a_{ij})$ is used also, for short. In this notation, the first index i is called the *row index*, while the second index j is called the *column index*.

Each of the horizontal arrays is called a *row*, thus

$$(a_{11}, a_{12},\ldots,a_{1j},\ldots,a_{1n}),(a_{21}, a_{22},\ldots,a_{2j},\ldots,a_{2n}),\ldots,(a_{i1}, a_{i2},\ldots,a_{ij},\ldots,a_{in}),\ldots,(a_{m1}, a_{m2},\ldots,a_{mj},\ldots,a_{mn})$$

are called the first row, second row,…, i-th row,…, m-th row, respectively. On the other hand, each of the vertical arrays is called a *column*, thus

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{i1} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{i2} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{mj} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{in} \\ \vdots \\ a_{mn} \end{pmatrix}$$

are called the first column, second column,…, j-th column, …, n-th column, respectively. Such an $\mathbf{A}$ is called a matrix with m rows and n columns, an (m,n)-matrix, or an $m \times n$ matrix.

An (m,n)-matrix with all the entries 0 is called the *zero matrix* and written simply as $\mathbf{0}$, thus

$$\mathbf{0} := \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

1.2. Vectors

A matrix with only one row, or only one column is called a *vector*, thus
$(a_1, a_2, \ldots, a_j, \ldots, a_n)$

is a *row vector*, while

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{pmatrix}$$

is a *column vector*.

The rows and columns of an (m,n)-matrix $\mathbf{A}$ above are thus called, the first row vector, second row vector,…, i-th row vector,…, m-th row vector, and the first column vector, second column vector,…, j-th column vector, …, n-th column vector.

A $(1,1)$-matrix, i.e., a number or a symbol, is called a *scalar*.

1.3. Addition and Scalar Multiplication of Matrices

The *addition* of two (m,n)-matrices $\mathbf{A}=(a_{ij})$ and $\mathbf{B}=(b_{ij})$ are defined by

$$\mathbf{A}+\mathbf{B}:=(a_{ij}+b_{ij})=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1j}+b_{1j} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2j}+b_{2j} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1}+b_{i1} & a_{i2}+b_{i2} & \cdots & a_{ij}+b_{ij} & \cdots & a_{in}+b_{in} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mj}+b_{mj} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

when the addition of the entries makes sense. The multiplication of a scalar c with an (m,n)-matrix $\mathbf{A}=(a_{ij})$ is defined by

$$c\mathbf{A}:=(ca_{ij})=\begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1j} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2j} & \cdots & ca_{2n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ ca_{i1} & ca_{i2} & \cdots & ca_{ij} & \cdots & ca_{in} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mj} & \cdots & ca_{mn} \end{pmatrix}$$

when the multiplication of a scalar with the entries makes sense.

1.4. Multiplication of Matrices

What makes matrices most interesting and powerful is the multiplication, which does wonders as explained below.

Suppose that the entries appearing in our matrices are numbers which admit multiplication. Then the *multiplication* $\mathbf{AB}$ of two matrices $\mathbf{A}$ and $\mathbf{B}$ is defined when *the number of columns of* $\mathbf{A}$ *is the same as the number of rows of* $\mathbf{B}$.

Let $\mathbf{A}=(a_{ij})$ be an (l,m)-matrix and $\mathbf{B}=(b_{jk})$ an (m,n)-matrix. Then their product is the (l,n)-matrix defined by

$$\mathbf{AB}:=(c_{ik}), \qquad \text{with} \quad c_{ik}:=\sum_{j=1}^{m} a_{ij}b_{jk},$$

or more concretely,

$$\mathbf{AB} = \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1m}b_{m1} & \cdots & a_{11}b_{1k} + \cdots + a_{1m}b_{mk} & \cdots & a_{11}b_{1n} + \cdots + a_{1m}b_{mn} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1}b_{11} + \cdots + a_{im}b_{m1} & \cdots & a_{i1}b_{1k} + \cdots + a_{im}b_{mk} & \cdots & a_{i1}b_{1n} + \cdots + a_{im}b_{mn} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{l1}b_{11} + \cdots + a_{lm}b_{m1} & \cdots & a_{l1}b_{1k} + \cdots + a_{lm}b_{mk} & \cdots & a_{l1}b_{1n} + \cdots + a_{lm}b_{mn} \end{pmatrix}.$$

Of particular interest is the product $\mathbf{Av}$ of an (m,n)-matrix $\mathbf{A} = (a_{ij})$ with a column vector $\mathbf{v}$ of size n, which is the column vector of size m defined by

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_j \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + \cdots + a_{1n}v_n \\ \vdots \\ a_{i1}v_1 + \cdots + a_{in}v_n \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n \end{pmatrix},$$

as well as the product $\mathbf{uA}$ of a row vector $u = (u_1,\ldots,u_m)$ of size m with $\mathbf{A}$, which is the row vector of size n defined by

$$(u_1,\ldots,u_i,\ldots,u_m) \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$
$$=(u_1a_{11}+\cdots+u_ma_{m1},\ldots,u_1a_{1j}+\cdots+u_ma_{mj},\ldots,u_1a_{1n}+\cdots+u_ma_{mn}).$$

The *transpose* $\mathbf{A}^{\mathrm{T}}$ of an (m,n)-matrix $\mathbf{A} = (a_{ij})$ is the (n,m)-matrix defined by

$$\mathbf{A}^{\mathrm{T}} := (a'_{ji}), \qquad \text{with} \quad a'_{ji} := a_{ij},$$

or more concretely,

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}^{\mathrm{T}} := \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{i1} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{i2} & \cdots & a_{m2} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{1j} & a_{2j} & \cdots & a_{ij} & \cdots & a_{mj} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{in} & \cdots & a_{mn} \end{pmatrix}.$$

For an (l,m)-matrix $\mathbf{A}$ and an (m,n)-matrix $\mathbf{B}$, it is easy to see that

$$(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}},$$

when the multiplication of the numbers concerned is *commutative*.

When $\mathbf{A}$ and $\mathbf{B}$ are (n,n)-matrices, both products $\mathbf{AB}$ and $\mathbf{BA}$ make sense, but they *need not be the same* in general.

2. Determinants

2.1. Square Matrices

Square matrices, namely matrices with the same number of rows and columns, are most interesting.

Special among them is the *identity matrix* of size n, denoted by $\mathbf{I}$ or $\mathbf{I}_n$ and defined by

$$\mathbf{I} = \mathbf{I}_n := \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} = (\delta_{ij}),$$

where δ_{ij} is known as *Kronecker's delta* defined by

$$\delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

For an arbitrary (m,n)-matrix $\mathbf{A}$, the following clearly holds:

$$\mathbf{I}_m\mathbf{A} = \mathbf{A} \qquad \text{and} \qquad \mathbf{A}\mathbf{I}_n = \mathbf{A}.$$

The matrix $c\mathbf{I}$ with the same entry c along the diagonal and 0 elsewhere is called a *scalar matrix*.

More generally, a square matrix $\mathbf{D} = (d_{ij})$ of size n is called a *diagonal matrix* if $d_{ij} = 0$ for $i \neq j$, that is,

$$\mathbf{D} = \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix}.$$

2.2. Determinants

Let $\mathbf{A} = (a_{ij})$ be a square matrix of size n (also said to be of *order* n), that is, an (n,n)-matrix or an $n \times n$ matrix. When the entries a_{ij} are *numbers* (rational numbers, real numbers, complex numbers, or more generally elements of a commutative ring to be introduced in *Rings and Modules*), for which *addition*, *subtraction* and *commutative multiplication* are possible, associated to $\mathbf{A}$ is a number called the *determinant* of $\mathbf{A}$ and denoted by $|\mathbf{A}|$ or by $\det(\mathbf{A})$.

When $n = 1$ or $n = 2$, the determinant is defined to be

$$|a_{11}| := a_{11}, \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} := a_{11}a_{22} - a_{12}a_{21}.$$

For $n = 3$, the formula is a bit more complicated.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} := a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

The determinant of $\mathbf{A} = (a_{ij})$ for general n is defined as follows:

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} := \sum_{\sigma} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)},$$

where σ runs through the *permutations* of the indices $\{1, 2, \ldots, \ldots, n\}$, and $\operatorname{sgn}(\sigma)$ is the *signature* of σ to be defined elsewhere in *Groups and Applications*, since it is not so practical to compute the determinant using this formula. Instead, there is an inductive way of computing the determinant: If how to compute the determinants of square matrices of size $n-1$ is known, then the determinant of a square matrix $\mathbf{A}$ of size n is defined by

$$|\mathbf{A}| := \sum_{j=1}^{n} a_{1j}\Delta_{1j} = a_{11}\Delta_{11} + a_{12}\Delta_{12} + \cdots + a_{1n}\Delta_{1n},$$

where, for i and j in general, Δ_{ij} is the (i,j)-*cofactor* of $\mathbf{A}$ defined by

$$\Delta_{ij} := (-1)^{i+j} \det(\ \mathbf{A} \text{ with the } i\text{-th row and the } j\text{-th column removed }).$$

This formula is known as the expansion of $|\mathbf{A}|$ with respect to the first row. In fact, it can be shown that the expansion with respect to the i-th row for any $i = 1, \ldots, n$ gives rise to the same number:

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{ij}\Delta_{ij} = a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in}.$$

A similar formula holds when the role of rows and columns is interchanged, that is, the expansion of $|\mathbf{A}|$ with respect to the j-th column holds as well. In particular, $\left|\mathbf{A}^{\mathrm{T}}\right| = |\mathbf{A}|$.

For square matrices $\mathbf{A}$ and $\mathbf{B}$ of size n, it can be shown that

$$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|.$$

-
-
-

Bibliography

Chatelin, F. (1993): *Eigenvalues of Matrices*. With exercises by Mario Ahués and the author. Translated from the French and with additional material by Walter Ledermann. xviii+382 pp., John Wiley & Sons, Ltd., Chichester, ISBN 0-471-93538-7 [This is one of the advanced books that deal with algorithmic and computational aspects of matrices and linear algebra.]

Hogbert, L. ed. (2007). Handbook of Linear Algebra, 1400 pp.,Chapman & Hall/CRC, Boca Raton, Florida, ISBN 987-1-58488-510-8 [A complete treatment of linear algebra.]

Lang, S. (1989). Linear Algebra, Reprint of the Third Edition, x+285pp., Undergraduate Texts in Mathematics, Springer-Verlag, New York, ISBN 0-387-96412-6 [A textbook intended for a one-term course at the junior or senior level. Originally published from Addison-Wesley Publishing Company.]

Roman, S. (2005). Advanced Linear Algebra, Second Edition, xvi+482 pp., Graduate Texts in Mathematics 135, Springer-Verlag, New York, ISBN 0-387-97837-2 [A textbook intended for advanced courses in linear algebra for graduate students or advanced undergraduate students.]

Biographical Sketch

Tadao ODA, born 1940 in Kyoto, Japan

Education: BS in Mathematics, Kyoto University, Japan (March, 1962). MS in Mathematics, Kyoto University, Japan (March, 1964). Ph.D. in Mathematics, Harvard University, U.S.A. (June, 1967).

Positions held: Assistant, Department of Mathematics, Nagoya University, Japan (April, 1964-July, 1968) Instructor, Department of Mathematics, Princeton University, U.S.A. (September, 1967-June, 1968) Assistant Professor, Department of Mathematics, Nagoya University, Japan (July, 1968-September, 1975) Professor, Mathematical Institute, Tohoku University, Japan (October, 1975-March, 2003) Professor Emeritus, Tohoku University (April, 2003 to date)