Metadata of the chapter that will be visualized online

Book Title	Encyclopedia of Social Network Analysis and Mining	
Book Copyright Year	2014	
Copyright Holder	Springer Science+Business Media New York	
Title	Matrix Algebra, Basics of	

Author	Degree	
	Given Name	**Ayman**
	Particle	
	Family Name	**Badawi**
	Suffix	
	Phone	
	Fax	
	Email	abadawi@aus.edu
Affiliation	Division	Department of Mathematics
	Organization	American University of Sharjah
	Street	Box 26666
	City	Sharjah
	Country	UAE

M

Matrix Algebra, Basics of

Ayman Badawi
Department of Mathematics, American University of Sharjah, Sharjah, UAE

Synonyms

Addition and subtraction of matrices; Augmented matrix; Consistent and inconsistent system; Cramer rule; Determinant of a matrix; Echelon form; Elementary matrix; Invertible (nonsingular) matrices; Multiplication of matrices; Symmetric and skew-symmetric matrices; Transpose of a matrix

Glossary

Matrix $n \times m$ A matrix is a block consisting of n row and m column. An entry in a matrix B located in the ith row and jth column of B is denoted by b_{ij}

Consistent and Inconsistent System of Linear Equations A system of linear equations is said to be consistent if it has a solution, and it is called inconsistent if it has no solutions

Augmented Matrix An augmented matrix of a system of linear equations written in matrix-form $CX = B$ is a matrix of the form $[C|B]$, where C is the coefficient matrix of the system and B is the constant column of the system

Transpose of a Matrix The transpose of a matrix A is denoted by A^T such that $a_{ij}^T = a_{ji}$

Symmetric Matrix and Skew-Symmetric Matrix A square matrix A, $n \times n$, is said to be symmetric if $A^T = A$, and it is called a skew-symmetric if $A^T = -A$

Identity Matrix I_n Let $n \geq 2$ be a positive integer. Then $B = I_n$ is the square matrix, $n \times n$, where $b_{ij} = 1$ if $i = j$ and $b_{ij} = 0$ if $i \neq j$ If A is an $n \times m$ matrix, then $AI_m = A$ and $I_n A = A$.

Elementary Matrix An elementary matrix is a matrix which differs from the identity matrix (I_n) by one single elementary row operation

Equivalent Matrices Two matrices are equivalent if each is obtained from the other by applying a sequence of row operations

Invertible (Nonsingular) Matrix A square matrix A, $n \times n$, is said to be invertible or nonsingular if there exists a matrix $n \times n$ denoted by A^{-1} such that $AA^{-1} = A^{-1}A = I_n$

Determinant of a Matrix The determinant is a value associated with a square matrix. It can be computed from the entries of the matrix by a specific arithmetic expression, while other ways to determine its value exist as well. Determinants occur throughout mathematics. The use of determinants in calculus includes the Jacobian determinant in the substitution rule for integrals of functions of several variables. They are used to define the characteristic polynomial of a matrix that is an essential tool in eigenvalue problems in linear algebra

R. Alhajj, J. Rokne (eds.), *Encyclopedia of Social Network Analysis and Mining*, DOI 10.1007/978-1-4614-6170-8,

Cramer Rule Cramer's rule is an explicit formula for the solution of a system of linear equations with as many equations as unknowns, valid whenever the system has a unique solution. It expresses the solution in terms of the determinants of the (square) coefficient matrix and of matrices obtained from it by replacing one column by the constant column of right hand sides of the equations. It is named after Gabriel Cramer (1704–1752)

Definition

In this entry, we describe all basic matrix operations: Addition, subtraction, and multiplication. We show the importance of matrices in studying system of linear equations (augmented matrix and row operations). We show different methods used in calculating determinant of a square matrix. We show the importance of determinant in solving system of linear equations (Cramer rule) and in finding the inverse of a matrix (Adjoint method).

Introduction

Graphs are very useful ways of presenting information about social networks. However, when there are many actors and/or many kinds of relations, they can become so visually complicated that it is very difficult to see patterns. It is also possible to represent information about social networks in the form of matrices. Representing the information in this way also allows the application of mathematical and computer tools to summarize and find patterns. Social network analysts use matrices in a number of different ways. So, understanding a few basic things about matrices from mathematics is necessary. For example, the simplest and most common matrix is binary. That is, if a tie is present, a one is entered in a cell; if there is no tie, a zero is entered. This kind of a matrix is the starting point for almost all network analysis and is called an "adjacency matrix" because it represents who is next to or adjacent to whom in the "social space" mapped by the relations that we have measured. The following is an example of a binary matrix:

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Matrices and linear algebra are surely inseparable subjects, and they are important "concepts" needed in many aspects of real life science. The subject of linear algebra can be partially explained by the meaning of the two terms comprising the title. We can understand "linear" to mean anything that is "straight" or "flat". For example, in the xy-plane we are accustomed to describing straight lines as the set of solutions to an equation of the form $y = mx + b$, where the slope m and the y-intercept b are constants that together describe the line. Living in three dimensions, with coordinates described by triples (x, y, z), they can be described as the set of solutions to equations of the form $ax + by + cz = d$, where a, b, c, d are constants that together determine the plane. While we might describe planes as "flat", lines in three dimensions might be described as "straight". From a multivariate calculus course, we recall that lines are sets of points described by equations such as $x = 3t - 4, y = -7t + 2, z = 9t$, where t is a parameter that can take on any value.

Another view of this notion of "flatness" is to recognize that the sets of points just described are solutions to equations of a relatively simple form. These equations involve addition and multiplication only. Here are some examples of typical equations:

$$2x + 3y - 4z = 134$$
$$x_1 + 5x_2 - x_3 + x_4 + x_5 = 0$$
$$9a - 2b + 7c + 2d = -7$$

What we will not see in a linear algebra course are equations like:

$$xy + 5yz = 13\, x_1 + x_2^3/x_4 - x_3x_4x_5^2 = 0$$

$$\cos(ab) + log(c - d) = -2$$

A system of linear equations in several unknowns is naturally represented using the formalism of matrices.

The word "algebra" is used frequently in mathematical preparation courses. Most likely, we have spent a good 10–15 years learning the algebra of the real numbers, along with some introduction to the very similar algebra of complex numbers. However, there are many new algebras to learn and use, and likely, linear algebra and matrix operations will be our second algebra. Like learning a second language, the necessary adjustments can be challenging at times, but the rewards are many. And it will make learning our third and fourth algebras even easier. Perhaps, "groups" and "rings" are excellent examples of other algebras with very interesting properties and applications.

The brief discussion above about lines and planes suggests that linear algebra has an inherently geometric nature, and this is true. Examples in two and three dimensions can be used to provide valuable insight into important concepts of this subject.

The material presented here can be found in every textbook on basic linear algebra. Since there are so many textbooks on basic linear algebra, and we cannot list all of them, we refer to a few books here. For example, Axler (1997), Bernstein (2005), Beezer (2004), Blyth and Robertson (2002), Kaw (2011), Lang (1986), Lay (2003), Robbiano (2011), and Shores (2007).

System of Linear Equations

Matrices play an important role in solving a system of linear equations as we will see later on in this section. Let R be the set of all real numbers and C be the set of all complex numbers. Then $R^n = \{(a_1, a_2, \ldots, a_n) \mid a_1, a_2, \ldots, a_n \in R\}$ and $C^n = \{(a_1, a_2, \ldots, a_n) \mid a_1, a_2, \ldots, a_n \in C\}$. An element of R^n (C^n) is called a point.

A system of linear equations is a collection of m equations with n variable $x_1, x_2, x_3, \ldots, x_n$ of the form

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n &= b_3 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n &= b_m
\end{aligned} \qquad (1)$$

where $a_{ij}, b_j \in R\ (\in C)$.

A point $(a_1, \ldots, a_n) \in R^n$ $(\in C^n)$ is said to be a solution to a system of linear equations with n variables, $x_1, x_2, x_3, \ldots, x_n$ as in (1) if we substitute a_1 for x_1, a_2 for x_2, a_3 for x_3, …, a_n for x_n; then, for every equation of the system the left side will equal the right side, i.e., each equation is true simultaneously.

Let F be a set. Then $|F|$ denotes the cardinality of the set F, i.e., the number of the elements in F.

Theorem 1 *Let $F \subseteq R^n (\subseteq C^n)$ be the set of all solutions to a system of linear equations with n variables. Then either $|F| = 1$ (i.e., the system has unique solution) or F is an empty set (i.e., the system has no solution) or $|F| = \infty$ (i.e., the system has infinitely many solutions).*

Example 1 $F = \{(2, 1)\}$ is the set of all solutions to the system $2x_1 + x_2 = 5$, $x_1 + 2x_2 = 4$ (i.e., $x_1 = 2, x_2 = 1$, and hence, the solution is unique).

$F = \{(a_1, 2a_1 + 3) \mid a_1 \in R\}$ is the set of all solutions to the system $-2x_1 + x_2 = 3$, $-4x_1 + 2x_2 = 6$ (i.e., the system has infinitely many solutions).

The system $x_1 + 2x_2 = 0, 2x_1 + 4x_2 = 1$ has no solutions.

A system of linear equations is called **consistent** if it has a solution, and it is called **inconsistent** if it has no solutions. A system of linear equations as in (1) is called **homogeneous** if $b_1 = b_2 = \cdots = b_m = 0$.

Theorem 2 *Every homogeneous system of linear equations is consistent (i.e., (0, 0, …, 0) is always a solution of such system). If a system of linear equations is consistent and it has more variables than equations, then the system has infinitely* *many solutions. In particular, if a homogenous system has more variables than equations, then it has infinitely many solutions.*

Two systems of linear equations are **equivalent** if their solution sets are equal.

Theorem 3 *If we apply one or two or all of the following equation operations to a system of linear equations as many times as we want, then the original system and the transformed system are equivalent.*

1. *Swap the locations of two equations in the list of equations.*
2. *Multiply each term of an equation by a nonzero quantity.*
3. *Multiply each term of one equation by some quantity and add these terms to a second equation, on both sides of the equality. Leave the first equation the same after this operation but replace the second equation by the new one.*

In light of Theorem 3, one can view each equation of a system of linear equations as a row of a matrix and each equation operation in Theorem 3 as a row operation on a matrix. Hence, we have the following well-known **row operations**.

Let A be an $m \times n$ matrix (i.e., A has m rows and n columns). Then each of the following is called a row operation on A.

1. Swap the locations of two rows.
2. Multiply each entry of a single row by a nonzero quantity.
3. Multiply each entry of one row by some quantity and add these values to the entries in the same columns of a second row. Leave the first row the same after this operation but replace the second row by the new values.

We will use the following notations to describe the row operations stated above:

1. $R_i \leftrightarrow R_j$: Swap the location of rows i and j.
2. αR_i: Multiply row i by the nonzero scalar α.
3. $\alpha R_i + R_j \to R_j$: Multiply row i by the scalar α and add to row j, so that row j will change but no change in row i.

Two matrices, A, B, of the same size , say $m \times n$, are said to be **row-equivalent** if and only if A is obtained from B by applying a sequence of row operations on B.

The following type of matrices is needed in order to achieve our main goal and solve a system of linear equations using augmented matrices.

A matrix A, $m \times n$, is called in **reduced row-echelon form** if it meets all of the following conditions:

1. If there is a row where every entry is zero, then this row lies below any other row that contains a nonzero entry.
2. The leftmost nonzero entry of a row is equal to 1.
3. The leftmost nonzero entry of a row is the only nonzero entry in its column.
4. Consider any two different leftmost nonzero entries, one located in row i, column j and the other located in row s, column t. If $s > i$, then $t > j$.

A matrix A, $m \times n$, is said to be in **row-echelon form** if and only if A satisfies conditions (1), (2), and (4) as above.

Example 2 The matrix $A = \begin{bmatrix} 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ is in reduced row-echelon form.
The matrix $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ is in row-echelon form but not in reduced row-echelon form.
The matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is neither in row-echelon form nor in reduced row-echelon form.

Theorem 4 *Let A be a matrix, $m \times n$. Then A is row-equivalent to a unique matrix, $m \times n$, in reduced row-echelon form.*

Consider the system in (1). The augmented matrix of the system is

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & |b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & |b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & |b_m \end{bmatrix}$$

where $a_{1i}, a_{2i}, \ldots, a_{mi}$ are the coefficients of the variable x_i in the system (1). In view of Theorem 3, we have the following result.

Theorem 5 *Let A be the augmented matrix of a given system of linear equations and let E be the reduced row-echelon form of A. Then the solution set of the given system is equal to the solution set of the system that has E as its augmented matrix.*

Theorem 6 *Let A be the augmented matrix of a given system of linear equations and let E be the reduced row-echelon form of A. Then the given system is consistent if and only if none of the equations that correspond to the matrix E has a form zero = nonzero.*

Suppose A is the augmented matrix of a consistent system of linear equations and let E be the reduced row-echelon form of A. Suppose j is the index of a column of B that contains the leading 1 for some row. Then the variable x_j is said to be **dependent or leading**. A variable that is not dependent is called **independent or free**. If x_k is an independent variable, we understand that x_k can take any real (complex) value.

Consider the following system:

$$x_1 + 2x_2 - x_3 + 2x_4 = 1$$

$$-2x_1 - 4x_2 + 3x_3 - 4x_4 = 2$$

$$-x_1 - 2x_2 + x_3 - 2x_4 = -1$$

Whose augmented matrix is equivalent to the matrix

$$E = \begin{bmatrix} 1 & 2 & 0 & 2 & |5 \\ 0 & 0 & 1 & 0 & |4 \\ 0 & 0 & 0 & 0 & |0 \end{bmatrix}$$

in reduced row-echelon form, which corresponds to the system:

$$x_1 + 2x_2 + 2x_4 = 5$$

$$x_3 = 4$$

$$0 = 0$$

No equation of this system has a form *zero* = *nonzero*. Therefore, the original system is consistent. The dependent (leading) variables are x_1 and x_3. The independent (free) variables are x_2 and x_4. Thus, x_2, x_4 can take any real (complex) values. We write the leading variables in terms of the dependent variables. Thus, we have:

$$x_1 = 5 - 2x_2 - 2x_4$$

$$x_3 = 4$$

$$x_2, x_4 \in R(C)$$

Hence, the solution set of the original system is

$$F = \{(5 - 2x_2 - 2x_4, x_2, 4, x_4) | x_2, x_4 \in R(C)\}.$$

Matrix Operations

Let $M_{n\times m}$ be the set of all $n \times m$ matrices with entries from R (C) for some positive integers n, m. If $A \in M_{n\times m}$, then a_{ij} denotes the entry in the matrix A that is located in the ith row and the jth column of A. If $A \in M_{n\times m}$, then we say that $size(A) = n \times m$.

Theorem 7 (Addition, subtraction, and multiplication by a scalar) *Let A, B be two matrices and $\alpha \in R(C)$. Then $A + B$ and $A - B$ is defined if and only if size(A) = size(B). Furthermore, suppose that size(A) = size(B), $A + B = C$, $A - B = D$, and $\alpha A = F$. Then $c_{ij} = a_{ij} + b_{ij}$, $d_{ij} = a_{ij} - b_{ij}$, and $f_{ij} = \alpha a_{ij}$.*

Example 3 Let $A = \begin{bmatrix} 3 & 4 & 1 \\ -1 & 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} -2 & 0 & 1 \\ 4 & -2 & 1 \end{bmatrix}$, and $\alpha = -2$. Then $A + B = C = \begin{bmatrix} 1 & 4 & 2 \\ 3 & -2 & 3 \end{bmatrix}$, $A - B = D = \begin{bmatrix} 5 & 4 & 0 \\ -5 & 2 & 1 \end{bmatrix}$, and $-2A = \begin{bmatrix} -6 & -8 & -2 \\ 2 & 0 & -4 \end{bmatrix}$.

For a matrix $A \in M_{n\times m}$, let A_{r_i} denote the ith row of A and let A_{c_i} denote the ith column of A.

Theorem 8 (Matrix multiplication) *Let A be an $m \times n$ matrix and B be a $v \times k$ matrix. Then the matrix multiplication AB is defined if and only if $n = v$. Furthermore, suppose that $n = v$ and let $AB = D$. Then the following statements hold:*

1. *Size(D) = $m \times k$ and $d_{ij} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots + a_{i,n}b_{n,j}$ [dot product].*
2. *$D_{c_i} = b_{1i}A_{c_1} + b_{2i}A_{c_2} + \cdots + b_{ni}A_{c_n}$ [linear combination of the columns of A].*

3. $D_{r_i} = a_{i1}B_{r_1} + a_{i2}B_{r_2} + \cdots + a_{in}B_{r_n}$ *[linear combination of the rows of B].*

Example 4 Let $A = \begin{bmatrix} 1 & 2 & 5 \\ -3 & 0 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 5 & 2 \\ -3 & 7 & -4 \end{bmatrix}$. Then AB is defined by Theorem 8. Let $AB = D$. Then

1. Size$(D) = 2 \times 3$ and $d_{23} = (-3)(3) + (0)(2) + (6)(-4)$.
2. The second column of D is $D_{c_2} = -2\begin{bmatrix} 1 \\ -3 \end{bmatrix} + 5\begin{bmatrix} 2 \\ 0 \end{bmatrix} + 7\begin{bmatrix} 5 \\ 6 \end{bmatrix}$.
3. The second row of D is $D_{r_2} = -3[1 \;\; -2 \;\; 3] + 0[0 \;\; 5 \;\; 2] + 6[-3 \;\; 7 \;\; -4]$.

Note that BA is undefined by Theorem 8. In fact, it is possible that for some matrices A, B that AB and BA are defined but $AB \neq BA$.

Theorem 9 *Let i be a positive integer and A be a matrix. Then $A^i = A \times A \times \cdots \times A$ (i times) is defined if and only if A is a square matrix, i.e., $A \in M_{n\times n}$ for some positive integer n.*

Theorem 10 *1. Let $\alpha, \beta \in R(C)$, $A, B \in M_{n\times m}$ and let $C \in M_{m\times k}$. Then $(\alpha A + \beta B)C = \alpha AC + \beta BC$.*
2. Let $\alpha, \beta \in R(C)$, $A, B \in M_{n\times m}$ and let $C \in M_{k\times n}$. Then $C(\alpha A + \beta B)C = \alpha CA + \beta CB$.
3. Let $A \in M_{n\times m}$, $B \in M_{m\times k}$, and $C \in M_{k\times i}$. Then $ABC = (AB)C = A(BC)$.

Theorem 11 (Matrix-form of a system of linear equations) *Consider the system in (1). Let $C = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, and $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$. Then (in view of Theorem 8(2)), the matrix-form of the system in (1) is $CX = B$, where C is called the coefficient matrix of the system, X is called the variables-column of the system, and B is called the constant column of the system.*

Theorem 12 *Let $CX = B$ be the matrix-form of a given system of linear equations with m equations and n variables (hence, size$(C) = m \times n$, size$(X) = n \times 1$, and size$(B) = m \times 1$). Then the given system is consistent if and only if there are some real (complex) numbers, $r_1, r_2, \ldots, r_n$, such that $B = r_1C_{c_1} + r_2C_{c_2} + \cdots + r_nC_{c_n}$, i.e., B is a linear combination of the columns of C.*

Example 5 Consider the system:

$$x_1 + 2x_3 - x_3 = -1$$

$$3x_1 + 5x_2 + 2x_3 = 7$$

$$-x_1 + 2x_2 - 6x_3 = -13$$

Then $C = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 5 & 2 \\ -1 & 2 & -6 \end{bmatrix}$ is the coefficient matrix, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is the variables-column, and $B = \begin{bmatrix} -1 \\ 7 \\ -13 \end{bmatrix}$ is the constant column. Thus, the matrix-form of the system is $CX = B$.

Since $B = 1\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} + 0\begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix} + 2\begin{bmatrix} -1 \\ 2 \\ -6 \end{bmatrix}$ is a linear combination of the columns of C, we conclude that the system is consistent by Theorem 12, and the point (1, 0, 2) is in the solution set of the system (i.e., $x_1 = 1, x_2 = 0, x_3 = 2$ is a solution to the system).

Let $n \geq 1$ be a positive integer. Then I_n is an $n \times n$ matrix where $i_{11} = i_{22} = \cdots = i_{nn} = 1$ and $i_{kj} = 0$ if $k \neq j$. We call I_n an **identity matrix**. Note that if $n = 1$, then $I_n = 1$.

Theorem 13 *Let A be a $k \times m$ matrix. Then $AI_m = A$ and $I_kA = A$.*

Theorem 14 (Row operations and matrix multiplication) *Let A be an $n \times m$ matrix and let W be a row operation. Assume that we applied W exactly once on A and we obtained the matrix B, also assume we applied W on the matrix I_n exactly once and we obtained the matrix E. Then $EA = B$.*

A matrix that is obtained from I_n by applying one row operation on I_n exactly once is called an **elementary matrix**.

Example 6 Let A be a 2×5 matrix and assume we applied a sequence of row operations on A, and we obtained the matrix B as below:

$$A \ \overline{R_1 \leftrightarrow R_2} \ A_1 \ \overline{2R_2 + R_1 \to R_1} \ A_2 \ \overline{3R_1} \ B.$$

Since we performed exactly three row operations on A, we should be able to find three elementary matrices, E_1, E_2, E_3, such that $E_3E_2E_1A = B$.

$$I_2 \ \overline{R_1 \leftrightarrow R_2} \ E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = E_1$$

$$I_2 \ \overline{2R_2 + R_1 \to R_1} \ E_2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$I_2 \ \overline{3R_1} \ E_3 = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

Hence, $E_3E_2E_1A = B$.

Let A be a square matrix (i.e., $A \in M_{n \times n}$). We say A is **nonsingular or invertible** if there exists a matrix denoted by A^{-1} such that $AA^{-1} = A^{-1}A = I_n$. If we cannot find a matrix B such that $AB = BA = I_n$, then we say A is **singular or non-invertible**.

Theorem 15 *Let $A \in M_{n \times n}$ and suppose that A is invertible. Then A^{-1} is unique. Furthermore, suppose that $BA = I_n$ for some matrix B. Then $BA = AB = I_n$, and hence $B = A^{-1}$.*

Theorem 16 *Let $A \in M_{n \times n}$ and suppose that B is the reduced row echelon form of A. Then A is invertible if and only if $B = I_n$.*

Theorem 17 (Calculating A^{-1}) *Let $A \in M_{n \times m}$ and suppose that we joint I_n to the matrix A, and we formed a new matrix denoted by $[A|I_n]$. Then*

1. *Suppose that we applied a sequence of row operations on the matrix $[A|I_n]$, and we obtained the matrix $[D|F]$ (i.e., A is row-equivalent to D and I_n is row-equivalent to F). Then $FA = D$.*
2. *Suppose A is a square matrix (i.e., $n = m$), and we applied a sequence of row operations on the matrix $[A|I_n]$, and we obtained the matrix $[D|F]$ where D is the reduced row-echelon form of A. If $D = I_n$, then $F = A^{-1}$.*

Example 7 Let $A = \begin{bmatrix} 4 & 3 \\ 1 & 1 \end{bmatrix}$. We use Theorem 29(2) in order to find A^{-1}. We form the matrix $[A|I_2]$ and we apply a sequence of row operations on the matrix $[A|I_2]$ in order to obtain the matrix $[D|F]$ where D is the reduced row operation form of A. We see that $D = I_2$ and $F = \begin{bmatrix} 1 & -3 \\ -1 & 4 \end{bmatrix}$. Thus, $F = A^{-1}$ by Theorem 29(2).

Theorem 18 *1. Let $A, B \in M_{n \times n}$ be invertible matrices. Then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*
2. Let $A \in M_{n \times n}$ be invertible and $\alpha \in R(C)$ such that $\alpha \neq 0$. Then αA is invertible and $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$.

Theorem 19 *Let $CX = B$ be the matrix-form of a given system of linear equations with n equations and n variables. Then the system has a unique solution if and only if C is invertible. Furthermore, if C^{-1} is the inverse of C, then $X = C^{-1}B$.*

Example 8 Consider the following system in matrix-form $CX = B$, where $C = \begin{bmatrix} 4 & 3 \\ 1 & 1 \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and $B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Since C is invertible by Example 7, we have $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = C^{-1}B = \begin{bmatrix} 1 & -3 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Thus, $\{(-1, 2)\}$ is the solution set of the system.

Let A be an $n \times m$ matrix. The **transpose of A** is denoted by A^T where $a^T_{ij} = a_{ji}$, and hence $size(A^T) = m \times n$.

A square matrix A, $n \times n$, is called **symmetric** if $A^T = A$, and it is called **skew-symmetric** if $A^T = -A$.

Theorem 20 *1. Assume that $A, B \in M_{n\times m}$ and $\alpha, \beta \in R(C)$. Then $(\alpha A + \beta B)^T = \alpha A^T + \beta B^T$.*

2. Assume that AB is defined for some matrices A, B. Then $(AB)^T = B^T A^T$.

3. Assume A is a square matrix and $\alpha \in R(C)$. Then $\alpha(A + A^T)$ is symmetric and $\alpha(A - A^T)$ is skew-symmetric.

4. Assume A is a square matrix, then $A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$, i.e., A is a sum of a symmetric matrix and a skew-symmetric matrix.

Theorem 21 *Let $A \in M_{n\times n}$. Then A is invertible if and only if A^T is invertible. Furthermore, if A is invertible, then $(A^T)^{-1} = (A^{-1})^T$.*

Determinant and Cramer Rule

Let $A \in M_{n\times n}$ $(n \geq 2)$. Then A_{ij} denotes the matrix obtained from A after deleting the ith row and the jth column of A. Some authors called such matrix a **minor of** A. Note that $size(A_{ij}) = (n-1) \times (n-1)$. If B is a square matrix, then $det(B)$ or $|B|$ denotes the determinant of B.

Theorem 22 *Let A be a 2×2 matrix, say $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Then $det(A) = a_{11}a_{22} - a_{12}a_{21}$.*

Theorem 23 (Calculating $det(A)$) *Let $A \in M_{n\times n}$. Then*

1. Assume that we selected the ith row of A. Then

$$\begin{aligned} det(A) &= (-1)^{i+1} a_{i1} det(A_{i1}) \\ &\quad + (-1)^{i+2} a_{i2} det(A_{i2}) \\ &\quad + \cdots + (-1)^{i+n} a_{in} det(A_{in}) \end{aligned}$$

2. Assume that we selected the jth column of A. Then

$$\begin{aligned} det(A) &= (-1)^{j+1} a_{1j} det(A_{1j}) \\ &\quad + (-1)^{j+2} a_{2j} det(A_{2j}) \\ &\quad + \cdots + (-1)^{j+n} a_{nj} det(A_{nj}) \end{aligned}$$

3. $det(A)$ is unique and it does not rely on the row or the column we select, and thus, it is always recommended that we select a row or a column of A that has more zeros in order to calculate $det(A)$.

Example 9 Let $A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & 5 & 0 \\ 3 & 2 & 0 \end{bmatrix}$. To find $det(A)$, we observe that the 3rd column of A has more zeros. Hence, by Theorem 23(2) , we have

$$det(A) = (-1)^{3+1} 3 det \begin{bmatrix} -4 & 5 \\ 3 & 2 \end{bmatrix} = -69$$

Let $A \in M_{n\times n}$. If $a_{ij} = 0$ whenever $i > j$, then A is called an **upper triangular matrix**. If $a_{ij} = 0$ whenever $i < j$, then A is called a **lower triangular matrix**. If $a_{ij} = 0$ whenever $i \neq j$, then A is called a **diagonal matrix**. If A is upper triangular or lower triangular or diagonal, then A is said to be a **triangular matrix**.

Theorem 24 *Let $A \in M_{n\times n}$ be a triangular matrix. Then $det(A) = a_{11}a_{22}.\cdots.a_{n-1n-1}a_{nn}$.*

Theorem 25 *1. Let $A \in M_{n\times n}$ be an invertible upper triangular matrix. Then A^{-1} is an upper triangular matrix.*

2. Let $A \in M_{n\times n}$ be an invertible lower triangular matrix. Then A^{-1} is a lower triangular matrix.

3. Let $A \in M_{n\times n}$ be an invertible diagonal matrix. Then A^{-1} is a diagonal matrix.

Theorem 26 *Let $A \in M_{n\times n}$. If one of the following statements hold, then $det(A) = 0$.*

1. Two rows or two columns of A are identical.

2. One row of A is a multiple of another row of A, or one column of A is a multiple of another column of A.

3. One row or one column of A is entirely zeros.

Theorem 27 (The effect of row operations on determinant) *Let $A \in M_{n\times n}$. Then,*

1. Suppose that we applied row operation number one exactly once on A: $A \ \overline{R_i \leftrightarrow R_j} \ B$. Then $det(B) = -det(A)$.

2. *Suppose that we applied row operation number two exactly once on A:* $A \xrightarrow{\alpha R_i} B$. *Then* $det(B) = \alpha det(A)$.
3. *Suppose that we applied row operation number three exactly once on A:* $A \xrightarrow{\alpha R_i + R_j \to R_j} B$. *Then* $det(B) = det(A)$.

Theorem 28 (Most used method to find a determinant) *Let* $A \in M_{n\times n}$. *It is always recommended that we apply row operations on A in order to transform A to a triangular matrix, then we use Theorem 27 and Theorem 24 to calculate* $det(A)$.

Example 10 Let $A = \begin{bmatrix} 2 & 1 & 2 & 1 \\ -2 & 1 & 2 & 6 \\ 4 & 2 & 5 & 1 \\ -2 & -1 & -2 & 3 \end{bmatrix}$. We use Theorem 28 in order to calculate $det(A)$.

$$A \xrightarrow{R_1 + R_2 \to R_2,\ \ -2R_1 + R_3 \to R_3,\ \ R_1 + R_4 \to R_4} B = \begin{bmatrix} 2 & 1 & 2 & 1 \\ 0 & 2 & 4 & 7 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Since B is obtained from A by applying row operation number three exactly three times on A and row operation number three has no effect on $det(A)$, we conclude that $det(A) = det(B)$ by Theorem 27(3). Hence, $det(A) = det(B) = 8$ by Theorem 24.

Example 11 Let $A \in M_{4\times4}$ and suppose that $A \xrightarrow{R_1 \leftrightarrow R_2} B \xrightarrow{2R_3} C = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 0 & 1 & 1 & -4 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 4 \end{bmatrix}$. By Theorem 24, $det(C) = 16$. By Theorem 27(1), $16 = det(C) = 2det(B)$, and hence $det(B) = 8$. By Theorem 27(2), $8 = det(B) = -det(A)$, and hence $det(A) = -8$.

Theorem 29 (Characterizing invertible matrices in terms of determinant) *Let* $A \in M_{n\times n}$. *Then A is invertible (nonsingular) if and only if* $det(A) \neq 0$.

Theorem 30 (Characterizing consistent and inconsistent systems of linear equations in terms of determinant) *Let* $CX = B$ *be the matrix-form of a system of linear equations with n equations and n variables (i.e.,* $size(C) = n \times n$*). Then:*

1. *The system has a unique solution if and only if* $det(C) \neq 0$.
2. *Assume that the given system is consistent. Then the system has infinitely many solutions if and only if* $det(C) = 0$.

Theorem 31 1. *Let* $A, B \in M_{n\times n}$. *Then* $det(AB) = det(A)det(B)$, *and it is not always true that* $det(A + B) = det(A) + det(B)$.
2. *Let* $A \in M_{n\times n}$. *Then* $det(A^T) = det(A)$.
3. *Let* $A \in M_{n\times n}$ *be an invertible matrix. Then* $det(A^{-1}) = \frac{1}{det(A)}$.
4. *Let* $A \in M_{n\times n}$ *and* $\alpha \in R(C)$. *Then* $det(\alpha A) = \alpha^n det(A)$.

Theorem 32 (Adjoint method: calculating A^{-1} using determinant) *Let* $A \in M_{n\times n}$ *be an invertible matrix. Then the (i, j)-entry of* $A^{-1} =$

$$a_{ij}^{-1} = \frac{(-1)^{i+j} det(A_{ji})}{det(A)}$$

(Recall: A_{ji} *is the matrix obtained from A after deleting the* j*th row and* i*th column of A.)*

Example 12 Let $A = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix}$. Then $det(A) = a_{11}a_{22}a_{33} = (3)(2)(1) = 6$ by Theorem 24. Hence, A is invertible by Theorem 29. Thus, the (2, 3)-entry of $A^{-1} = a_{23}^{-1} = \frac{(-1)^5 det(A_{32})}{det(A)}$

by Theorem 32. Now, $A_{32} = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ and $det(A_{32}) = 9$. Thus, $a_{23}^{-1} = \frac{-9}{6} = \frac{-3}{2}$.

Let $CX = B$ be the matrix-form of a system of linear equations with n equations and n variables, say $x_1, x_2, \ldots, x_n$. Then C_i indicates the matrix obtained from C after replacing the ith column of C by the constant column B.

Theorem 33 (Cramer rule: solving $n \times n$ system of linear equations) *Let $CX = B$ be the matrix-form of a system of linear equations with n equations and n variables, say $x_1, x_2, \ldots, x_n$ and suppose that $det(C) \neq 0$ (and hence, the system has a unique solution by Theorem 30). Then $x_i = \dfrac{det(C_i)}{det(C)}$.*

Example 13 Consider the system in the matrix-form $CX = B$, where $C = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ -2 & -4 & 2 \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $B = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$. Then $det(C) = 12$, $C_1 = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ -2 & -4 & 2 \end{bmatrix}$, $C_2 = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 2 & 2 \\ -2 & -2 & 2 \end{bmatrix}$, $C_3 = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ -2 & -4 & -2 \end{bmatrix}$. Thus, by Theorem 33 we have: $x_1 = det(C_1)/det(C) = -12/12 = 1$, $x_2 = det(C_2)/det(C) = 12/12 = 1$, and $x_3 = det(C_3)/det(C) = 0/12 = 0$.

Conclusions

In this entry, we explicitly explained and stated all major results on basic matrix operations. We illustrated all different methods used in solving system of linear equations using matrices. The concept of elementary matrices and their strong relation with row operations is explained in details. Major results on determinant and its use are illustrated clearly in this article.

Cross-References

- ▸ Eigenvalues, Singular Value Decomposition
- ▸ Least Squares
- ▸ Matrix Analysis of Networks
- ▸ Matrix Decomposition

References

Axler S (1997) Linear algebra done right. Springer, Berlin

Beezer RA (2012) A first course in linear algebra. Congruent Press, Tacoma

Bernstein DS (2005) Matrix mathematics: theory, facts, and formulas with application to linear systems theory. Princeton University Press, Princeton

Blyth TS, Robertson EF (2002) Basic linear algebra. Springer, Berlin

Kaw A (2011) Introduction to matrix algebra, 2nd edn. University of South Florida, Tampa

Lang S (1986) Introduction to linear algebra. Springer, Berlin

Lay DC (2003) Linear algebra and its applications, 3rd edn. Pearson, Toronto

Robbiano L (2011) Linear algebra for everyone. Springer, Berlin

Shores TS (2007) Applied linear algebra and matrix analysis. Springer, Berlin