Hebron University

Faculty of Graduate Studies

On Determinant Of Non-Square Matrices

By

Shireen Azmi Sabri AL-Dweak

Supervisors

Dr. Mahmoud Shalalfeh

Submitted In Partial Fulfillment of the Requirements For The Degree of Master of Mathematics, Faculty of Graduate Studies, At Hebron University At Hebron, Palestine.

2021

On Determinant Of Non-Square Matrices

By

Shireen Azmi Sabri AL-Dweak

This Thesis was defended successfully on –January – 2021 and approved by

Committee Members		**Signature**
1. **Dr. Mahmoud Shalalfeh** **(Hebron University)**	**Supervisor**	………………..
2. **Dr. Bassam Manasrah** **(Hebron University)**	**Internal Examiner**	………………..
3. **Dr. Iyad Alhriba** **(Palestine Polytechnic University)**	**External Examiner**	………………..

الإهداء

إلى من أشتاق إليه بكل جوارحي ... إلى أبي الذي فارقنا بجسده، ولكن روحه ما زالت تُرفرف في سماء حياتي.

إلى نبع المحبة والإيثار والكرم ... إلى من أمدَّتني بالنصح والإرشاد والدعاء ... أمي أطال الله في عمرها وأمدَّها بالصحة والعافية لتظل عوناً لي.

إلى من لم تبخل بوقت أو جهد لمساعدتي يوما ما .... اختي سوسن

إلى إخوتي واخواتي .... سندي وعضدي ومشاطري أفراحي وأحزاني.

إلى جميع أهلي داخل الوطن الغالي وخارجه

إلى كل من دعا لي بالخير

أهديكم خلاصة جُهدي العلمي

Acknowledgments

First of all, I thank my God for all the blessing he bestowed on me and continues to bestow on me.

I would like to express my gratitude and I thank my kind, sympathetic mother who prays and asks my God to help me. I ask my God to accept her prayer which will benefit me in this life and after death.

I would also like to express my gratitude and appreciation for my sister Sawsan Dweak for her help, patience, encouragement, supporting this project.

My sincere thanks go to my advisor Dr. Mahmoud Shalalfeh who offered me advice and assistance of every stage of my thesis including wise supervision and constant encouragement. His fruitful suggestions were very helpful. I am indebted to Dr. Mahmoud for his valuable and constructive comments which substantially improved the final product and for the time and effort he put.

I would also like to acknowledge to all my teachers in Hebron University, department of mathematics.

My thanks are extended to Dr. Bassam Manasrah and Dr. Iyad Alhribat for serving in my committee.

I thank my family members for the support and encouragement and for their effective cooperation and care in preparing this thesis.

Finally it is pleasure to record my thanks to all my friends, and to all employees in wedad naser iddeen school for their spiritual support.

Declaration

The work provided in this thesis, unless otherwise referenced, is the result of the researcher's work, and has not been submitted elsewhere for any other degree or qualification.

Student's name: Shireen dweak

Signature:

Date :

Contents

Abstract

We study the definition of the determinant of a non-square matrix, using cofactor definition and Radic definition, and we proved that they are identical by proving the uniqueness of the determinant function that satisfies the four characterizing properties of determinant function.

We also study the connection between the area of any polygon in the Cartesian plane and determinant function for $2 \times n$ matrices. We used several methods to find the mathematical isotope and prove the properties for inverse and adjoint for a matrix as well as solving systems of equations in several ways.

الملخص

قمنا بدراسة تعريف محدد المصفوفة غير المربعة باستخدام تعريف العامل المساعد وتعريف رادك ، وأثبتنا أنهما متطابقان من خلال إثبات أن كلاً منهما يحقق الشروط الاربع للدالة المحددة.

درسنا أيضًا العلاقة بين مساحة أي مضلع في المستوى الديكارتي والدالة المحددة لمصفوفات من الرتبة $2 \times n$

و أستخدمنا عدة طرق لإيجاد النظير الرياضي وإثبات خصائص معكوس المصفوفة ومضاد المصفوفة بالإضافة إلى حل أنظمة المعادلات الخطية بعدة طرق.

Preface

In the books of a linear algebra we studied the concept of matrix and its types, and we studied the definition of the determinant of the square matrix, its properties and its applications. Here we will study the definition of the determinant of the non-square matrix, its verified properties and its applications in finding the area of polygons and finding solutions to the system of linear equations.

My thesis consists of six chapters. Each chapter is divided into sections. A number like 2.1.3 indicates item (definition, theorem, corollary or lemma) number 3 in section 1 of chapter 2. Each chapter begins with a clear statement of the pertinent definition and theorems together with illustrative and descriptive material. At the end of this thesis we present a collection of references.

In chapter (1) we introduce the basic results and definitions which shall be needed in the following chapters. The topics include results about matrices and matrix operations, properties of algebraic operations on matrices, Determinants of Matrices, The Inverse of a Matrix, Cofactor Expansion, Adjoint of Matrix, Linear Systems, Reduced Row-Echelon Form, Gauss-Jordan reduction, Cramer's rule, Rank and Nullity and Vectors in the Euclidean space R^n. This chapter is absolutely fundamental. The results have been stated without proofs, for theory may be looked in any text book in linear algebra. A reader who is familiar with these topics may this chapter and refer to it only when necessary.

Chapter (2) will be devoted to give a defined determinant of a non-square matrix. we will start by introducing the define a determinant function of non-square matrix in terms of characterizing properties that we want it to have. In section (2) we define minors and cofactors. In section (3 and 4), we will study the method for finding a determinant of a non-square matrix $(m \times n,\ m \leq n)$ using cofactor expansion, also study the effect of elementary row operations on determinant. In section (5) We will study the method for finding a determinant of a non-square matrix $(m \times n,\ m \leq n)$ using Radic's definition**.** Finally, we can proof the cofactor definition and Radic definition are determinant function, and the cofactor definition and Radic definition are the same.

Chapter (3) we study Radic definition for determinant of a rectangular matrix in more detailed way. We present new identities for the determinant of a rectangular matrix. We develop some important properties of this determinant. We generalize several classical important determinant identities, and description how the determinant is affected by operation on columns, such as interchanging columns, reversing columns or decomposing a single column.

Chapter (4) we will study an application for determinants of non-square matrices in calculating the area of polygons in R^2, and proof the area of a polygon is the determinant function, the area equals the determinant **(**the cofactor definition and Radic definition).

Chapter (5) we will study existence of inverses for non-square matrices. Also, we compute an inverse of a rectangular matrix using solution of a linear system and an adjoint of matrices. In section (2) we study some important properties for inverse and adjoint of non-square matrices. In section (3) we discuss Pseudo inverse method which gives an inverse of matrices.

Chapter (6) we will discuss some results concerning the solutions of a linear system $Ax = b$ using inverses as well as the pseudo-inverse and adjoint of a rectangular $m \times n$ matrix A, and General solution theorem. In section (3) we study want to generalize this method of Cramer's for an $m < n$ system of linear equations. Finally in section (4) we study shall consider some particular cases and examples to illustrate the results of what we have done in the previous sections especially in applying pseudo inverse to some certain examples.

Chapter One

Preliminaries

This chapter contains some definitions and basic results about matrices and matrix operations, properties of algebraic operations on matrices, Determinants of Matrices**,** The Inverse of a Matrix, Cofactor Expansion, Adjoint of Matrix, Linear Systems, Reduced Row-Echelon Form, Gauss-Jordan reduction, Cramer's rule, Rank and Nullity and Vectors in the Euclidean space R^n.

1.1 Matrices and matrix operations

Definition 1.1.1 $[\,9\,, p. 11\,]$. An $m \times n$ matrix A is a rectangular array of $m\,n$ real (or complex) numbers arranged in m horizontal rows and n vertical columns:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m1} & \dots & a_{mn} \end{bmatrix} \qquad (1)$$

The i^{th} row of A is $[a_{i1} \quad a_{i2} \quad a_{i3} \quad \cdots \quad a_{in}]$ $\quad (1 \le i \le m)$,

The j^{th} column of A is $\begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{mj} \end{bmatrix}$ $\quad (1 \le j \le n)$.

We shall say that A is m by n (written as $m \times n$). If $m = n$, we say that A is a square matrix of order n, and the numbers $a_{11}, a_{22}, \dots, a_{nn}$ form the main diagonal of A. We refer to the number a_{ij} which is in the i^{th} row and j^{th} column of A, as the $i,j\;th$ element of A, or the (i,j) entry of A, and we often write the matrix as $A = (a_{ij})$.

Definition 1.1.2 $[\,9\,, p. 16\,]$. **(The Transpose of a Matrix)** If $A = (a_{ij})$ is an $m \times n$ matrix, then the $n \times m$ matrix $A^T = (a_{ij}^T)$, where $a_{ij}^T = a_{ji}$ $(\,1 \le i \le m,\; 1 \le j \le n)$, is called the transpose of A. Thus the transpose of A is obtained by interchanging the rows and columns of A. Before operations.

Definition 1.1.3 [2 , $p.$ 27]. Two matrices are defined to be equal if they have the same size and their corresponding entries are equal.

Definition 1.1.4 [9 , $p.$ 12]. **(Diagonal Matrix)** a square matrix $A = (a_{ij})$ for which every term off the main diagonal is zero, that is, $a_{ij} = 0$ for $i \neq j$, is called a diagonal matrix.

Definition 1.1.5 [2 , $p.$ 14].

a. (Matrix Addition) If $A = \left(a_{ij}\right)$ and $B = \left(b_{ij}\right)$ are $m \times n$ matrices, then the sum of A and B is the $m \times n$ matrix $C = \left(c_{ij}\right)$, defined by $c_{ij} = a_{ij} + b_{ij}$

$(1 \leq i \leq m,\ 1 \leq j \leq n)$.

That is, C is obtained by adding corresponding elements of A and B.

b. (Scalar Multiplication) If $A = (a_{ij})$ is an $m \times n$ matrix, and r is a real number, then the scalar multiple of A by r, rA, is the $m \times n$ matrix $B = \left(b_{ij}\right)$, where $b_{ij} = r\, a_{ij}$, $(1 \leq i \leq m,\ 1 \leq j \leq n)$.

That is, B is obtained by multiplying each element of A by r.

c. (Matrix Multiplication) If $A = (a_{ij})$ is an $m \times p$ matrix, and $B = (b_{ij})$ is a $p \times n$ matrix, then the product of A and B, denoted $A\,B$, is the $m \times n$ matrix $C = \left(c_{ij}\right)$, defined by

$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$, $(1 \leq i \leq m,\ 1 \leq j \leq n)$.

The following proprieties for operations on matrices will be stated without proof.

1.2 Properties of algebraic operations on matrices

Theorem 1.2.1 [9 , $p.$ 35].

Let A, B, C and D be an $m \times n$ matrices. and r and s are real numbers, then

(1) $A + B = B + A$.

(2) $A + (B + C) = (A + B) + C$.

(3) There is a unique $m \times n$ matrix O such that $A + O = A$ for any $m \times n$ matrix A. The matrix O is called the $m \times n$ additive identity or zero matrix.

(4) For each $m \times n$ matrix A, there is a unique $m \times n$ matrix D such that

$$A + D = O. \qquad (1)$$

We shall write D as $-A$, so that (1) can be written as

$$A + (-A) = O.$$

The matrix $-A$ is called the additive inverse or the negative of A.

(5) If A, B and C are of the appropriate sizes, then $A(B\,C) = (A\,B)\,C$.

(6) If A, B and C are of the appropriate sizes, then $A(B + C) = A\,B + A\,C$.

(7) If A, B and C are of the appropriate sizes, then $(A + B)C = A\,C + B\,C$.

(8) $r(s\,A) = (r\,s)A$.

(9) $(r + s)A = r\,A + s\,A$.

(10) $r(A + B) = r\,A + r\,B$.

(11) $A(r\,B) = r(A\,B) = (rA)B$.

For a proof for these properties [9].

Theorem 1.2.2 [9, p. 41]. (**Properties of Transposing a matrix**)

If r is a scalar, A and B are matrices, then

(a) $\left(A^T\right)^T = A$.

(b) $(A + B)^T = A^T + B^T$ and $(A - B)^T = A^T - B^T$.

(c) $(A\,B)^T = B^T A^T$.

(d) $(r\,A)^T = r\,A^T$.

For a proof for these properties [9].

1.3 Determinants of Matrices

Definition 1.3.1 [8]. **(Determinant function)** A determinant function assigns to each square matrix A a scalar associated to the matrix, denoted by $det(A)$ or $|A|$ such that:

(1) The determinant of an $n \times n$ identity matrix "I" is 1. $|I| = 1$.

(2) If the matrix B is identical to the matrix A except the entries in one of the rows of B are each equal to the corresponding entries of A multiplied by the same scalar c, then $|B| = c\,|A|$.

(3) If the matrices A, B and C are identical except for the entries in one row, and for that row an entry in A is found by adding the corresponding entries in B and C, then $|A| = |B| + |C|$.

(4) If the matrix B is the result of exchanging two rows of A, then the determinant of B is the negation of the determinant of A. ($|B| = -|A|$)

Theorem 1.3.2 [8]. A determinant function has the following four properties.

(a) The determinant of any matrix with an entire row of $0's$ is 0.

(b) The determinant of any matrix with two identical rows is 0.

(c) If one row of a matrix is a scalar multiple of another row, then its determinant is 0.

(d) If a scalar multiple of one row of a matrix is added to another row, then the resulting matrix has the same determinant as the original matrix.

Theorem 1.3.3 [8]. There is at most one determinant function.

Definition 1.3.4 [9 , $p.$ 92]. Let $A = (a_{ij})$ be an $n \times n$ matrix. The determinant of A (written $\det(A)$ or $|A|$) is defined by

$$\det(A) = |A| = \sum (\pm)\, a_{1j_1} a_{2j_2} \dots a_{nj_n},$$

where the summation ranges over all permutations $j_1 j_2 \dots j_n$ of the set $S = \{1,2,\dots,n\}$. The sign is taken as $+$ or $-$ according to whether the permutation $j_1 j_2 \dots j_n$ is even or odd.

Theorem 1.3.5 [9 , $p.$ 95]. **(Properties of Determinants of Matrices)**

(a) If A is a square matrix. then $\det(A) = \det(A^T)$

(b) If matrix B results from matrix A by interchanging two rows (columns) of A, then $det(B) = -det(A)$

(c) If two rows (columns) of A are equal, then $det(A) = 0$

(d) If B is obtained from A by multiplying a row (column) of A by real number c, then $det(B) = c\, det(A)$.

(e) If $B = (b_{ij})$ is obtained from $A = (a_{ij})$ by adding to each element of the r^{th} row (column) of A a corresponding element of the s^{th} row (column) $r \neq s$ of A, then $det(B) = det(A)$.

(f) The determinant of a product of two matrices is the product of their determinants, that is, $det(AB) = det(A)det(B)$.

(g) If A has a row (column) consisting of all zeros, then $det(A) = 0$

For proofs of these properties [9].

Theorem 1.3.6 [9 , $p.$ 93].

1. Let A , B and C be an $n \times n$ matrices that differ only in a single row, say the r^{th} row, and assume that the r^{th} row of C can be obtained by adding corresponding entries in the r^{th} rows of A and B. Then

$$det(C) = det(A) + det(B)$$

The same result holds for columns.

2. (Decomposing a column). If a column K in a square matrix A is a sum of two columns (eg. $K = K_1 + K_2$), then the determinant $|A|$ is a sum of tow determinants of matrices obtained from A by replacing K by K_1 and K_2 respectively.

1.4 Cofactor Expansion and Adjoint

Definition 1.4.1 [9 , $p.$ 103]. **(Minor and Cofactor)** Let $A = (a_{ij})$ be an $n \times n$ matrix. Let M_{ij} be the $(n-1) \times (n-1)$ submatrix of A obtained by deleting the i^{th} row and j^{th} column of A. The determinant $det(M_{ij})$ is called the minor of a_{ij}. The cofactor A_{ij} of a_{ij} is defined as $A_{ij} = (-1)^{i+j} det(M_{ij})$.

Theorem 1.4.2 [9 , $p.$ 104].

Let $A = (a_{ij})$ be an $n \times n$ matrix. Then for each $1 \leq i \leq n$,

$$det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$$

(expansion of $det(A)$ about the i^{th} row).

And for each $1 \leq j \leq n$,

$$det(A) = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}$$

(expansion of $det(A)$ about the j^{th} column).

Definition 1.4.3 [9 , $p.$ 108]. **(Adjoint of Matrix)** Let $A = (a_{ij})$ be an $n \times n$ matrix. Then $n \times n$ matrix $adj\, A$, called the adjoint of A, is the matrix whose $(i,j)th$ element is the cofactor A_{ji} of a_{ji}. (The transpose of the matrix of cofactors), thus

$$adj(A) = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}.$$

Theorem 1.4.4 [9 , $p.$ 108]. (**Properties of the Adjoint**)

(a) If $A = (a_{ij})$ is an $n \times n$ matrix. Then $A(adjA) = (adjA)A = det(A)I_n$

(b) If $A = (a_{ij})$ is invertible $n \times n$ matrix. then $\det(adj\ A) = \det(A)^{n-1}$

(c) If $A = (a_{ij})$ is invertible $n \times n$ matrix. then $adj(adj\ A) = (\det A)^{n-2}\ A$

(d) $adj(A\ B) = adj(B)\ adj(A)$

(e) $\left(adj(A)\right)^T = adj(A^T)$

(f) $adj(\ k\ A\ ) = k^{n-1} adj(A)$, where k is any scalar

For proofs of these properties [9].

1.5 The Inverse of a Matrix

Definition 1.5.1 [9 , $p.$ 19]. **(Inverse of a Matrix)** An $n \times n$ matrix A is called non-singular (or invertible) if there exists an $n \times n$ matrix B such that

$$A\ B = BA = I_n.$$

The matrix B is called an inverse of A. If there exists no such matrix B, then A is called singular (or noninvertible).

It is easy to show that an inverse of a matrix is unique, if it exists, and so it is legitimate to say the inverse of A and write it as A^{-1}, thus $A\ A^{-1} = A^{-1}\ A = I_n$.

Theorem 1.5.2 [9 , $p.$ 71]. (**Properties of the Inverse of a matrix**)

(a) If A is a nonsingular matrix, then A^{-1} is a nonsingular and $\left(A^{-1}\right)^{-1} = A$

(b) If A and B are nonsingular matrices, then $A\ B$ is nonsingular and

$$(A\ B)^{-1} = B^{-1}\ A^{-1}$$

(c) If A is a nonsingular matrix, then $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

(d) For any nonzero scalar k, then $(kA)^{-1} = \frac{1}{k}\ A^{-1}$

For proofs of these properties [9].

Corollary 1.5.3 [9 , $p.$100]. If A is nonsingular, then $det(A) \neq 0$ and

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)}$$

Theorem 1.5.4 [9 , $p.$100]. A square matrix A is nonsingular if and only if

$$det(A) \neq 0$$

Theorem 1.5.5 [9 , $p.$107]. If $A = \left(a_{ij}\right)$ is an $n \times n$ matrix. Then

$$\begin{aligned} a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn} = 0 \quad & for\ i \neq k \\ a_{1j}A_{1k} + a_{2j}A_{2k} + \cdots + a_{nj}A_{nk} = 0 \quad & for\, j \neq k \end{aligned}$$

Corollary 1.5.6 [2 , $p.$106] If A is an $n \times n$ matrix and $det(A) \neq 0$, then

$$A^{-1} = \frac{1}{\det(A)} adj(A) = \begin{bmatrix} \frac{A_{11}}{\det(A)} & \frac{A_{21}}{\det(A)} & \cdots & \frac{A_{n1}}{\det(A)} \\ \frac{A_{12}}{\det(A)} & \frac{A_{22}}{\det(A)} & \cdots & \frac{A_{n2}}{\det(A)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_{1n}}{\det(A)} & \frac{A_{2n}}{\det(A)} & \cdots & \frac{A_{nn}}{\det(A)} \end{bmatrix}.$$

1.6 Linear Systems

A linear system (A system of linear equations) with m-equations and n-unknowns consists of m simultaneous equations each one is an equation in n-variables. That is, it has the form:

$$\begin{matrix} a_{11}x_1 + & a_{12}x_2 & + \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 + & a_{22}x_2 & + \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 & + \cdots & + & a_{mn}x_n & = & b_m \end{matrix} \qquad (1)$$

A solution for the linear system is a sequence of n real numbers $s_1,\ s_2,\ \ldots,\ s_n$ which when substituted in the equations of the linear system all become true statements.

The principal question for this kind of systems is to find the set of solutions to this system, that is to find all n-tuple $\left(x_1, x_2, \ldots,\ x_n\right)$ that satisfy (1)

A general system of linear equations lies in one and only one of the following categories

i. The system has no solution.

ii. The system has exactly one solution.

iii. The system has infinitely many solutions.

If there are fewer equations than variables in a linear system, then the system either has no solution or it has infinitely many solutions.

Definition 1.6.1 [9 , $p.$ 3]. **(Consistent and Inconsistent)** A system of equations that has at least one solution is called consistent. Otherwise, it is called inconsistent.

The key idea in finding the solution of a linear system is to apply elementary row operations

Theorem 1.6.2 [9 , $p.$ 7]. If any finitely many operations of the following is (are) applied to the linear system (1)

1. Interchange two equations.

2. Multiply an equation by a nonzero constant.

3. Add a multiple of one equation to another.

To get a new system. Then both systems have the same solution.

The operations listed in Theorem 1.6.2 are called **elementary row operations**.

Definition 1.6.3 [9 , $p.$ 49]. **(Row equivalent)** Let A and B be two $m \times n$ matrices.

We say that A is row equivalent to B if B can be obtained by applying a finite sequence of elementary row operations to A.

Now define the following matrices:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Then

$$\boldsymbol{Ax} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \qquad \vdots \qquad \ddots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$$

The entries in the product $\boldsymbol{Ax}$ are merely the left sides of the equations in (1). Hence the linear system (1) can be written in matrix form as $\boldsymbol{Ax} = \boldsymbol{b}$.

The matrix A is called the coefficient matrix of the linear system (1), and the matrix

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right],$$

obtained by adjoining **b** to A, is called the augmented matrix of the linear system (1). The augmented matrix of (1) will be written as $[\boldsymbol{A}|\boldsymbol{b}]$. Conversely, any matrix with more than one column can be thought of as the augmented matrix of a linear system. The coefficient and augmented matrices play key roles in solving linear systems.

Definition 1.6.4 [9 , p. 59]. **(Homogenous Systems)** A system of linear equations is said to be homogenous if the constant terms are all zero. That is, the system has form $\boldsymbol{Ax} = \boldsymbol{0}$.

We note here that a homogenous system is always consistent,

in fact $x_1 = x_2 = \cdots = x_n = 0$ is a solution which is called the trivial solution.

1.7 Gaussian Elimination, Cramer's rule

Definition 1.7.1 [2 , p. 8]. **(Reduced Row-Echelon Form)** A matrix having the following properties 1, 2 and 3 (but not necessarily 4) is said to be in row echelon form (r.e.f). A matrix that satisfies all the 4 conditions is said to be in reduced row echelon form (r.r.e.f)

1- If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. (We call this a leading 1.)
2- Any row that consists entirely of zeros is placed at the bottom of the matrix.
3- In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.
4- Each column that contains a leading 1 has zeros everywhere else.

Any $m \times n$ matrix A can be transformed into a reduced row echelon form $A^{'}$ which is row equivalent to A, and this $A^{'}$ is unique .

Definition 1.7.2 [9 , $p.$ 54] **(Gauss-Jordan reduction)** The Gauss-Jordan reduction procedure for solving the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is as follows.

Step1. Form the augmented matrix $[\boldsymbol{A}|\boldsymbol{b}]$.

Step 2. Transform the augmented matrix to reduced row-echelon form by using elementary row operations.

Step 3. The linear system that corresponds to the matrix in reduced row-echelon form that has been obtained in step 2 has exactly the same solutions as the given linear system. For each nonzero row of the matrix in reduced row-echelon form, solve the corresponding equation for the unknown that corresponds to the leading entry of the row. The rows consisting entirely of zeros can be ignored, since the corresponding equation will be satisfied for any values of the unknowns.

Definition 1.7.3 [2 , $p.$ 50]. **(Elementary matrix)** An $n \times n$ matrix is called an elementary matrix if can be obtained from the $n \times n$ identity matrix I_n by preforming a single elementary row operations.

Theorem 1.7.4 [2 , $p.$ 19]. A homogeneous system of linear equations with more unknowns than equations has infinitely many solutions.

Theorem 1.7.5 [2 , $p.$ 109]. **(Cramer's rule).** If $\boldsymbol{Ax} = \boldsymbol{b}$ is a system of n linear equations in n unknowns such that $\det(A) \neq 0$, then the system has the unique solution, given as follows:

$$x_1 = \frac{\det(A_1)}{\det(A)} \quad , \quad x_2 = \frac{\det(A_2)}{\det(A)} \quad , \quad .... \quad , \quad x_n = \frac{\det(A_n)}{\det(A)}.$$

Where A_j is the matrix obtained by replacing the entries in the j^{th} column of A by the entries in the matrix $\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

1.8 Vectors in the Euclidean space R^n

In this section we state some basic definitions regarding the vector space R^n

Definition 1.8.1 $[2, p. 167]$. **(Vectors in n-space)** If n is a positive integer, then an ordered n-tuple is a sequence of n real numbers $(a_1, a_2, \ldots, a_n)$. The set of all ordered n-tuples is called n-space and is denoted by R^n and an element $(a_1, a_2, \ldots, a_n)$ is called a vector.

A vector in the plane is a 2-vector $u = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, where x_1 and y_1 are real numbers, called the components of u.

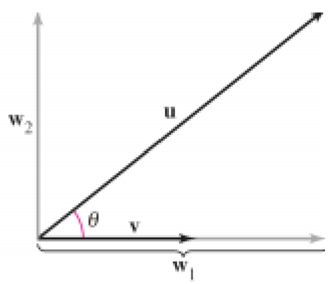
Definition 1.8.2 $[9, p. 148]$. **(Norm of a vector)** The length (also called magnitude or norm) of the vector $u = (u_1, u_2, \ldots, u_n)$ in R^n is $\|u\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$.

Definition 1.8.3 $[9, p. 148]$. **(Inner product)** If $u = (u_1, u_2, \ldots, u_n)$ and

$v = (v_1, v_2, \ldots, v_n)$ are vectors in R^n, then their dot product is defined by $u.v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$. The dot product in R^n is also called the standard inner product

Definition 1.8.4 $[2, p. 174]$. **(orthogonality)** Two vectors u and v in R^n are called orthogonal if $u.v = 0$.

Definition 1.8.5 $[2, p. 136]$. **(orthogonal projection)**



The vector u is the sum of w_1 and w_2, where w_1 is parallel to v and w_2 is perpendicular to v. the vector w_1 is called the orthogonal projection of u on v or sometimes the vector component of u along v. It is denoted by $proj_v u$.

The vector w_2 is called the vector component of u orthogonal to v. Since we have $w_2 = u - w_1$, thus the vector can be written $w_2 = u - proj_v u$.

Theorem 1.8.6 $[2, p. 136]$. If u and v are vectors in 2-space or 3-space and if $v \neq 0$, then $proj_v u = \frac{u.v}{\|v\|^2} v$ (vector component of u along v).

$u - proj_v u = u - \frac{u.v}{\|v\|^2} v$ (vector component of u orthogonal to v).

1.9 Rank and Nullity

Definition 1.9.1 $[2, p. 222]$. **(subspace)** A subset W of a vector space V is called a subspace of V if W is itself a vector space under the addition and scalar multiplication defined on V.

Definition 1.9.2 $[9, p. 207]$**. (Linear combination)** Let $v_1, v_2, \dots, v_n$ be vectors in a vector space V. A vector v in V is called a linear combination of $v_1, v_2, \dots, v_n$ if $v = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n$ for some numbers $k_1, k_2, \dots, k_n$.

Definition 1.9.3 $[2, p. 233]$**.(Linear independence)** If $S = \{v_1, v_2, \dots, v_n\}$ is a nonempty set of vectors, then the vector equation $k_1 v_1 + k_2 v_2 + \cdots + k_n v_n = 0$ has at least one solution, namely $k_1 = 0$, $k_2 = 0, \dots, k_n = 0$. If this is the only solution, then S is called a linearly independent set. If there are other solutions, then S is called a linearly dependent set.

Definition 1.9.4 $[9, p. 213]$**. (Span)** The vector $v_1, v_2, \dots, v_n$ in a vector space V are said to span V if every vector in V is a linear combination of $v_1, v_2, \dots, v_n$. Moreover, if these vectors are distinct and we denote them as a set $S = \{v_1, v_2, \dots, v_n\}$, then we also say that the set S spans V, or that $v_1, v_2, \dots, v_n$ spans V, or that V is spanned by S, or $span\, S = V$.

Definition 1.9.5 [2 , $p.$ 259]. **(row space and column space)** If A an $m \times n$ matrix, then the subspace of R^n spanned by the row vectors of A is called the row space of A, and the subspace of R^m spanned by column vectors is called the column space of A.

Definition 1.9.6 [2 , $p.$ 273]. **(Rank and nullity)** The common dimension of the row space and column space of a matrix A is called the rank of A and is denoted by $rank(A)$, the dimension of the null space which is the solution space of $\boldsymbol{Ax} = \boldsymbol{0}$ of A is called the nullity of A and is denoted by $nullity(A)$.

Theorem 1.9.7 [2 , $p.$ 275]. If A is an $n \times n$ matrix, then

(a) $rank(A)$= the number of leading variables in the solution of $\boldsymbol{Ax} = \boldsymbol{0}$.

(b)$nullity(A)$ = the number of parameters in the solution of $\boldsymbol{Ax} = \boldsymbol{0}$.

Theorem 1.9.8 [9 , $p.$ 249]. (**Rank-Nullity Theorem**) If A is an $m \times n$ matrix, Then

$$rank(A) + \text{nullity } (A) = n$$

Theorem 1.9.9[2 , $p.$ 275]. If A is any $n \times m$ matrix, then $rank(A) = rank(A^T)$

Theorem 1.9.10 [9 , $p.$ 250]. if A is an $n \times n$ matrix, then $rank(A) = n$ if and only if $det(A) \neq 0$

Definition 1.9.11 [9 , $p.$ 327] **(Linear transformation from R^n to R^m)**

A linear transformation T from R^n into R^m is a function assigning a unique vector $T(x)$ in R^m to each x in R^n such that:

(a) $T(x + y) = T(x) + T(y)$, for every x and y in R^n .

(b) $T(k\,x) = k\,T(x)$ for every x in R^n and every scalar k.

If $n = m$, the linear transformation $T: R^n \to R^n$ is called a linear operator on R^n.

The matrix $A_{m\times n}=[T(e_1):T(e_2):\dots:T(e_n)]$ is called the **standard matrix** for the linear transformation and $T:R^n\to R^m$ is a multiplication by A, that is $T_A(x)=Ax$.

Definition 1.9.12 $[2, p.395]$. **(Kernel and Range)** If $T:V\to W$ is a linear transformation, then the set of vectors in V that T maps into 0 is called the kernel of T; it is denoted by $ker(T)$. The set of all vectors in W that are images under T of at least one vector in V is called the range of T; it is denoted by $R(T)$.

Theorem 1.9.13 $[2, p.397]$. If A is an $m\times n$ matrix, and $T_A:R^n\to R^m$ is multiplication by A and define

$rank(T_A)=\dim\big(R(T_A)\big)$, nullity $(T_A)=\dim(\ker(T_A))$, then

(i) nullity $(T_A)=$ nullity (A)

(ii) $rank(T_A)=rank(A)$

Theorem 1.9.14 $[2, p.281]$. **(Invertible matrix theorem)** If A is an $n\times n$ matrix, and if $T_A:R^n\to R^n$ is multiplication by A, then the following are equivalent.

(a) A is invertible

(b) $\boldsymbol{Ax}=\mathbf{0}$ has only the trivial solution.

(c) The reduced row-echelon form of A is I_n

(d) A is expressible as a product of elementary matrices.

(e) $\boldsymbol{Ax}=\boldsymbol{b}$ is consistent for every $n\times 1$ matrix b.

(f) $\boldsymbol{Ax}=\boldsymbol{b}$ has exactly one solution for every $n\times 1$ matrix b.

(g) $\det(A)\neq 0$

(h) The range of T_A is R^n

(i) T_A is one–to–one.

(j) A has rank n.

(k) A has nullity 0.

For proofs of these properties [2].

Theorem 1.9.15 $[9\,, p.365\,]$. Let $T: R^n \to R^m$ be a linear transformation defined by $T(x) = Ax,\; x$ in $R^n,$ where A is an $m \times n$ matrix.

(1) T is one-to-one if and only if $rank(A) = n$

(2) T is onto if and only if $rank(A) = m$

Lemma 1.9.16 [10]. Consider the vector spaces R^m and R^n, and let $T_A : R^m \to R^n$ and $T_{A^T}: R^n \to R^m$ the adjoint operator, then the following statement holds

(1) $Rang(A) = R^n \leftrightarrow \exists\, \gamma > 0$ such that $\left\|A^T z\right\|_{R^m} \geq \gamma \, \|z\|_{R^n} \,, z \in R^n$.

(2) $\overline{Rang(A)} = R^n \leftrightarrow Ker\left(A^T\right) = \{0\} \leftrightarrow A^T \;\; is \;\; 1-1$ (see theorem 1.9.15)

Chapter Two

Determinant

In the history of matrices, mathematicians are interested in finding the value of the determinant for square matrices only, actually the definition of determinant and its properties are discussed only for square matrices. To break this we generalize the concept of determinant from a square matrix to a non-square matrix, and we also study their properties, methods of computation and some application.

2.1 Determinant function

There are many ways that general $m \times n$ determinants can be defined. We'll first define a determinant function in terms of characterizing properties that we want it to have. Then we'll use the construction of a determinant following the method given in the section, and through it we will prove that cofactor expansion definition of the determinant and radic definition are the same.

We generalize the idea given in [8] for the case of rectangular matrices:

Definition 2.1.1 An $m \times n$ matrix $A = (a_{ij})$ where $m \le n$ is said to have form $\tilde{I}$ if i) Entries of A are 0 or 1

ii) Eeach row have exactly one non-zero entry

That is, $\tilde{I}_{m\times n}$ consists of all matrices of the form $[A_1 \quad A_2 \quad \cdots \quad A_n]$ where

$$A_j \in \{e_1, e_2, \dots, e_m\} \cup \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \{e_1, e_2, \dots, e_m\} \text{ is the standard basis for } R^m$$

Such a matrix is called a permutation matrix in literature.

Example, $A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ is an example of an element in $\tilde{I}$

Definition 2.1.2 A determinant function assigns to each $m \times n$ $(m \le n)$ matrix A a scalar associated to the matrix, denoted $det(A)$ or $|A|$ such that

A1: The determinant of an $m \times n$ $(m \le n)$ in $\tilde{I}_{m\times n}$ is $(-1)^{p+q}$

where $p = 1 + 2 + \cdots + m$ and $q = j_1 + j_2 + \cdots + j_n$ ($j_1, j_2, \ldots, j_n$) represent columns that contain 1's.

A2: If the matrix B is identical to the matrix A except the entries in one of the rows of B are each equal to the corresponding entries of multiplied by the same scalar c, then $|B| = c|A|$.

A3: If the matrices A, B and C are identical except for the entries in one row, and for that row an entry in A is found by adding the corresponding entries in B and C, then $|A| = |B| + |C|$.

A4: If the matrix B is the result of exchanging two rows of A, then the determinant of B is the negation of the determinant of A.

These conditions are enough to characterize the determinant, but they don't show such a determinant function exists and is unique. We'll show both existence and uniqueness, but start with uniqueness. First, we'll note a couple of properties that determinant functions have that follow from the definition.

Theorem 2.1.3 A determinant function has the following four properties.

(a) The determinant of any matrix $A_{m \times n}$ ($m \leq n$) with an entire row of $0's$ is 0.

(b) The determinant of any matrix $A_{m \times n}$ ($m \leq n$) with two identical rows is 0.

(c) If one row of a matrix $A_{m \times n}$ ($m \leq n$) is a scalar multiple of another row, then its determinant is 0.

(d) If a multiple of one row of a matrix $A_{m \times n}$ ($m \leq n$) is added to another row, then the resulting matrix has the same determinant as the original matrix.

Proof: Property (a) follows from the second statement (A2) in the definition.

If A has a whole row of $0's$, then using that row and $c = 0$ in the second statement (A2) of the definition, then $B = A$.

So, $|A| = 0|B|$. Therefore, $\det(A) = 0$.

Property (b) follows from the fourth statement (A4) in the definition.

If you exchange the two identical rows, the result is the original matrix, but its determinant is negated. The only scalar which is its own negation is 0.
Therefore, the determinant of the matrix is 0.
Property (c) follows from the second statement (A2) in the definition and Property (b).
Property (d) follows from the third statement (A3) in the definition and Property (c). ■

Now we can show the uniqueness of determinant function.

Theorem 2.1.4 There is at most one determinant function.

Proof: The four properties that determinants are enough to find the value of the determinant of a matrix.

Suppose a matrix $A_{m\times n}$ $(m \leq n)$ has more than one nonzero entry in a row. Then using the third statement (A3) in definition 2.1.2.
Now, $\det(A) = \det(A_1) + \det(A_2) + \cdots + \det(A_n)$
Where A_j is the matrix that looks just like A except in that row, all the entries are 0 expect the j^{th} one which is the j^{th} entry of that row in A.
That means we can reduce the question of evaluating determinants of general matrices to evaluating determinants of matrices that have at most one nonzero entry in each row.

By Property (a) in the theorem 2.1.3, if the matrix has a row of all $0's$, its determinant is 0. Thus, we only need to consider matrices that have exactly one nonzero entry in each row.
Using the second statement (A2) in definition 2.1.2, we can further assume that the nonzero entry in that row is 1.

Now, we're down to evaluating determinants that only have entries of $0's$ and $1's$ with exactly one 1 in each row.

If two of those rows have the $1's$ in the same column, then by Property (b) in the theorem 2.1.3, that matrix has determinant 0.
Now the only matrices left to consider are matrices from $\tilde{I}$.
Using alternation, the fourth condition (A4) in definition 2.1.2, the rows can be interchanged until the $1's$ only lie on the entry a_{ij}

Finally, we're left with a matrix in $\tilde{I}$, but by the first condition (A1) in definition 2.1.2, its determinant is $(-1)^{p+q}$.
Thus, the value of the determinant of every matrix is determinant by the definition. There can be only one determinant function.

■

We need some way to construct a function with those properties, and well do that with a "cofactor construction" and " radic construction".

2.2 Minors and cofactors.

To every non-square matrix $A = (a_{ij})$ of order $m \times n$, we can associate a number (real or complex) called determinant of the non-square matrix A, where $a_{ij} = (i,j)^{th}$ element of A.

This may be thought of as a function which associates to each non-square matrix over a field F a unique number from F (real or complex). If M is the set of non-square matrices, K is the set of real numbers and $f: M \to K$ is defined by $f(A) = k,$ where $A \in M$ and $k \in K,$ then $f(A)$ is called the determinant of A. It is also denoted by $|A|$ or det(A).

The determinant can also be viewed as a function of the columns of the matrix. Let these columns be $A_1, A_2, .., A_n$, then we write the determinant as

$$|A_1, A_2, .., A_n| \quad \text{or} \quad det(A_1, A_2, .., A_n)$$

Definition 2.2.1 [3] (**Determinants of order $\mathbf{1 \times n}$**)

If $A = [a_{11} \quad a_{12} \quad a_{13} \quad \cdots \quad a_{1n}]$, then the determinant of A is

$$|A| = a_{11} - a_{12} + a_{13} - \cdots + (-1)^{1+n} a_{1n}$$

$$= \sum_{i=1}^{n} (-1)^{1+i} a_{1i}$$

Example 2.2.2 $|1 \quad 5 \quad 9| = 1 - 5 + 9 = 5$

For larger matrices, we use cofactor expansion to find the determinant of A. First of all, let’s define a few terms.

Definition 2.2.3 [3] **(Minor)** Let $A = (a_{ij})$ be an $m \times n$ matrix. for each entry a_{ij} of A, we define the minor M_{ij} of a_{ij} to be the determinant of the $(m-1) \times (n-1)$ matrix which remains when the i^{th} row and j^{th} column are deleted from A.

Example 2.2.4 Let $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$

To find M_{11}, look at element $a_{11} = 1$, delete the entries from column 1 and row 1 that corresponding to $a_{11} = 1$, see the image below.

$$\begin{bmatrix} 1 & - & - \\ - & 4 & 6 \end{bmatrix}$$

Then M_{11} is the determinant of remaining matrix, i.e.,

$$M_{11} = |4 \quad 6| = 4 - 6 = -2$$

Similarly, M_{22} can be found by looking at the element $a_{22} = 4$ and delete the same row and column where this element is found, i.e., deleting the second row, second column:

$$\begin{bmatrix} 1 & - & 5 \\ - & 4 & - \end{bmatrix}$$

Then, $M_{22} = |1 \quad 5| = 1 - 5 = -4$

It is easy to see that for the matrix A the minors of ramming elements are

$$\begin{bmatrix} - & 3 & - \\ 2 & - & 6 \end{bmatrix}$$

$M_{12} = |2 \quad 6| = 2 - 6 = -4$

$$\begin{bmatrix} - & - & 5 \\ 2 & 4 & - \end{bmatrix}$$

$M_{13} = |2 \quad 4| = 2 - 4 = -2$

$$\begin{bmatrix} - & 3 & 5 \\ 2 & - & - \end{bmatrix}$$

$M_{21} = |3 \quad 5| = 3 - 5 = -2$

$$\begin{bmatrix} 1 & 3 & - \\ - & - & 6 \end{bmatrix}$$

$M_{23} = |1 \quad 3| = 1 - 3 = -2$

Definition 2.2.5 [3] **(Cofactor).** Let $A = (a_{ij})$ be an $m \times n$ matrix. for each entry a_{ij} of A, we define the cofactor of a_{ij}, which is denoted by C_{ij} as

$$C_{ij} = (-1)^{i+j} M_{ij}$$

The matrix of cofactors of A is C with c_{ij} is cofactor of a_{ij} and the adjoint of A is $adj(A)$ is the transpose of C.

Basically, the cofactor is either M_{ij}, or $-M_{ij}$ where the sign depends on the location of the element in the matrix. For that reason, it is easier to know the pattern of cofactor instead of actually remembering the formula. If you start in the position corresponding to a_{11} with a positive sign, the sign of the cofactor has an alternating pattern. you can see this by looking at a matrix containing the sign of the cofactors:

$$\begin{bmatrix} + & - & + & \cdots & \cdots \\ - & + & - & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

The element 1 in matrix A (example 2.2.4) has place sign + and minor -2 so its cofactor is + (-2) = -2

The element 4 in matrix A (example 2.2.4) has place sign + and minor -4 so its cofactor is + (- 4) = - 4

Proceeding in this way we can find all the cofactors.

The original matrix, its matrix of minors and its matrix of cofactors are:

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}, \qquad M = \begin{bmatrix} -2 & -4 & -2 \\ -2 & -4 & -2 \end{bmatrix},$$

$$C = \begin{bmatrix} -2 & 4 & -2 \\ 2 & -4 & 2 \end{bmatrix}, \qquad adj(A) = \begin{bmatrix} -2 & 2 \\ 4 & -4 \\ -2 & 2 \end{bmatrix}$$

2.3 Determinant Using Cofactor Expansion

In this section we shall deal with matrices of size $m \times n$ where $n \leq m$ or $n \geq m$. A matrix $A_{m \times n}$ with $n \neq m$ is called a rectangular matrix. When $m \leq n$, A is said to be a horizontal matrix, otherwise A is said to be a vertical matrix.

Definition 2.3.1 [3] Let $A = (a_{ij})$ be an $m \times n$ matrix with $m \leq n$. (horizontal matrix). The determinant of A is defined as

$$det(A) = a_{11}\,C_{11} + a_{12}\,C_{12} + a_{13}\,C_{13} + \cdots + a_{1n}\,C_{1n} = \sum_{j=1}^{n} a_{1j}\,C_{1j}$$

This is called cofactor expansion along the first row.

The determinant of a vertical matrix A is defined to be the determinant of the horizontal matrix A^T.

The following theorem asserts that we can evaluate the determinant of a larger horizontal matrix by selecting any row, multiplying each element in that row by its corresponding cofactor, and summing the result, a result which is true in the case of square matrices (Theorem 1.4.2).

The following theorem is from [7] but we give here another proof

Theorem 2.3.2 (**The cofactor expansion theorem**)

Let $A = (a_{ij})$ be an $m \times n$ matrix. If $m \leq n$, the determinant of A is

$$det(A) = a_{i1}\,C_{i1} + a_{i2}\,C_{i2} + a_{i3}\,C_{i3} + \cdots + a_{in}\,C_{in}$$

$$= \sum_{j=1}^{n} a_{ij}\,C_{ij}$$

This is called the determinant using cofactor expansion along the i^{th} row.

The proof of this theorem will be given after proving Theorem 2.4.1

Example 2.3.3. Evaluate the determinant of $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$ by cofactor expansion

i) along the first row

ii) along the second row

Solution:

$$det(A) = \begin{vmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{vmatrix} = 1\,|4 \quad 6| - 3\,|2 \quad 6| + 5\,|2 \quad 4|$$

$$= 1\,(-2) - 3\,(-4) + 5(-2) = 0$$

$$det(A) = \begin{vmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{vmatrix} = -2\,|3 \quad 5| + 4\,|1 \quad 5| - 6\,|1 \quad 3|$$

$$= -2\,(-2) + 4\,(-4) - 6(-2) = 0$$

Note that : when A is an $m \times n$ matrix with $m \geq 3$ or $n \geq 3,$ the cofactors C_{ij} are determinants of $(m-1) \times (n-1)$ matrices. To compute these determinants, we apply cofactor expansion again, and obtain determinants of $(m-2) \times n-2)$ matrices. We keep applying cofactor expansion until we hit $1 \times n$ determinants, which we know how to compute (see definition 2.2.1).

Example 2.3.4 Evaluate the determinant of $A = \begin{bmatrix} 2 & 3 & 4 & 1 \\ 1 & 5 & 0 & 2 \\ 2 & 4 & 1 & 3 \end{bmatrix}$ using Theorem 2.3.2

Solution: Since $m < n,$ expanding a long any row, say first row

$$\begin{vmatrix} 2 & 3 & 4 & 1 \\ 1 & 5 & 0 & 2 \\ 2 & 4 & 1 & 3 \end{vmatrix} = 2\,\begin{vmatrix} 5 & 0 & 2 \\ 4 & 1 & 3 \end{vmatrix} - 3\,\begin{vmatrix} 1 & 0 & 2 \\ 2 & 1 & 3 \end{vmatrix} + 4\,\begin{vmatrix} 1 & 5 & 2 \\ 2 & 4 & 3 \end{vmatrix} - 1\,\begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & 1 \end{vmatrix}$$

$$= 2(5(-2) - 0 + 2(3)) - 3(1(-2) - 0 + 2(1))$$

$$+ 4(1(1) - 5(-1) + 2(-2)) - 1(1(3) - 5(1) + 0)$$

$$= 2(-4) - 3(0) + 4(2) - 1(-2) = 2$$

In calculating a determinant using cofactor expansion, it is usually a good idea to choose a row or column containing as many zeros as possible.

Theorem 2.3.5 [3] If A is a horizontal matrix with a row of zeros, then

$\det(A) = 0$

Proof: Since the determinant of A can be found by a cofactor expansion along any row, we can use the row of zeros

$$\det(A) = 0\, C_{i1} + 0\, C_{i2} + \cdots + (-1)^{i+n}\; 0\; C_{in}$$

$$= 0$$ ■

This theorem represents property (a) in the theorem 2.1.3

2.4 Effect of elementary row operations on determinates

The evaluation of the determinant of an $m \times n$ $(m \leq n)$ matrix using the definition 2.3.1 involves the summation of ${}_{n}p_{(m-1)}$, each term being a product of m factors. As m, n increases, this computation becomes too cumbersome and so another technique has been devised to evaluate the determinant which works quite efficiently. This technique uses row operations to put a matrix into a form in which the determinant is easily calculated, keeping track of the row operations used, and how they affect the determinant, we can backtrack, and determine what the original determinant was.

We will look at the effect of each elementary row operation on the determinant.

The following theorem is from [3], [7], but we give here another proof

Theorem 2.4.1. If A and B are $m \times n$ matrices with $m \leq n$, and B is obtained from A by interchanging two rows of A, then $det\,(A) = -det\,(B)$

Proof: Base case : Let $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \end{bmatrix}$, and let

$B = \begin{bmatrix} a_{21} & a_{22} & \dots & a_{2n} \\ a_{11} & a_{12} & \dots & a_{1n} \end{bmatrix}$. Then B is the only matrix that can be obtained from A by swapping rows. And we see that

$$\det(B) = a_{21}\left(a_{12} - a_{13} + \cdots + (-1)^n a_{1n}\right) - a_{22}\left(a_{11} - a_{13} + \cdots + (-1)^n a_{1n}\right) + a_{23}\left(a_{11} - a_{12} + \cdots + (-1)^n a_{1n}\right) + \dots + (-1)^{n+1} a_{2n}\left(a_{11} - a_{12} + \cdots + (-1)^n a_{1n-1}\right)$$

$$= \left(a_{21} a_{12} - a_{21} a_{13} + \cdots + (-1)^n a_{21} a_{1n}\right) - \left(a_{22} a_{11} - a_{22} a_{13} + \cdots + (-1)^n a_{22} a_{1n}\right) + \left(a_{23} a_{11} - a_{23} a_{12} + \cdots + (-1)^n a_{23} a_{1n}\right) + \dots + (-1)^{n+1} \left(a_{2n} a_{11} - a_{2n} a_{12} + \cdots + (-1)^n a_{2n} a_{1n-1}\right) \qquad (1)$$

Also ,

$\det(A) = a_{11}\left(a_{22} - a_{23} + \cdots + (-1)^n a_{2n}\right) - a_{12}\left(a_{21} - a_{23} + \cdots + (-1)^n a_{2n}\right) + a_{13}\left(a_{21} - a_{22} + \cdots + (-1)^n a_{2n}\right) + \ldots + (-1)^{1+n} a_{1n}\left(a_{21} - a_{22} + \cdots + (-1)^n a_{2n-1}\right)$

$= \left(a_{11} a_{22} - a_{11} a_{23} + \cdots + (-1)^n a_{11} a_{2n}\right) - \left(a_{12} a_{21} - a_{12} a_{23} + \cdots + (-1)^n a_{12} a_{2n}\right) + \left(a_{13} a_{21} - a_{13} a_{22} + \cdots + (-1)^n a_{13} a_{2n}\right) + \ldots + (-1)^{1+n} \left(a_{1n} a_{21} - a_{1n} a_{22} + \cdots + (-1)^n a_{1n} a_{2n-1}\right) \qquad (2)$

From (1) and (2) , clearly

$$det\,(A) = -det\,(B)$$

Induction hypothesis: For all $k \times n$ with $(k \le n)$ matrices A, if B is obtained from A by swapping two rows , then $det\,(A) = -det\,(B)$

Induction step : Let A be a $(k+1) \times (n)$ with $(k+1 \le n)$ matrix, and let B be a matrix obtained from A by swapping two rows. Say row r and $r+1$ of A were swapped when making B.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ a_{(r+1)1} & a_{(r+1)2} & \cdots & a_{(r+1)n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{(k+1)1} & a_{(k+1)2} & \cdots & a_{(k+1)n} \end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{(r+1)1} & a_{(r+1)2} & \cdots & a_{(r+1)n} \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{(k+1)1} & a_{(k+1)2} & \cdots & a_{(k+1)n} \end{bmatrix}$$

We may evaluate $det\,(B)$ by cofactor expansion along its first row.

$$det(B) = b_{11}B_{11} + b_{12}B_{12} + b_{13}B_{13} + \cdots + b_{1n}B_{1n}$$

To compute $det(B)$, we will need to look at the submatrices $B(i\,,j\,)$. Our choice of $i = 1$ means that $B(1\,,j\,)$ can be obtained from $A(1\,,j\,)$ by swapping the rows r and $r+1$, as we swapped to get B from A. This means that $B(1\,,j\,)$ is a $k \times n$ matrix that is obtained from $A(1\,,j\,)$ by swapping two rows, and thus, by our inductive hypothesis, $detB(1\,,j\,) = -\det A(1\,,j)$.

Now,

$$det\ (B) = b_{11}(-1)^{1+1}det\ B(1,1) + \cdots + b_{1n}(-1)^{1+n}\ det\ B(1,n)$$

$$= a_{11}(-1)^{1+1}detB(1,1) + \cdots + a_{1n}(-1)^{1+n}\ detB(1,n)$$

$$= a_{11}(-1)^{1+1}(-1)detA(1,1) + \cdots + a_{1n}(-1)^{1+n}(-1)\ detA(1,n)$$

$$= -\left(a_{11}(-1)^{1+1}detA(1,1) + \cdots + a_{1n}(-1)^{1+n}\ detA(1,n)\right)$$

$$\det(B) = -\det(A)$$

This proves the result for the interchange of two adjacent rows in an $m \times n$ matrix. To see that this result holds for arbitrary row interchanges, we note that the interchange of two rows, say row r and s where $r < s$ can be performed by $2(s-r)-1$ interchanges of adjacent rows. As the number of interchanges is odd and each one changes the sign of the determinant, the net effect is a change of sign as desired.

■

We are now able to prove the cofactor expansion theorem 2.3.2

Proof: Let B be the matrix obtained by moving the i^{th} row of A to the top, using $i-1$ interchanges of adjacent rows. Thus $det\ (B) = (-1)^{i-1}\det(A)$, but $b_{1j} = a_{ij}$ and $B_{1j} = A_{ij}$ for $j \in [1,n]$ and so

$$\det\ (B) = \begin{vmatrix} a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{(i-1)1} & \cdots & a_{(i-1)j} & \cdots & a_{(i-1)n} \\ a_{(i+1)1} & \cdots & a_{(i+1)j} & \cdots & a_{(i+1)n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{vmatrix}$$

Hence,

$$\det(A) = (-1)^{i-1}\det(B) = (-1)^{i-1}\sum_{j=1}^{n}(-1)^{1+j}b_{1j}\det(B_{1j})$$

$$= (-1)^{i-1}\sum_{j=1}^{n}(-1)^{1+j}a_{ij}\det(A_{ij}) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij})$$

Giving the formula for cofactor expansion along the i^{th} row.

■

Corollary 2.4.2 [3] If any two rows of a horizontal matrix are identical, then the value of its determinant is zero.

Proof: Let $|A|$ be the determinant of the horizontal matrix A. Assume that row i and row j in A are identical**.** By Theorem 2.4.1 interchange row i and row j, the determinant of the resulting matrix is $-|A|$. But the original matrix and the resulting matrix are the same

That is $|A| = -|A|$. Hence, we obtain $|A| = 0$. ■

This theorem represents property (b) in the theorem 2.1.3

The following theorem is from [3], [7], but we give here another proof

Theorem 2.4.3 Let A and B be $m \times n$ matrices with $m \leq n$, and B is obtained from A by multiplying all the entries of some row of A by a scalar k. Then

$$det\ (B) = k\ det\ (A)$$

Proof: If we expand along the i^{th} row of B to calculate its determinant, we get

$$\det(B) = b_{i1}B_{i1} + \cdots + (-1)^{i+n}b_{in}\,B_{in}\ .$$

But the reason we have chosen the i^{th} row of B is that we know that $b_{ij} = k\,a_{ij}$ for $j = 1\,, \ldots, n$. Moreover, since the submatrices $B(i\,,j\,)$ will all have row i removed, and since this is the only place where B differs from A, we see that $A(i\,,j\,) = B(i\,,j\,)$. Thus, the cofactor B_{ij} for b_{ij} is the same as the cofactor A_{ij} for a_{ij} . So we have that

$$\begin{aligned}\det(B) &= ka_{i1}B_{i1} + \cdots + (-1)^{i+n}k\, a_{in}\, B_{in} \\ &= k\left(a_{i1}A_{i1} + \cdots + (-1)^{i+n}\, a_{in}\, A_{in}\right) \\ &= k\det(A)\end{aligned}$$

■

If we know the determinant of matrix A, we can use this information to calculate the determinant of the matrix $k\,A$, where k is a constant.

Corollary 2.4.4 Let A and B be an $m \times n$ matrices with $m \le n$, and B is obtained from A by multiplying all the entries of rows of A by a scalar k. Then

$$\det(k\,A) = k^m \det(A)$$

Proof: Since all m rows of A are multiplied by the scalar k to get $k\,A$, using the above theorem m times gives

$$\begin{aligned}\det(k\,A) &= (k)(k)\ldots(k)\det(A) \\ &= k^m \det(A)\end{aligned}$$

■

The following theorem is from [3], but we give here another proof

Theorem 2.4.5 Let A and B be an $m \times n$ matrices with $m \le n$, and B is obtained from A by adding a multiple of one row of A to another row of A. Then

$$det\,(B) = det\,(A)$$

Proof: Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \ldots & a_{sn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, and suppose that B is the matrix obtained from A by adding k times row s to row r,

$$B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \ldots & a_{sn} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{s1} + a_{r1} & ka_{s2} + a_{r2} & \cdots & ka_{s2} + a_{rn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Then we can compute the determinant of B by expanding along row r, getting

$$\det(B) = b_{r1}C_{r1} + \cdots + (-1)^{r+n} b_{rn} C_{rn}$$

Our choice of r gets us that $b_{rj} = ka_{sj} + a_{rj}$ for all $j = 1, \ldots, n$. So, now let's consider the submatrices $B(r, j)$ and $A(r, j)$ are the cofactor of A and B. This means that $B(r, j)$ can obtained from $A(r, j)$ by adding k times row s to row r. And since $B(r, j)$ and $A(r, j)$ are $(m-1) \times (n-1)$ matrices, we get

$det\, B(r, j) = \det A(r, j)$. That is,

$$det\,(B) = b_{r1}(-1)^{r+1} detB(r, 1) + \cdots + b_{rn}(-1)^{r+n}\, detB(r, n)$$

$$= (ka_{s1} + a_{r1})(-1)^{r+1} detA(r, 1) + \cdots +$$

$$+(ka_{sn} + a_{rn})(-1)^{r+n}\, detA(r, n)$$

$$\det(B) = ka_{s1}(-1)^{r+1} detA(r, 1) + a_{r1}(-1)^{r+1} detA(r, 1) + \cdots$$

$$+ ka_{sn}(-1)^{r+n} detA(r, n) + a_{rn}(-1)^{r+n} detA(r, n)$$

$$= \left(ka_{s1}(-1)^{r+1} detA(r, 1) + \cdots + \mathrm{k}\, a_{sn}(-1)^{r+n} detA(r, n)\right)$$

$$+ a_{r1}(-1)^{r+1} detA(r, 1) + \cdots + a_{rn}(-1)^{r+n} detA(r, n)$$

$$\det(B) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \ldots & a_{sn} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{s1} & ka_{s2} & \cdots & ka_{s2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \ldots & a_{sn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix},$$

any matrix in which one row is a multiple of another has determinant zero, thus, $\det(B) = 0 + \det(A)$

$$\det(B) = \det(A)$$ ■

This theorem represents property (d) in the theorem 2.1.3

Since we know how elementary row operations affect the determinant, we can compute the determinant of a given matrix by computing the determinant of its r.r.e.f (see definition 1.7.1) and taking into account the effect of the row operations. The same procedure that we used in books of linear algebra for determinant of square matrices.

The following table describes the effect of applying row operations on computing the determinant of a horizontal matrix.

	Type of ERO	Effect on determinant
1	Add a multiple of one row to another row	No effect
2	Multiply a row by a constant k	Determinate is multiplied by k
3	Interchange two rows	Determinant changes sign

We mention here that these properties correspond to their counter parts for determinants of square matrices (See Theorem 1.3.5)

Example 2.4.6 Find the determinant of $A = \begin{bmatrix} 2 & 4 & 6 \\ 1 & 3 & 5 \end{bmatrix}$

Solution: We use row reduction until A is in reduced row echelon form. At each step we keep track of the effect on the determinant.

$$\begin{vmatrix} 2 & 4 & 6 \\ 1 & 3 & 5 \end{vmatrix} \xrightarrow{R_1 \leftrightarrow R_2:\ \det\times(-1)} -\begin{vmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{vmatrix}$$

$$\xrightarrow{\mathrm{R}_2 - 2R_1 \to \mathrm{R}_2 : \det unchanged} -\begin{vmatrix} 1 & 3 & 5 \\ 0 & -2 & -4 \end{vmatrix} \xrightarrow{\left(-\frac{1}{2}\right)R_2:\ det\left(-\frac{1}{2}\right)} -\begin{vmatrix} 1 & 3 & 5 \\ 0 & 1 & 2 \end{vmatrix}$$

$$\xrightarrow{\mathrm{R}_1 - 3R_2 \to \mathrm{R}_1 : \det unchanged} \frac{1}{2}\begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{vmatrix}$$

$$\frac{1}{2}\begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{vmatrix} = \frac{1}{2}\left(1|1 \quad 2| - \ |0 \quad 1|\right)$$

$$= \frac{1}{2}\left((-1) - \ (-1)\right) = 0$$

Note that, the definition of the determinant (see theorem 2.4.1, 2.4.3, 3.1.1) satisfies the axioms of determinant function 2.1.2 (*).

Also, the cofactor definition of the determinant (see theorem 2.3.5, 2.4.2, 2.4.5) satisfies the properties of determinant function 2.1.3.

2.5 Radic's determinant

Many definitions have been proposed for the determinant of non-square matrices. Earlier works have been mainly focused on utilizing the determinant of square blocks to define the determinant of the non-square matrix. They studied many useful properties of this determinant. Radic (1969) proposed the following efficient definition that has some of the major properties of the determinants of square matrices.

Definition 2.5.1 [14]. Let $A = (a_{ij})$ be an $m \times n$ matrix with $m \leq n$. The determinant of A is defined as

$$\det(A) = \sum_{1 \leq j_1 < \cdots < j_m \leq n} (-1)^{r+S} \, det \begin{bmatrix} a_{1j_1} & \cdots & a_{1j_m} \\ \vdots & \ddots & \vdots \\ a_{mj_1} & \cdots & a_{mj_m} \end{bmatrix} \qquad (2.5.1)$$

Where $j_1, j_2, \ldots, j_m \in N$, $r = 1 + 2 + \cdots + m$ and $s = j_1 + j_2 + \cdots + j_m$.

If $m > n$, then $det(A) = det(A^T)$.

The determinant of a square matrix and the determinant 2.5.1 of a $m \times n$ matrix, where $m \leq n$, have several common standard properties, including the following:

(1) If a row of matrix A is a linear combination of some other rows, then $det(A) = 0$

(2) If a row of A is multiplied by a number k, then the determinant of the resulting matrix is equal to $k.det(A)$.

(3) Interchanging two rows of A results in changing the sign of the determinant.

(4) If the matrix A has two identical rows, then $det(A) = 0$.

Proof: Let A be an $m \times n$ matrix with $m \leq n$, by Radic's definition

$$\det(A) = \sum_{1\leq j_1<\cdots<j_m\leq n} (-1)^{r+S}\, det \begin{bmatrix} a_{1j_1} & \cdots & a_{1j_m} \\ \vdots & \ddots & \vdots \\ a_{mj_1} & \cdots & a_{mj_m} \end{bmatrix}$$

Where $j_1, j_2, \ldots, j_m \in N$, $r = 1 + 2 + \cdots + m$ and $s = j_1 + j_2 + \cdots + j_m$.

(1) If a row of matrix A is a linear combination of some other rows, then all $det \begin{bmatrix} a_{1j_1} & \cdots & a_{1j_m} \\ \vdots & \ddots & \vdots \\ a_{mj_1} & \cdots & a_{mj_m} \end{bmatrix}$ contains a row is a linear combination of some other rows, and therefore, all $det \begin{bmatrix} a_{1j_1} & \cdots & a_{1j_m} \\ \vdots & \ddots & \vdots \\ a_{mj_1} & \cdots & a_{mj_m} \end{bmatrix} = 0$, where $j_1, j_2, \ldots, j_m \in N$, are square matrices, hence $det(A) = 0$

(2) Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, and suppose that B is the matrix obtained from A by multiplying row i by k , $B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ then,

$$\det(B) = \sum_{1\leq j_1<\cdots<j_m\leq n} (-1)^{r+S}\, det \begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{ij_1} & ka_{ij_2} & \ldots & ka_{ij_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}$$

Since all $det \begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{ij_1} & ka_{ij_2} & \cdots & ka_{ij_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}$ where $j_1, j_2, \ldots, j_m \in N$ are square matrices**,**

$$det\begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{ij_1} & ka_{ij_2} & \cdots & ka_{ij_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix} = kdet\begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{ij_1} & a_{ij_2} & \cdots & a_{ij_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}, \qquad j_1, j_2, \dots, j_m \in N$$

can be taken k out from all determinant square in common to produce $det(B) = kdet(A)$.

(3) Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, and suppose that B is the matrix

obtained Interchanging two rows of A, $B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \dots & a_{sn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$

then,

$$\det(B) = \sum_{1\le j_1<\cdots<j_m\le n} (-1)^{r+S}\, det\begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{sj_1} & a_{sj_2} & \dots & a_{sj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rj_1} & a_{rj_2} & \cdots & a_{rj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix},$$

Since all $det\begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{sj_1} & a_{sj_2} & \dots & a_{sj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rj_1} & a_{rj_2} & \cdots & a_{rj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}$ where $j_1, j_2, \dots, j_m \in N$ are square matrices**,**

$$det\begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{sj_1} & a_{sj_2} & \dots & a_{sj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rj_1} & a_{rj_2} & \cdots & a_{rj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix} = -det\begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rj_1} & a_{rj_2} & \dots & a_{rj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{sj_1} & a_{sj_2} & \cdots & a_{sj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}, \qquad j_1, j_2, \dots, j_m \in N$$

can be taken sign negative out from all determinant square in common to produce $det(B) = -det(A)$.

(4) Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, then

$$\det(A) = \sum_{1 \le j_1 < \cdots < j_m \le n} (-1)^{r+S} \, det \begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rj_1} & a_{rj_2} & \cdots & a_{rj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rj_1} & a_{rj_2} & \cdots & a_{rj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}$$

Since all $det \begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rj_1} & a_{rj_2} & \cdots & a_{rj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rj_1} & a_{rj_2} & \cdots & a_{rj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}$ where $j_1, j_2, \dots, j_m \in N$ are square matrices**,**

then all $det \begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rj_1} & a_{rj_2} & \cdots & a_{rj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{rj_1} & a_{rj_2} & \cdots & a_{rj_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix} = 0$, hence $det(A) = 0$

(5) We need to show that radic definition of the determinant satisfies the axioms (A3) of determinant function

Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, $B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$,

$$C = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Then,

$$\det(C) = \sum_{1\le j_1<\cdots<j_m\le n} (-1)^{r+S}\, det \begin{bmatrix} a_{1j_1} + b_{1j_1} & a_{1j_2} + b_{1j_2} & \cdots & a_{1j_m} + b_{1j_m} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}$$

Since all $det \begin{bmatrix} a_{1j_1} + b_{1j_1} & a_{1j_2} + b_{1j_2} & \cdots & a_{1j_m} + b_{1j_m} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}$ where $j_1, j_2, \dots, j_m \in N$ are square matrices**,**

Therefore, $det \begin{bmatrix} a_{1j_1} + b_{1j_1} & a_{1j_2} + b_{1j_2} & \cdots & a_{1j_m} + b_{1j_m} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}$

$$= det \begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix} + det \begin{bmatrix} b_{1j_1} & b_{1j_2} & \cdots & b_{1j_m} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix}$$

Hence, $\det(C) =$

$$\sum_{1\le j_1<\cdots<j_m\le n} (-1)^{r+S} \left(det \begin{bmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_m} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix} + det \begin{bmatrix} b_{1j_1} & b_{1j_2} & \cdots & b_{1j_m} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_m} \end{bmatrix} \right)$$

$= \det(A) + \det(B)$

■

Note that, the definition of the determinant (see properties 2,3,5) satisfies the axioms of determinant function 2.1.2 (**).

Also, the Radic definition of the determinant (see properties 1.4) satisfies the properties of determinant function 2.1.3.

Corollary 2.5.2 The determinant obtained by cofactor expansion and Radic definition are the same.

Proof: the cofactor definition and Radic definition are determinant function, since the four characterizing properties of determinant listed in definition 2.1.2 are satisfied by the cofactor definition 2.3.1 and Radic definition 2.5.1 of determinants, and because of uniqueness of determinant function (see theorem 2.1.4), the cofactor definition and Radic definition are the same. (see * , **)

■

Example 2.5.3 Evaluate the determinant of $A = [a_1 \quad a_2 \quad a_3]$ using Radic definition

Solution: $|A| = (-1)^{1+1}a_1 + (-1)^{1+2}a_2 + (-1)^{1+3}a_3$

$$= a_1 - a_2 + a_3$$

Example 2.5.4 Evaluate the determinant of $A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$ using Radic definition

Solution :

$$|A| = (-1)^{(3)+(1+2)} \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} + (-1)^{(3)+(1+3)} \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + (-1)^{(3)+(2+3)} \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}$$

The evaluation of the determinant of an $m \times n$ matrix $(m \leq n)$ using Radic's definition reduces to evaluation of $\binom{n}{m}$ determinant of $m \times m$ matrices.

Note, Let $A_1, A_2, \dots, A_n$ be columns of the matrix $A_{m\times n}$ $(m \leq n)$. Then Radic's determinant of A is a function in the columns of A and can be written in the form

$$det(A) = det(A_1, \dots, A_n) = |A_1, \dots, A_n|$$

Theorem 2.5.5 [13] Let $A = [A_1, \dots, A_n]$ be $2 \times n$ matrix with $n \geq 2$. Then

$$det(A_1, \dots, A_n) = det(A_1, A_2 - A_3 + A_4 \dots + (-1)^n A_n) +$$

$$det(A_2, A_3 - A_4 + \dots + (-1)^{n-1} A_n) + \dots +$$

$$det(A_{n-1}, A_n)$$

Proof: By principle of mathematical induction, (P.M.I)

Base case, $n = 3$,

$$det(A_1, A_2, A_3) = det(A_1, A_2) - det(A_1, A_3) + det(A_2, A_3)$$

Since $det(A_1, A_2), det(A_1, A_3)$ are square then

$$det(A_1, A_2) - det(A_1, A_3) = det(A_1, A_2 - A_3)$$

Then, $det(A_1, A_2, A_3) = det(A_1, A_2 - A_3) + det(A_2, A_3)$

Induction hypothesis: We assume that it is true for $n = k$

$$det(A_1, \dots, A_k) = det(A_1, A_2 - A_3 + A_4 \dots + (-1)^k A_k) +$$

$$det(A_2, A_3 - A_4 + \dots + (-1)^{k-1} A_k) + \dots + det(A_{k-1}, A_k)$$

We will show that the identity holds for $n = k + 1$

$$det(A_1, \dots, A_{k+1}) = (-1)^{(3)+(3)} det(A_1, A_2) + (-1)^{(3)+(4)} det(A_1, A_3) +$$

$$(-1)^{(3)+(5)} det(A_1, A_4) + (-1)^{(3)+(6)} det(A_1, A_5) + \dots +$$

$$(-1)^{(3)+(1+k+1)} det(A_1, A_{k+1}) + (-1)^{(3)+(5)} det(A_2, A_3) +$$

$$(-1)^{(3)+(6)} det(A_2, A_4) + (-1)^{(3)+(7)} det(A_2, A_5) + \dots +$$

$$(-1)^{(3)+(2+k+1)} det(A_2, A_{k+1}) + \dots + (-1)^{(3)+(k+k+1)} det(A_k, A_{k+1})$$

$$= det(A_1, A_2 - A_3 + A_4 \dots + (-1)^{k+1} A_{k+1})$$

$$+ det(A_2, A_3 - A_4 + \dots + (-1)^k A_{k+1}) + \dots + det(A_k, A_{k+1})$$ ■

Theorem 2.5.5 converts the computations of the determinant of an $2 \times n$ matrix according to Radic definition which needs computing $\binom{n}{2}$ determinants to computation of $n-1$ determinants of size 2×2.

($\binom{n}{2} = \frac{n(n-1)}{2} > n-1$)

Example 2.5.6 Evaluate the determinant of $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$ using theorem 2.5.5

Then**,** $det(A) = \begin{vmatrix} 1 & -2 \\ 2 & -2 \end{vmatrix} + \begin{vmatrix} 3 & 5 \\ 4 & 6 \end{vmatrix} = (-2+4) + (18-20) = 0$.

Theorem 2.5.7 [13] Let $A = [A_1, \dots, A_n]$ be $2 \times n$ matrix with $n \geq 2$. Then

$$det(A_1, A_2, \dots, A_{n-1}, A_n) = det(A_1, A_2, \dots, A_{n-1}) +$$

$$(-1)^n det(A_1 - A_2 + \dots + (-1)^n A_{n-1}, A_n)$$

Proof: By principle of mathematical induction, (P.M.I)

Base case: $n = 3$,

$$det(A_1, A_2, A_3) = det(A_1, A_2) - det(A_1, A_3) + det(A_2, A_3)$$

$$= det(A_1, A_2) - det(A_1 - A_2, A_3)$$

Induction hypothesis: We assume that it is true for $n = k$

$det(A_1, \dots, A_k)$

$$= (-1)^{3+3} det(A_1, A_2) + (-1)^{3+4} det(A_1, A_3) + (-1)^{3+5} det(A_1, A_4)$$

$$+ \cdots + (-1)^{4+k} det(A_1, A_k) + (-1)^{3+5} det(A_2, A_3)$$

$$+ (-1)^{3+6} det(A_2, A_4) + (-1)^{3+7} det(A_2, A_5) + \cdots$$

$$+ (-1)^{5+k} det(A_2, A_k) + \cdots + (-1)^{2+2k} det(A_{k-1}, A_k)$$

$$= det(A_1, A_2, A_3, \dots, A_{k-1}) + (-1)^k det(A_1 - A_2 + \dots + (-1)^k A_{k-1}, A_k)$$

We will show that the equality holds for $n = k + 1$

$det(A_1 , \dots , A_{k+1}) = (-1)^{3+3} det(A_1 , A_2) + (-1)^{3+4} det(A_1 , A_3) +$

$(-1)^{3+5} det(A_1 , A_4) + \cdots + (-1)^{4+k} det(A_1 , A_k) + (-1)^{3+k} det(A_1 , A_{k+1}) +$

$+(-1)^{3+5} det(A_2 , A_3) + (-1)^{3+6} det(A_2 , A_4) + (-1)^{3+7} det(A_2 , A_5) +$

$\cdots + (-1)^{5+k} det(A_2 , A_k) + (-1)^{6+k} det(A_2 , A_{k+1}) + \cdots +$

$(-1)^{4+2k} det(A_k , A_{k+1})$

$$= det(A_1 , A_2 , A_3 , \dots , A_{k-1}) + (-1)^k det(A_1 - A_2 + \dots + (-1)^k A_{k-1}, A_k)$$
$$+ (-1)^{k+1} det(A_1 , A_{k+1}) + (-1)^{k+1} det(A_2 , A_{k+1}) + \cdots$$
$$+ det(A_k , A_{k+1})$$

$$= det(A_1 , A_2 , A_3 , \dots , A_{k-1}) + (-1)^k det(A_1 - A_2 + \dots + (-1)^k A_{k-1}, A_k)$$
$$+ (-1)^{k+1} det(A_1 - A_2 + \dots + (-1)^{k+1} A_k, A_{k+1})$$

$$det(A_1 , \dots , A_{k+1}) = det(A_1 , A_2 , A_3 , \dots , A_k)$$
$$+(-1)^{k+1} det\left(A_1 - A_2 + \dots + (-1)^{k+1} A_k, A_{k+1}\right) \quad \blacksquare$$

Example 2.5.8. Evaluate the determinant of $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$, using theorem 2.5.7

Solution :

$|A| = \begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} + (-1)^{(3)} \begin{vmatrix} -2 & 5 \\ -2 & 6 \end{vmatrix} = (4 – 6) – (-12 + 10) = 0$

Chapter Three

Properties for non-square determinant

In this chapter we study Radic definition for determinant of a rectangular matrix in more detailed way. We present new identities for the determinant of a rectangular matrix. We develop some important properties of this determinant. We generalize several classical important determinant identities, and describe how the determinant is affected by operation on columns such as interchanging columns, reversing columns or decomposing a single column.

Although we present here properties of Radic determinant but we have proved in chapter 2 that Radic determinant and the determinant by cofactor expansion give the same value. So, we may use the term determinant to mean any of the common values.

3.1 Properties for determinant of a rectangular matrix.

In this section we will be mainly concerned with the properties of the determinants of square matrices (theorem 1.3.5) that are still valid when one goes to rectangular matrices.

The following theorem is from [3], but we give here another proof

Theorem 3.1.1 If every element in any fixed row of a horizontal matrix can be expressed as the sum of tow quantities then the given horizontal matrix determinant can be expressed as the sum of tow horizontal matrix determinant of the same order with the elements of the remaining rows of the both being the same.

Proof : Let $A = \begin{bmatrix} \alpha + a_{11} & \beta + a_{12} & \dots & \delta + a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$,

By cofactor expansion along first row of $\det(A)$

$$| A | = (\alpha + a_{11}) \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m2} & \cdots & a_{mn} \end{vmatrix} - (\beta + a_{12}) \begin{vmatrix} a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} +$$

$$(\gamma + a_{13}) \begin{vmatrix} a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} + \cdots + (-1)^{1+n} (\delta + a_{1n}) \begin{vmatrix} a_{21} & \cdots & a_{2(n-1)} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{m(n-1)} \end{vmatrix}$$

$\det(A) =$

$$\left(\alpha \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m2} & \cdots & a_{mn} \end{vmatrix} - \beta \begin{vmatrix} a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} + \gamma \begin{vmatrix} a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} + \right.$$

$$\left. \cdots + (-1)^{1+n} \delta \begin{vmatrix} a_{21} & \cdots & a_{2(n-1)} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{m(n-1)} \end{vmatrix} \right) + \left(a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m2} & \cdots & a_{mn} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} + \right.$$

$$\left. a_{13} \begin{vmatrix} a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} + \cdots + (-1)^{1+n} a_{1n} \begin{vmatrix} a_{21} & \cdots & a_{2(n-1)} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{m(n-1)} \end{vmatrix} \right)$$

$$= \begin{vmatrix} \alpha & \beta & \cdots & \delta \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix} \quad \blacksquare$$

Note*:* 1- is those axiom A3 in definition 2.1.2

2- This property is valid for square matrices as well (see theorem 1.3.6)

Theorem 3.1.2 [1]. Let $1 \le m \le n$, and A be an $m \times m$ matrix, and B be an $m \times n$ matrix, then $det(A\,B) = \det(\mathrm{A})\det(\mathrm{B})$

Proof: Let $B = [B_1, \dots, B_n]$, then

$det(A\,B) = \det(\mathrm{A}\,[B_1, \dots, B_n]\,) = det\,([\mathrm{A\,B}_1, \dots, \mathrm{A\,B}_n])$

By Radic definition

$$\det(A\,B) = \sum_{1 \le j_1 < \cdots < j_m \le n} (-1)^{r+S}\, det\left(\left[\mathrm{A\,B}_{j_1}, \dots, \mathrm{A\,B}_{j_m}\right]\right)$$

Where $r = 1 + 2 + \cdots + m$ and $s = j_1 + j_2 + \cdots + j_m$

$$\det(A\,B) \;=\; \sum_{1\le j_1<\cdots<j_m\le n} (-1)^{r+S} \det(\mathrm{A})\, det\left(\left[\mathrm{B}_{\mathrm{j}_1}, \dots, \mathrm{B}_{\mathrm{j}_m}\right]\right)$$

Since Radic definition gives square matrix (see theorem1.3.5.f)

$$= \det(\mathrm{A}) \sum_{1\le j_1<\cdots<j_m\le n} (-1)^{r+S}\;\; det\left(\left[\mathrm{B}_{\mathrm{j}_1}, \dots, \mathrm{B}_{\mathrm{j}_m}\right]\right)$$

$$= \det(\mathrm{A})\det(\mathrm{B})$$

■

Example 3.1.3 Prove theorem 3.1.2 for $A = \begin{bmatrix} 2 & 2 \\ -1 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 5 & 7 \\ 2 & 1 & 4 \end{bmatrix}$

Then**,** $A\,B = \begin{bmatrix} 10 & 12 & 22 \\ 3 & -2 & 5 \end{bmatrix}$,

$\det(\mathrm{A}) = 8$, $\det(\mathrm{B}) = 8$, $\det(A.B) = 64$

$det(A\,B) = \det(\mathrm{A})\,\det(\mathrm{B}) = 8 \times 8 = 64$

A sufficient condition for the equation is that A is square and $A\,B$ is defined.

We note here that $B\,A$ is not defined.

In fact there is no determinant function that satisfies $det(AB) = det(A)det(B)$ for all matrices A, B.

Example 3.1.4 Let $A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & -2 \\ 2 & 1 & 0 \end{bmatrix}$

Then, $A\,B = \begin{bmatrix} 5 & 3 & 3 \\ 4 & 0 & -5 \end{bmatrix}$,

$\det(\mathrm{A}) = -4$, $\det(\mathrm{B}) = 7$, $\det(AB) = 10$

$\det(\mathrm{A})\,\det(\mathrm{B}) = -28 \ne det(A\,B)$

We notice from this example that the determinant is not distributed in the case that the first matrix is rectangular, so for the theorem to be true, the first matrix must be the square.

Example 3.1.5 Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 2 & 0 \\ -4 & 2 \end{bmatrix}$

Then, $A\,B = \begin{bmatrix} 2 & 1 \\ 6 & 2 \end{bmatrix}$, $B\,A = \begin{bmatrix} 2 & 3 & 0 \\ 2 & 2 & 0 \\ 0 & 2 & 0 \end{bmatrix}$

$\det(\text{A}) = 1$, $\det(\text{B}) = -2$, $\det(A\,B) = -2$, $\det(B\,A) = 0$

Now, $\det(A\,B) = \det(\text{A})\det(\text{B}) = \det(\text{B})\det(\text{A}) \neq det(B\,A)$

$$-2 = (1)(-2) = (-2)(1) \neq 0$$

Lemma 3.1.6 [1] Let A be an $m \times n$ matrix, $1 \leq m < n$, and

$m + n$ odd, then $\underset{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}{det}[(a_{ij})] = 0$, where $a_{1j} = 1$ for all j, $1 \leq j \leq n$.

Proof: by induction on even integer n for all odd integer numbers m,

$1 \leq m < n.$

Base case : If $n = 2$, then $m = 1$ we have $\det([1\,,1]) = 1 - 1 = 0$

Induction hypothesis : We assume that it is true for even n and odd m,

$1 \leq m < n,$

$D = det[A_1\,, A_2, \dots, A_n] = det\begin{bmatrix} 1 & 1 & \dots & 1 \\ B_1 & B_2 & \dots & B_n \end{bmatrix}$ where $B_j = \begin{bmatrix} a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}$, $(1 \leq j \leq n)$

Expanding the determinant with respect to the first row yields

$$D = det[B_2\,, B_3, \dots, B_n] - det[B_1\,, B_3, \dots, B_n] + \cdots + (-1)^{n+1} det[B_1\,, B_2, \dots, B_{n-1}] = 0$$

Induction step : We will show that the identity holds for $n + 2$

(which is even number) and all odd m, $1 \leq m < n + 2$

$$D = det[A_1\,, A_2, \dots, A_n\,, A_{n+1}\,, A_{n+2}] = det\begin{bmatrix} 1 & 1 & \dots & 1 & 1 & 1 \\ B_1 & B_2 & \dots & B_n & B_{n+1} & B_{n+2} \end{bmatrix}$$

Expanding the determinant with respect to the first row yields

$$D = det[B_2, B_3, \dots, B_{n+2}] - det[B_1, B_3, \dots, B_{n+2}] + \cdots$$
$$+ (-1)^{n+1} det[B_1, B_2, \dots, B_{n-1}, B_{n+1}, B_{n+2}]$$
$$+ (-1)^{1+n+1} det[B_1, B_2, \dots, B_n, B_{n+2}] + (-1)^{1+n+2} det[B_1, B_2, \dots, B_{n+1}]$$

By theorem 2.5.7, all the resulting determinants are deleted from each other.

Then , $D = 0$.

Another case can be established in the same way(odd n and even m). ■

This property does not apply if the matrices are square $(n \times n)$ because the sum of the order is even (whether n is odd or even)

Example 3.1.7 Evaluate the determinant of $A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 5 & 6 \end{bmatrix}$

Then, $\det(A) = 1 \neq 0$, and $det(B) = 0$

Lemma 3.1.8 [1] .

1. $\det[A_1, A_2, \dots, A_n, 0_m] = det[A_1, A_2, \dots, A_n]$, and
2. $\det[A_1, A_2, \dots, A_{j-1}, 0_m, A_{j+1}, \dots, A_n] = det[A_1, A_2, \dots, A_{j-1}, -A_{j+1}, \dots, -A_n]$

Where $m \leq n$, $A_K = [a_{1,k}, \dots, a_{m,k}]^T$ for $k \in \{1, \dots, n\} - \{j\}$ and 0_m is an m by 1 zero vector .

Proof : proof of the first formula,

Let $A_j = \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}$, $A_j = \begin{bmatrix} a_{1,j} \\ B_j \end{bmatrix}$ where $B_j = \begin{bmatrix} a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}$, $(1 \leq j \leq n)$

By expanding the determinant with respect to the first row, we get

$$\mathrm{D} = \det[A_1, A_2, \dots, A_n, 0_m] = det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ B_1 & B_2 & \cdots & B_n & 0_m \end{bmatrix}$$

$$D = (-1)^2 a_{1,1}\, det[B_2, B_3, \dots, B_n, 0_m] + (-1)^3\, a_{12} det[B_1, B_3, \dots, B_n, 0_m] + \cdots$$

$$+ (-1)^{1+n}\, a_{1n} det[B_1, B_2, \dots, B_{n-1}, 0_m]$$

$$+ (-1)^{2+n}\, (0)\, det[B_1, B_2, \dots, B_n]$$

$$D = a_{1,1}\, det[B_2, B_3, \dots, B_n, 0_m] - a_{12} det[B_1, B_3, \dots, B_n, 0_m] + \cdots$$

$$+ (-1)^{1+n}\, a_{1n} det[B_1, B_2, \dots, B_{n-1}, 0_m]$$

$$\det[A_1, A_2, \dots, A_n, 0_m] = det\,[A_1, A_2, \dots, A_n]$$

Now, proof the second formula, by (P.M.I)

Base case : For $n = 2$, $m = 1$, we have $det[a_1 \quad 0] = a_1 - 0 = a_1$

Induction hypothesis: Assume that for all n and m, $1 \le m < n$, it is true that

$$\det[A_1, A_2, \dots, A_{j-1}, 0_m, A_{j+1}, \dots, A_n] = det[A_1, A_2, \dots, A_{j-1}, -A_{j+1}, \dots, -A_n]$$

Induction step: We will show that the identity holds for $n + 1$, $1 \le m < n + 1$?

Let $A_j = \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}$, $A_j = \begin{bmatrix} a_{1,j} \\ B_j \end{bmatrix}$ where $B_j = \begin{bmatrix} a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}$, $(1 \le j \le n)$

By expanding the determinant with respect to the first row, we get

$$\text{D} = \det[A_1, A_2, \dots, A_{j-1}, 0_m, A_{j+1}, \dots, A_n, A_{n+1}]$$

$$= det \begin{bmatrix} a_{1,1} & \dots & a_{1,j-1} & 0 & a_{1,j+1} & \dots & a_{1,n} & a_{1,n+1} \\ B_1 & \dots & B_{j-1} & 0_m & B_{j+1} & \dots & B_n & B_{n+1} \end{bmatrix}$$

$$= a_{1,1}\, det[B_2, \dots, B_{j-1}, o_m, B_{j+1} \dots, B_n, B_{n+1}] + \cdots$$

$$+ (-1)^{j-1+1} a_{1,j-1} det[B_1, \dots, B_{j-2}, 0_m, B_{j+1} \dots, B_n, B_{n+1}] + 0$$

$$+ (-1)^{j+1+1} a_{1,j+1} det[B_1, \dots, B_{j-1}, 0_m, B_{j+2} \dots, B_n, B_{n+1}] + \cdots$$

$$+ (-1)^{n+2} a_{1,n+1} det[B_1, \dots, B_{j-1}, 0_{m+1}, B_{j+2} \dots, B_n]$$

By inductive hypothesis

$$D = a_{1,1}\, det\left[B_2 , \dots , B_{j-1}, -B_{j+1} \dots , -B_{n+1}\right] + \cdots$$

$$+ \; (-1)^{j-1+1} a_{1,j-1} det\left[B_1 , \dots , B_{j-2}, - B_{j+1} \dots , -B_{n+1}\right] + \cdots$$

$$+ \; (-1)^{j+1+1} a_{1,j+1} det\left[B_1 , \dots , B_{j-1}, - B_{j+2} \dots , -B_{n+1}\right] + \cdots$$

$$+ (-1)^{n+2} a_{1,n+1} det\left[B_1 , \dots , B_{j-1}, - B_{j+2} \dots , -B_n\right]$$

$$= det\begin{bmatrix} a_{1,1} & \dots & a_{1,j-1} & -\, a_{1,j+1} & \dots & -a_{1,n+1} \\ B_1 & \dots & B_{j-1} & -\, B_{j+1} & \dots & -B_{n+1} \end{bmatrix}$$

$$= det\left[A_1 , A_2 , \dots , A_{j-1} , -A_{j+1} , \dots , -A_{n+1}\right]$$ ■

This property does not apply if the matrices are square since if all the elements of a column are zeros, then the value of the determinant is zero. (see theorem 1.3.5)

Example 3.1.9 An example that illustrates Theorem 3.1.8 is

Let $A = \begin{bmatrix} 3 & 5 & 7 \\ 2 & 1 & 4 \end{bmatrix}$, $\det(A) = 8$

So, $\begin{vmatrix} 3 & 5 & 7 & 0 \\ 2 & 1 & 4 & 0 \end{vmatrix} = \begin{vmatrix} 3 & 5 & 7 \\ 2 & 1 & 4 \end{vmatrix} = 8,$

and $\begin{vmatrix} 3 & 0 & 5 & 7 \\ 2 & 0 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 3 & -5 & -7 \\ 2 & -1 & -4 \end{vmatrix} = 8$

Theorem 3.1.10 [1]. Suppose $1 \le m < n$, and $m + n$ be an odd integer, $A = \left(a_{i,j}\right) = [A_1, \dots , A_n]$ be an $m \times n$ matrix, and X be an arbitrary $m \times 1$ column vector, then

$$\det[A_1 + X , \dots , A_n + X] = det[A_1 , \dots , A_n]$$

Proof: by (P.M.I)

Base case : For $n = 2$, $m = 1$, we have

$$\det\left[a_{1,1} + X \quad a_{1,2} + X\right] = a_{1,1} - a_{1,2} = det\left[a_{1,1} \quad a_{1,2}\right]$$

Induction hypothesis: Assume the assertion is true for all n even and m odd with $1 \leq m < n$,

$$\det[A_1 + X, \dots, A_n + X] = det[A_1, \dots, A_n]$$

We have to prove that is true for $n+2$ even and all m odd ?

With $1 \leq m < n+2$,

Let $A_j = \begin{bmatrix} a_{1j} \\ B_j \end{bmatrix}$ where $B_j = \begin{bmatrix} a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$, $(1 \leq j \leq n+2)$ and $X = \begin{bmatrix} x_1 \\ X^{/} \end{bmatrix}$ where $X^{/} = \begin{bmatrix} x_2 \\ \vdots \\ x_m \end{bmatrix}$

By expanding the determinant with respect to the first row , we get

$$\det[A_1 + X, \dots, A_{n+2} + X] = det \begin{bmatrix} a_{1,1} + x_1 & \dots & a_{1,n+2} + x_1 \\ B_1 + X^{/} & \dots & B_{n+2} + X^{/} \end{bmatrix}$$

$$= (-1)^{1+1}(a_{1,1} + x_1) det\left[B_2 + X^{/}, \dots, B_{n+2} + X^{/} \right] + \cdots$$

$$+ (-1)^{n+2+1}(a_{1,n+2} + x_1) det\left[B_1 + X^{/}, \dots, B_{n+1} + X^{/} \right]$$

By induction hypothesis for each component we obtain:

$$det[A_1 + X, \dots, A_{n+2} + X] = (-1)^{1+1}(a_{1,1} + x_1) det[B_2, \dots, B_{n+2}]$$

$$+ \cdots + (-1)^{n+2+1}(a_{1,n+2} + x_1) det[B_1, \dots, B_{n+1}]$$

$$= (-1)^{1+1}(a_{1,1}) det[B_2, \dots, B_{n+2}] + \cdots + (-1)^{n+2+1}(a_{1,n+2}) det[B_1, \dots, B_{n+1}]$$

$$+ x_1 \left[(-1)^{1+1} det[B_2, \dots, B_{n+2}] + \cdots (-1)^{n+2+1} det[B_1, \dots, B_{n+1}] \right]$$

$$det[A_1 + X, \dots, A_{n+2} + X] = det \begin{bmatrix} a_{1,1} & \dots & a_{1,n+2} \\ B_1 & \dots & B_{n+2} \end{bmatrix} + x_1 det \begin{bmatrix} 1 & \dots & 1 \\ B_1 & \dots & B_{n+2} \end{bmatrix}$$

Applying lemma 3.1.6

$$det[A_1 + X, \dots, A_{n+2} + X] = det \begin{bmatrix} a_{1,1} & \dots & a_{1,n+2} \\ B_1 & \dots & B_{n+2} \end{bmatrix} + 0$$

$$= det[A_1, \dots, A_{n+2}]$$

■

The second case (n odd and m even) can be treated similarly.

Example 3.1.11 Let $A = \begin{bmatrix} 3 & 5 & 7 \\ 2 & 1 & 4 \end{bmatrix}$, take $X = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$,

Then, $det(A) = \det(A_1, A_2, A_3) = 8$

Also, $\det(A_1 + X, A_2 + X, A_3 + X) = \begin{vmatrix} 4 & 6 & 8 \\ 4 & 3 & 6 \end{vmatrix} = 8$

Corollary 3.1.12 [1] For $m + n$ odd**,** $1 \leq \mathrm{m} < \mathrm{n}$ and for all k,

$1 \leq \mathrm{k} \leq \mathrm{n}$, we have

$$\det[A_1, \dots, A_{k-1}, A_k, A_{k+1}, \dots, A_n]$$
$$= det[A_k - A_1, \dots, A_k - A_{k-1}, A_{k+1} - A_k, \dots, A_n - A_k]$$

Proof: $\det(\mathrm{A}) = \det[A_1, \dots, A_{k-1}, A_k, A_{k+1}, \dots, A_n]$

Applying theorem 3.1.10 with $\mathrm{X} = -\mathrm{A_k}$ and lemma 3.1.8

$$= det[A_1 - A_k, \dots, A_{k-1} - A_k, A_k - A_k, A_{k+1} - A_k, \dots, A_n - A_k]$$

$$= det[A_1 - A_k, \dots, A_{k-1} - A_k, 0_m, A_{k+1} - A_k, \dots, A_n - A_k]$$

$$= det[A_1 - A_k, \dots, A_{k-1} - A_k, A_k - A_{k+1}, \dots, A_k - A_n]$$

$$= det[A_k - A_1, \dots, A_k - A_{k-1}, A_{k+1} - A_k, \dots, A_n - A_k]$$

■

Example 3.1.13 let $A = \begin{bmatrix} 3 & 5 & 7 \\ 2 & 1 & 4 \end{bmatrix}$, then $\det(A) = 8$

Applying Corollary 3.1.12 with $A_k = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$

$\det(A) = \begin{vmatrix} 2 & 2 \\ -1 & 3 \end{vmatrix} = 8$

Note, This properties (theorem 3.1.10 and corollary 3.1.12) do not apply if the matrices are square ($n \times n$) because the sum of the order is even (Whether n is odd or even)

3.2 How determinant is affected by operations on columns.

In this section we describe how the determinant is affected by operations on columns such as interchanging columns, reversing columns or decomposing a single column.

(1) Decomposing column in a square matrix if

$A_{n\times n} = [A_1, A_2, \dots, A_k, \dots, A_n]$ and $A_k = B_k + C_k$, $k \epsilon \{1,2,\dots,n\}$ then

$| A | = | A_1, A_2, \dots, A_{k-1}, B_k, A_{k+1}, \dots, A_n| + | A_1, A_2, \dots, A_{k-1}, C_k, A_{k+1}, \dots, A_n|$.

In non-square matrices case what happens, we have

Theorem 3.2.1 [12]. Let $A = [A_1, A_2, \dots, A_k, \dots, A_n]$ be a $m \times n$ matrix, $m \leq n$, and $A_k = B_k + C_k$ for some $k \epsilon \{1,2,\dots,n\}$. Then

$$| A | = | A_1, A_2, \dots, A_{k-1}, B_k, A_{k+1}, \dots, A_n| + | A_1, A_2, \dots, A_{k-1}, C_k, A_{k+1}, \dots, A_n|$$

$$+ \sum_{\substack{1\leq j_1<\cdots<j_m \leq n \\ k\notin\{j_1, \dots, j_m\}}} (-1)^{r+j_1+j_2+\cdots+j_m+1} \left|A_{j_1}, A_{j_2}, \dots, A_{j_m}\right|$$

Where $r = 1 + 2 + \cdots + m$

Proof : After applying (2.5.1)

$$|A| = \sum_{1\leq j_1<\cdots<j_m \leq n} (-1)^{r+j_1+j_2+\cdots+j_m} \left|A_{j_1}, A_{j_2}, \dots, A_{j_m}\right|$$

We separate the sum of determinants into tow sums: the first one consisting of the determinants of matrices which contain the column $A_k = B_k + C_k$ and the second one consisting of other determinants.

$$| A | = \sum_{\substack{1\leq j_1<\cdots<j_m \leq n \\ k\in\{j_1, \dots, j_m\}}} (-1)^{r+j_1+j_2+\cdots+j_m} \left|A_{j_1}, \dots A_k, \dots, A_{j_m}\right|$$
$$+ \sum_{\substack{1\leq j_1<\cdots<j_m \leq n \\ k\notin\{j_1, \dots, j_m\}}} (-1)^{r+j_1+j_2+\cdots+j_m} \left|A_{j_1}, A_{j_2}, \dots, A_{j_m}\right|$$

$$= \sum_{\substack{1 \le j_1 < \cdots < j_m \le n \\ k \in \{j_1, \dots, j_m\}}} (-1)^{r+j_1+j_2+\cdots+j_m} \left|A_{j_1}, \dots, B_k, \dots, A_{j_m}\right|$$

$$+ \sum_{\substack{1 \le j_1 < \cdots < j_m \le n \\ k \in \{j_1, \dots, j_m\}}} (-1)^{r+j_1+j_2+\cdots+j_m} \left|A_{j_1}, \dots, C_k, \dots, A_{j_m}\right|$$

$$+ \sum_{\substack{1 \le j_1 < \cdots < j_m \le n \\ k \notin \{j_1, \dots, j_m\}}} (-1)^{r+j_1+j_2+\cdots+j_m} \left|A_{j_1}, A_{j_2}, \dots, A_{j_m}\right|$$

Now the third sum is added and subtracted so that it can be included into both the first and the second sum :

$$|A| = |A_1, A_2, \dots, A_{k-1}, B_k, A_{k+1}, \dots, A_n| + |A_1, A_2, \dots, A_{k-1}, C_k, A_{k+1}, \dots, A_n|$$

$$- \sum_{\substack{1 \le j_1 < \cdots < j_m \le n \\ k \notin \{j_1, \dots, j_m\}}} (-1)^{r+j_1+j_2+\cdots+j_m} \left|A_{j_1}, A_{j_2}, \dots, A_{j_m}\right|$$

$$= |A_1, A_2, \dots, A_{k-1}, B_k, A_{k+1}, \dots, A_n| + |A_1, A_2, \dots, A_{k-1}, C_k, A_{k+1}, \dots, A_n|$$

$$+ \sum_{\substack{1 \le j_1 < \cdots < j_m \le n \\ k \notin \{j_1, \dots, j_m\}}} (-1)^{r+j_1+j_2+\cdots+j_m+1} \left|A_{j_1}, A_{j_2}, \dots, A_{j_m}\right|$$

■

Example 3.2.2 Let $[A_1, A_2, A_3]$ be a 2×3 matrix and $A_1 = B_1 + C_1$. Then according to theorem 2.7.2 we have

$$|B_1 + C_1, A_2, A_3|$$

$$= |B_1, A_2, A_3| + |C_1, A_2, A_3| + \sum_{\substack{1 \le j_1 < j_2 \le 3 \\ 1 \notin \{j_1, j_2\}}} (-1)^{(1+2)+j_1+j_2+1} \left|A_{j_1}, A_{j_2}\right|$$

$$= |B_1, A_2, A_3| + |C_1, A_2, A_3| + (-1)^{3+2+3+1} |A_2, A_3|$$

$$= |B_1, A_2, A_3| + |C_1, A_2, A_3| - |A_2, A_3|.$$

(2) Interchanging columns in a square matrix results in changing the sign of the determinant. Non-square matrices in which the number of columns is equal to the number of rows increased by one have the same property.

Theorem 3.2.3 [12]. Let $A = [A_1, A_2, \dots, A_m, A_{m+1}]$ be a $m \times (m+1)$ matrix. Then for each $i, j \in \{1,2,\dots,m+1\}$ such that $i < j$, we have

$$det(A) = -det\big(A_1, A_2, \dots, A_{i-1}, A_j, A_{i+1}, \dots, A_{j-1}, A_i, A_{j+1}, \dots, A_m, A_{m+1}\big)$$

Proof: Let $r = 1 + 2 + \cdots + m$. Fix each $i, j \in \{1,2,\dots,m+1\}$ such that $i < j$. From all the determinants in the right–hand side of

$$|A| = \sum_{1 \le j_1 < \cdots < j_m \le n} (-1)^{r+j_1+j_2+\cdots+j_m} \left|A_{j_1}, A_{j_2}, \dots, A_{j_m}\right|$$

We distinguish the determinants in the expression which contain either A_i or A_j but not both of them. Thus, we have

$$|A| = (-1)^{\left(r+\frac{(m+1)(m+2)}{2}-i\right)}\left|A_1, A_2, \dots, A_{i-1}, A_{i+1}, \dots, A_{j-1}, A_j, A_{j+1}, \dots, A_{m+1}\right|$$

$$+ (-1)^{\left(r+\frac{(m+1)(m+2)}{2}-j\right)}\left|A_1, A_2, \dots, A_{i-1}, A_i, A_{i+1}, \dots, A_{j-1}, A_{j+1}, \dots, A_{m+1}\right|$$

$$+ \sum_{\substack{1 \le j_1 < \cdots < j_m \le n \\ i, j \in \{j_1, \dots, j_m\}}} (-1)^{r+j_1+j_2+\cdots+j_m} \left|A_{j_1}, \dots, A_i, \dots, A_j, \dots, A_{j_m}\right|$$

Notice that exactly $j - i - 1$ inversions are needed to move the column A_j to the position between A_{i-1} and A_{i+1} in the first summand. Similarly, in the second summand, also $j - i - 1$ inversions are needed to move the column A_i to the position between A_{j-1} and A_{j+1} .

In other summands we can simply interchange columns A_i and A_j with the sign change (square matrix $m \times m$). Thus, we have

$$|A| = (-1)^{\left(r+\frac{(m+1)(m+2)}{2}-i+(i-j-1)\right)} \times \left|A_1, A_2, \dots, A_{i-1}, A_j, A_{i+1}, \dots, A_{j-1}, A_{j+1}, \dots, A_{m+1}\right|$$

$$+ (-1)^{\left(r+\frac{(m+1)(m+2)}{2}-j+(j-i-1)\right)}\left|A_1, A_2, \dots, A_{i-1}, A_{i+1}, \dots, A_{j-1}, A_i, A_{j+1}, \dots, A_{m+1}\right|$$

$$- \sum_{\substack{1 \le j_1 < \cdots < j_m \le n \\ i,j \in \{j_1, \ldots, j_m\}}} (-1)^{(-1)^{r+j_1+j_2+\cdots+j_m}} \left| A_{j_1}, \ldots, A_j, \ldots, A_i, \ldots, A_{j_m} \right|$$

$$\det(A) = -(-1)^{\left(r+\frac{(m+1)(m+2)}{2}-j\right)} \left| A_1, A_2, \ldots, A_{i-1}, A_j, A_{i+1}, \ldots, A_{j-1}, A_{j+1}, \ldots, A_{m+1} \right|$$

$$- (-1)^{\left(r+\frac{(m+1)(m+2)}{2}-i\right)} \left| A_1, A_2, \ldots, A_{i-1}, A_{i+1}, \ldots, A_{j-1}, A_i, A_{j+1}, \ldots, A_{m+1} \right|$$

$$- \sum_{\substack{1 \le j_1 < \cdots < j_m \le n \\ i,j \in \{j_1, \ldots, j_m\}}} (-1)^{r+j_1+j_2+\cdots+j_m} \left| A_{j_1}, \ldots, A_j, \ldots, A_i, \ldots, A_{j_m} \right|$$

$$\det(A) = -det\left(A_1, A_2, \ldots, A_{i-1}, A_j, A_{i+1}, \ldots, A_{j-1}, A_i, A_{j+1}, \ldots, A_m, A_{m+1}\right) \qquad \blacksquare$$

Remark 3.2.4 Consider an $m \times n$ matrix A with m rows and n columns, $m \le n$. Let $A^{/}$ be a matrix obtained from A by interchanging two columns. Theorem 3.2.3 tells us that $\det(A) + \det\left(A^{/}\right) = 0$ when $n - m = 1$. However, in general, if $n - m > 1$ the sum $\det(A) + \det\left(A^{/}\right)$ is not zero as explained in the following example.

Example 3.2.5 Let $A = \begin{bmatrix} 1 & 2 & 7 & 4 & 3 \\ 3 & 0 & 1 & -1 & 2 \end{bmatrix}$, $\det(A) = 5$

And $A^{/} = \begin{bmatrix} 1 & 2 & 4 & 7 & 3 \\ 3 & 0 & -1 & 1 & 2 \end{bmatrix}$, $\det\left(A^{/}\right) = 5$

$\det(A) + \det\left(A^{/}\right)$ is not zero, since $n - m > 1$

Theorem 3.2.6 [1] (Cyclic). If $1 \le m < n$, and $m + n$ is an odd integer, then for all $i \in \{1, \ldots, n\}$. We have

$$(-1)^{(i+1)m} det[A_i, \ldots, A_n, \ldots, A_{i-1}] = det[A_1, \ldots, A_n]$$

Proof: It is sufficient to prove

$$(-1)^m det[A_n, A_1, \ldots, A_{n-1}] = det[A_1, \ldots, A_n]$$

Applying Theorem 3.1.10 with $X = -A_n$ and Lemma 3.1.8 we have

$$(-1)^m det[A_n, A_1, \dots, A_{n-1}] = (-1)^m det[0_m, A_1 - A_n, \dots, A_{n-1} - A_n]$$

$$= (-1)^m (-1)^m det[\, A_1 - A_n, \dots, A_{n-1} - A_n]$$

$$= det[\, A_1 - A_n, \dots, A_{n-1} - A_n]$$

$$= det[\, A_1 - A_n, \dots, A_{n-1} - A_n,\ A_n - A_n]$$

Applying Theorem 3.1.10 with $X = A_n$

$$(-1)^m det[A_n, A_1, \dots, A_{n-1}] = det[A_1, \dots, A_n]$$ ■

Theorem 3.2.7 [1] (Semi-Cyclic). If $1 \le m < n$, and $m + n$ even, then

for all $i \in \{1, \dots, n\}$. We have

$$(-1)^{(n-i)m} det[A_i, \dots, A_n, -A_1, \dots, -A_{i-1}] = det[A_1, \dots, A_n]$$

Proof: It is sufficient to prove

$$det[A_n, -A_1, \dots, -A_{n-1}] = det[A_1, \dots, A_n]$$

Applying Lemma 3.1.8 and Theorem 3.1.10 with $X = -A_n$ we have

$$det[A_n, -A_1, \dots, -A_{n-1}] = det[A_n, -A_1, \dots, -A_{n-1}, 0_m]$$

$$= det[A_n - A_n, -A_1 - A_n, \dots, -A_{n-1} - A_n, 0_m - A_n]$$

$$= det[0_m, -A_1 - A_n, \dots, -A_{n-1} - A_n, -A_n]$$

$$= det[\, A_1 + A_n, \dots, A_{n-1} + A_n, A_n]$$

$$= det[\, A_1 + A_n, \dots, A_{n-1} + A_n, A_n, 0_m\,]$$

$$= det[\, A_1, \dots, A_{n-1}, 0_m, 0_m - A_n\,]$$

$$= det[\, A_1, \dots, A_{n-1}, A_n\,]$$ ■

Example 3.2.8 Let $A = \begin{bmatrix} 1 & 2 & 7 & 4 & 3 \\ 3 & 0 & 1 & -1 & 2 \end{bmatrix}$, then $\det(A) = 5$

let $i = 2$, by theorem 3.2.6

$$\det(A_2) = (-1)^{(3)(2)} \begin{vmatrix} 2 & 7 & 4 & 3 & 1 \\ 0 & 1 & -1 & 2 & 3 \end{vmatrix} = 5$$

let $i = 3$, by theorem 3.2.6

$$\det(A_3) = (-1)^{(4)(2)} \begin{vmatrix} 7 & 4 & 3 & 1 & 2 \\ 1 & -1 & 2 & 3 & 0 \end{vmatrix} = 5$$

Let $B = \begin{bmatrix} 1 & 2 & 7 & 2 \\ 3 & 0 & 1 & 3 \end{bmatrix}$, then $\det(B) = 26$

Let $i = 2$, by theorem 3.2.7

$$\det(B_2) = (-1)^{(2)(2)} \begin{vmatrix} 2 & 7 & 2 & -1 \\ 0 & 1 & 3 & -3 \end{vmatrix} = 26$$

(3) Reversing columns in a $n \times n$ square matrix results in changing the sign of its determinant if and only if n is congruent to 2 or 3 (mod 4). Surprisingly, the determinant of non-square matrix also either change or does not change the sign after column reversing, depending on the number of rows and the number of columns of the matrix.

Theorem 2.7.13 [12]. Let $[A_1, A_2, \ldots, A_n]$ be a $m \times n$ matrix, $m \le n$. Then we have $|A_n, A_{n-1}, \ldots, A_2, A_1| = |A_1, A_2, \ldots, A_{n-1}, A_n|.(-1)^{\frac{m(2n+m+1)}{2}}$

$$= \begin{cases} |A_1, A_2, \ldots, A_{n-1}, A_n| & if\ m \equiv 0 (mod 4), \\ |A_1, A_2, \ldots, A_{n-1}, A_n| .(-1)^{n+1} & if\ m \equiv 1 (mod 4), \\ |A_1, A_2, \ldots, A_{n-1}, A_n|.(-1) & if\ m \equiv 2 (mod 4), \\ |A_1, A_2, \ldots, A_{n-1}, A_n| .(-1)^{n} & if\ m \equiv 3 (mod 4), \end{cases}$$

Proof: Let $r = 1 + 2 + \cdots + m = \frac{m(m+1)}{2}$ and $B_k = A_{n+1-k}$,

$k \in \{1,2, \ldots, n\}$. Since exactly $(m-1) + (m-2) + \cdots + 1 = \frac{(m-1)m}{2}$ inversions of (adjacent) columns are needed to reverse the columns of a $m \times m$ matrix, we have

$$|B_1, B_2, \ldots, B_n| = \sum_{1 \le i_1 < \cdots < i_m \le n} (-1)^{r + i_1 + i_2 + \cdots + i_m} |B_{i_1}, B_{i_2}, \ldots, B_{i_m}|$$

$$= \sum_{1\le i_1<\cdots<i_m\le n} (-1)^{r+i_1+i_2+\cdots+i_m+\frac{(m-1)m}{2}} \left|B_{i_m}, B_{i_{m-1}}, \dots, B_{i_1}\right|$$

$$= \sum_{1\le i_1<\cdots<i_m\le n} (-1)^{r+i_1+i_2+\cdots+i_m+\frac{(m-1)m}{2}} \left|A_{n+1-i_m}, A_{n+1-i_{m-1}}, \dots, A_{n+1-i_1}\right|.$$

Applying the following change of variables: $j_k = n+1-i_{m-k+1}$ for each $k \in \{1,2,\dots,m\}$, we get

$|A_n, A_{n-1}, \dots, A_2, A_1| = |B_1, B_2, \dots, B_n|$

$$= \sum_{1\le j_1<\cdots<j_m\le n} (-1)^{r+(n+1)-(j_1+j_2+\cdots+j_m)+\frac{(m-1)m}{2}} \left|A_{j_1}, A_{j_2}, \dots, A_{j_m}\right|$$

$$= \sum_{1\le j_1<\cdots<j_m\le n} (-1)^{r+m(n+1)-(j_1+j_2+\cdots+j_m)+\frac{(m-1)m}{2}} \left|A_{j_1}, A_{j_2}, \dots, A_{j_m}\right|$$

$$= |A_1, A_2, \dots, A_{n-1}, A_n|.(-1)^{m(n+1)+\frac{(m-1)m}{2}}$$

$$= |A_1, A_2, \dots, A_{n-1}, A_n|.(-1)^{\frac{m(2n+m+1)}{2}}$$

Finally, we state that

$$(-1)^{\frac{m}{2}(2n+m+1)} = \begin{cases} 1 & if\ m \equiv 0(mod4), \\ (-1)^{n+1} & if\ m \equiv 1(mod4), \\ (-1) & if\ m \equiv 2(mod4), \\ (-1)^{n} & if\ m \equiv 3(mod4), \end{cases}$$

Which is easy to verify. ■

Example 2.7.14 Let $A = \begin{bmatrix} 1 & 2 & 7 & 4 & 3 \\ 3 & 0 & 1 & -1 & 2 \end{bmatrix}$

Then $\det(A) = 5$, by theorem 2.7.13,

$$\begin{vmatrix} 3 & 4 & 7 & 2 & 1 \\ 2 & -1 & 1 & 0 & 3 \end{vmatrix} = (-1)^{\frac{2}{2}(10+2+1)} \det(A) = (-1)(5) = -5$$

$m \equiv 2(mod4) \to 2\ mod\ 4 = 2$

$$\Rightarrow \begin{vmatrix} 3 & 4 & 7 & 2 & 1 \\ 2 & -1 & 1 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 7 & 4 & 3 \\ 3 & 0 & 1 & -1 & 2 \end{vmatrix} (-1) = -5$$

Chapter Four

Applications of Determinant to Area of Polygons in R^2

In this Chapter we will study the application for determinants of non-square matrices in calculating the area of polygons in $\boldsymbol{R^2}$.

4.1 Areas of certain polygons in connection with determinant of rectangular matrices.

First we state and prove the following result that relates the determinant of a 2×2 matrix with the area of the parallelogram spanned by its columns

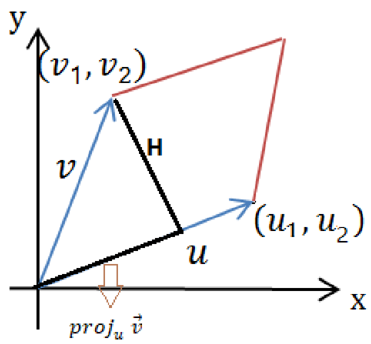
Theorem 4.1.1 [9] The absolute value of the determinant

$$\det \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix}$$

is equal to the area of the parallelogram in 2-space determined by the vectors

$u = (u_1, u_2)$ *and* $v = (v_1, v_2)$

Proof: The parallelogram defined by the columns of the above matrix is the one with vertices at $(0,0), (u_1, u_2),\ (u_1 + v_1, u_2 + v_2)\ and\ (v_1, v_2)$, as shown in the accompanying diagram.

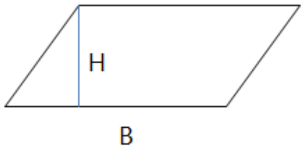


area of the parallelogram $A = B\,H$

$B = \|\vec{u}\|$, and $\quad H^2 + \|proj_u\,\vec{v}\|^2 = \|\vec{v}\|^2$

$H^2 = \|\vec{v}\|^2 - \|proj_u\,\vec{v}\|^2$

$H^2 = \vec{v}.\vec{v} - \left\|\frac{\vec{v}.\vec{u}}{\vec{u}.\vec{u}}\, u\right\|^2$ (see theorem 1.9.6)



$$H^2 = \vec{v}.\vec{v} - \left(\frac{\vec{v}.\vec{u}}{\vec{u}.\vec{u}}\,\vec{u}\,.\frac{\vec{v}.\vec{u}}{\vec{u}.\vec{u}}\,\vec{u}\right)$$

$$H^2 = \vec{v}.\vec{v} - \left(\frac{(\vec{v}.\vec{u})(\vec{v}.\vec{u})}{(\vec{u}.\vec{u})(\vec{u}.\vec{u})}\;(\vec{u}.\vec{u})\right)$$

$$H^2 = \vec{v}.\vec{v} - \left(\frac{(\vec{v}.\vec{u})^2}{(\vec{u}.\vec{u})}\right)$$

Now, $A^2 = B^2\,H^2$

$$A^2 = \vec{u}.\vec{u}\left(\vec{v}.\vec{v} - \frac{(\vec{v}.\vec{u})^2}{(\vec{u}.\vec{u})}\right)$$

$$A^2 = (\vec{u}.\vec{u})(\vec{v}.\vec{v}) - (\vec{v}.\vec{u})^2$$

But, $u = (u_1, u_2)\ and\ \ v = (v_1, v_2)$

So,

$A^2 = (u_1^2 + u_2^2)(v_1^2 + v_2^2) - (u_1v_1 + u_2\,v_2)^2$ (see definition 1.8.3)

$$A^2 = u_1^2\,v_1^2 + u_1^2\,v_2^2 + u_2^2\,v_1^2 + u_2^2\,v_2^2 - u_1^2\,v_1^2 - 2u_1v_1u_2\,v_2 - u_2^2\,v_2^2$$

$$A^2 = u_1^2\,v_2^2 - 2u_1v_1u_2\,v_2 + u_2^2\,v_1^2$$

$$A^2 = (u_1v_2 - u_2\,v_1)^2 = \left(\det\begin{bmatrix}u_1 & v_1\\ u_2 & v_2\end{bmatrix}\right)^2$$

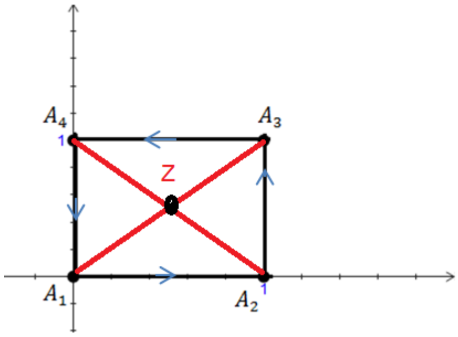
Area = $\left|\det\begin{bmatrix}u_1 & v_1\\ u_2 & v_2\end{bmatrix}\right|$

The area of the triangle whose heads are $u = (u_1, u_2)\ and\ \ v = (v_1, v_2)$ is

Area = $\frac{1}{2}\left|\det\begin{bmatrix}u_1 & v_1\\ u_2 & v_2\end{bmatrix}\right|$ ■

$det(u, v) = -det(v, u)$ and orientation

Also, the area of the quadrilateral can be found by dividing the shape into four triangles, as in the adjacent figure



area $A_1\,A_2\,A_3\,A_4$ = area of triangle $A_1\,A_2\,Z$ + area of triangle $A_2\,A_3\,Z$ + area of triangle $A_3\,A_4\,Z$ + area of triangle $A_4\,A_1\,Z$

area $A_1\,A_2\,A_3\,A_4 = \frac{1}{2}|A_1, A_2| + \frac{1}{2}|A_2, A_3| + \frac{1}{2}|A_3, A_4| + \frac{1}{2}|A_4, A_1|$

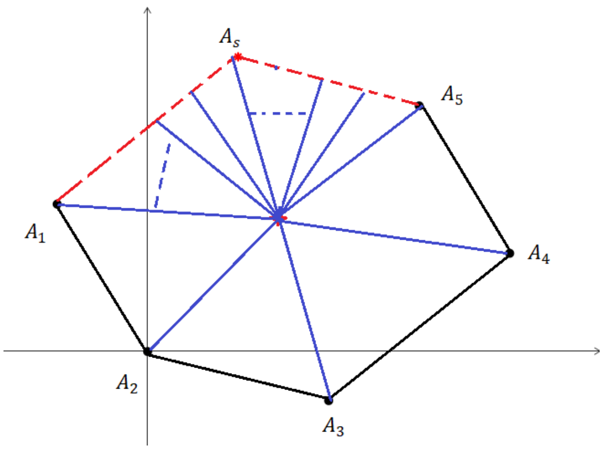
This division of quadrilateral into triangles gives us the idea of calculating the area of any polygon in R^2.

Theorem 4.1.2 The area of a polygon with vertices

$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ listed counter-clockwise around the perimeter is given by

$$A = \frac{1}{2}\left(\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} + \dots + \begin{vmatrix} x_n & x_1 \\ y_n & y_1 \end{vmatrix} \right)$$

From the adjacent figure it is clear that the area of the polygon is equal to the sum of the areas of triangles resulting from two consecutive vertices and the center of the polygon



It is clear that every real $m \times n$ matrix $A = [A_1, \dots, A_n]$ determines a polygon in R^m (the columns of the matrix correspond to the vertices of the polygon) and vice versa. The polygon which corresponds to the matrix $[A_1, \dots, A_n]$ will be denoted by $A_1 \dots A_n$.

area of $A_1 \dots A_n = \frac{1}{2}(|A_1, A_2| + |A_2, A_3| + \cdots + |A_{n-1}, A_n| + |A_n, A_1|)$.

In the following we shall restrict ourselves to the case when $m = 2$ (polygons in R^2).

Now, if $n = 3$. The area A with vertices $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ listed counter-clockwise around the perimeter is given by

$$A = \frac{1}{2}\left(\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} + \begin{vmatrix} x_3 & x_1 \\ y_3 & y_1 \end{vmatrix} \right)$$

$$= \frac{1}{2}\big((x_1y_2 - x_2y_1) + (x_2y_3 - x_3y_2) + (x_3y_1 - x_1y_3)\big)$$

$$= \frac{1}{2}\big(x_1(y_2 - y_3) - x_2(y_1 - y_3) + x_3(y_1 - y_2)\big)$$

$$= \frac{1}{2}\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$$

Theorem 4.1.3 The area A of the polygon determined by $A_1 \dots A_n$ is given by

$$A = \frac{1}{2}(|A_1, A_2| + |A_2, A_3| + \cdots + |A_{n-1}, A_n| + |A_n, A_1|)$$

$$= |A_1, A_2, \dots, A_n|$$

Proof: What we need here is to prove the second equality, we shall proceed by showing that the formula $\frac{1}{2}(|A_1, A_2| + |A_2, A_3| + \cdots + |A_{n-1}, A_n| + |A_n, A_1|)$ is indeed a determinant function

A1: The area of a triangle with vertices $(1,0), (0,1), (0,0)$ is $\frac{1}{2}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \frac{1}{2}$

A2: It is clear that multiplying a single row by a constant c means the dialation or contraction of one coordinate while keeping the other coordinate fixed and in this case new area equals $|c|$ times old area.

A3: A Let A a polygon with vertices $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ such that $x_i = u_i + v_i$, and B a polygon with vertices $(u_1, y_1), (u_2, y_2), \dots, (u_n, y_n)$, and C a polygon with vertices $(v_1, y_1), (v_2, y_2), \dots, (v_n, y_n)$

Suppose that, area of a polygon $A = [A_1, \dots, A_n] = \begin{bmatrix} u_1 + v_1 & u_2 + v_2 & \cdots & u_n + v_n \\ y_1 & y_2 & \cdots & y_n \end{bmatrix}$, but, area of a polygon $B = [B_1, \dots, B_n] = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \\ y_1 & y_2 & \cdots & y_n \end{bmatrix}$,

area of a polygon $C = [C_1, \dots, C_n] = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \\ y_1 & y_2 & \cdots & y_n \end{bmatrix}$,

that is, area of $A_1 \dots A_n = \frac{1}{2}(|A_1, A_2| + |A_2, A_3| + \cdots + |A_{n-1}, A_n| + |A_n, A_1|)$

$$= \frac{1}{2}\left(\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} + \cdots + \begin{vmatrix} x_n & x_1 \\ y_n & y_1 \end{vmatrix} \right)$$

$$= \frac{1}{2}\left(\begin{vmatrix} u_1 + v_1 & u_2 + v_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} u_2 + v_2 & u_3 + v_3 \\ y_2 & y_3 \end{vmatrix} + \cdots + \begin{vmatrix} u_n + v_n & u_1 + v_1 \\ y_n & y_1 \end{vmatrix} \right)$$

(use theorem 3.1.1)

$$= \frac{1}{2}\left(\begin{vmatrix} u_1 & u_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} v_1 & v_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} u_2 & u_3 \\ y_2 & y_3 \end{vmatrix} + \begin{vmatrix} v_2 & v_3 \\ y_2 & y_3 \end{vmatrix} + \cdots + \begin{vmatrix} u_n & u_1 \\ y_n & y_1 \end{vmatrix} + \begin{vmatrix} v_n & v_1 \\ y_n & y_1 \end{vmatrix} \right)$$

$$= \frac{1}{2}\left(\begin{vmatrix} u_1 & u_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} u_2 & u_3 \\ y_2 & y_3 \end{vmatrix} + \cdots + \begin{vmatrix} u_n & u_1 \\ y_n & y_1 \end{vmatrix} \right) + \frac{1}{2}\left(\begin{vmatrix} v_1 & v_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} v_2 & v_3 \\ y_2 & y_3 \end{vmatrix} + \begin{vmatrix} v_n & v_1 \\ y_n & y_1 \end{vmatrix} \right)$$

area of $A_1 \dots A_n$

$$= \frac{1}{2}(|B_1, B_2| + |B_2, B_3| + \cdots + |B_n, B_1|) + \frac{1}{2}(|C_1, C_2| + |C_2, C_3| + \cdots + |C_n, C_1|)$$

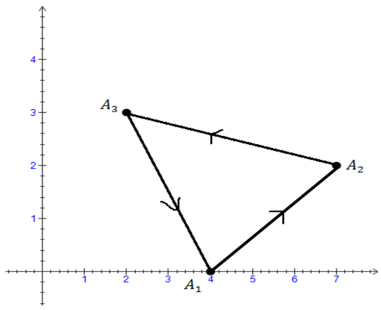
area of $A_1 \dots A_n =$ area of $B_1 \dots B_n$ + area of $C_1 \dots C_n$

A4: Exchanging two rows geometrically means the reflection of the heads of the polygon about in the straight line $y = x$, That is, the image of each point (x, y) under the transformation is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$, which keeps the area fixed.

Since the area of a polygon is a determinant function, which is unique by Theorem 2.1.4, we obtain that the area equals the previously defined determinant **(** the cofactor definition and Radic definition). ■

Example 4.1.4 Use matrices to find the area of triangle with vertices:

(4, 0), (7, 2), (2, 3)



area of triangle $= \frac{1}{2}\begin{vmatrix} 4 & 7 & 2 \\ 0 & 2 & 3 \end{vmatrix} = \frac{13}{2}$

4.2 Some properties of the determinant and their geometric interpretation

Theorem 4.2.1 [13]. Let $A_1 \dots A_n$ be a polygon in R^2. Then

$$\text{area of } A_1 \dots A_n = \frac{1}{2}\, |A_1 + A_2, A_2 + A_3, \dots, A_{n-1} + A_n, A_n + A_1|$$

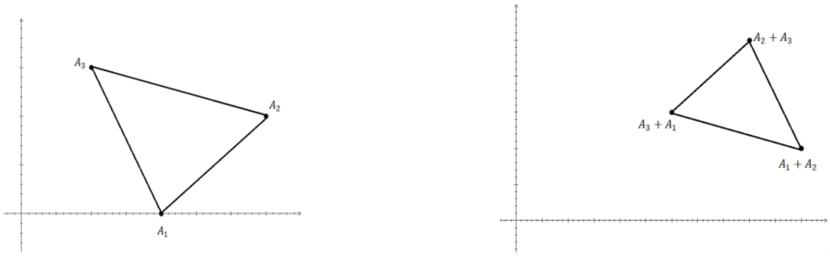
Proof: We need to show that

$$|A_1 + A_2, A_2 + A_3, \dots, A_{n-1} + A_n, A_n + A_1| =$$

$$|A_1\,, A_2| + |A_2\,, A_3| + \cdots + |A_{n-1}\,, A_n| + |A_n\,, A_1|$$

The proof will use the method of mathematical induction.

First, we have the theorem holds for $n = 3$, that is



From the adjacent figure it appears that the triangular shape resulted from the vertices shift process and therefore,

$$\text{area of } A_1A_2A_3 = \frac{1}{2}|A_1 + A_2, A_2 + A_3, A_3 + A_1|$$

$$= \frac{1}{2}(|A_1 + A_2\,, A_2 + A_3| + |A_2 + A_3\,, A_3 + A_1| + |A_3 + A_1\,, A_1 + A_2|)$$

Applying Theorem 3.1.10 with $X_1 = -A_2$, $X_2 = -A_3$, $X_3 = -A_1$ on the determinants in order

area of $A_1A_2A_3 = \frac{1}{2}(|A_1, A_3| + |A_2, A_1| + |A_3, A_2|)$

$$= \frac{1}{2}(|A_1, A_2| + |A_2, A_3| + |A_3, A_1|)$$

Second, for $n \geq 3$. Assume that is true for $n = k$, by theorem 2.5.7

$|A_1 + A_2, A_2 + A_3, \dots, A_{K-1} + A_K, A_k + A_1|$

$$= |A_1 + A_2, A_2 + A_3, \dots, A_{k-1} + A_k| + (-1)^k|A_1 + (-1)^k A_k, A_k + A_1|$$

$$= |A_1 + A_2, A_2 + A_3, \dots, A_{k-1} + A_k| + (-1)^k|A_1, A_k| + |A_k, A_1|$$

$$= |A_1, A_2| + |A_2, A_3| + \cdots + |A_{k-1}, A_k| + |A_k, A_1| + (-1)^k|A_1, A_k| + |A_k, A_1|$$

We shall prove the result holds for $k + 1$

$det[A_1 + A_2, A_2 + A_3, \dots, A_{k-1} + A_k, A_k + A_{k+1}, A_{k+1} + A_1]$

$$= \det[A_1 + A_2, A_2 + A_3, \dots, A_{k-1} + A_k]$$
$$+ (-1)^k \det[A_1 + (-1)^k A_k, A_k + A_1] + \det[A_k + A_{k+1}, A_{k+1} + A_1]$$

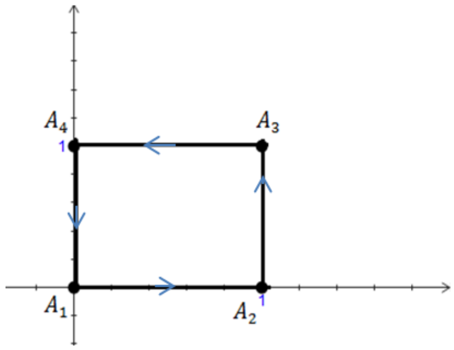
$$= \det[A_1 + A_2, A_2 + A_3, \dots, A_{k-1} + A_k] + (-1)^k \det[A_1, A_k] - \det[A_k, A_1]$$
$$+ \det[A_k, A_{k+1}] + \det[A_k, A_1] + \det[A_{k+1}, A_1]$$

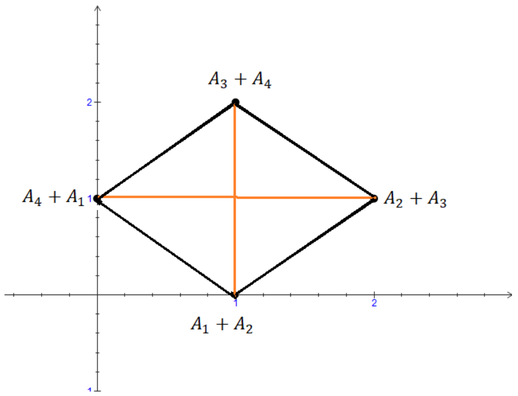
$$= \det[A_1 + A_2, A_2 + A_3, \dots, A_{k-1} + A_k] + (-1)^k \det[A_1, A_k] + \det[A_k, A_{k+1}]$$
$$+ \det[A_{k+1}, A_1]$$

$$= \det[A_1, A_2] + \det[A_2, A_3] + \cdots + \det[A_k, A_{k+1}] + \det[A_{k+1}, A_1]$$

■

Example 4.2.2 Use Theorem 4.2.1 to compute the area of the polygon in R^2 with vertices (0, 0), (1, 0), (1, 1), (0 ,1)





Now, Area $A_1\ A_2\ A_3\ A_4 = \frac{1}{2}\ |A_1 + A_2, A_2 + A_3, A_3 + A_4, A_4 + A_1|$

$$= \frac{1}{2}\begin{vmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{vmatrix} = 1$$

Corollary 4.2.3 [13]. If n is odd, then for every point X in R^2 it holds

$$|A_1 + X, \ldots, A_n + X| = |A_1, \ldots, A_n|$$

Proof: Since $m = 2$, if n is odd, then $m + n$ is odd, by theorem 3.1.10 it holds

that is, the area of the polygon $A_1 A_2 \ldots A_n$ will not be changed when a fixed vector $X = (x, y)$ is subtracted from all the vertices (heads) of the polygon.

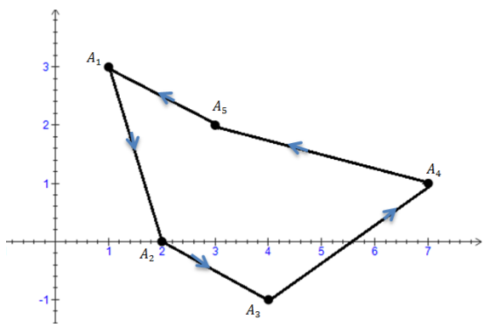
Example 4.2.4 The area of a polygon with vertices

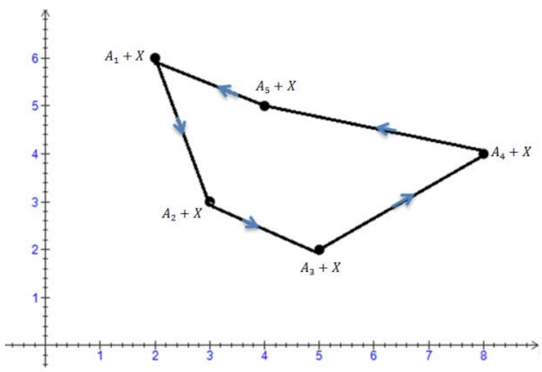
$$(1\ ,3), (2\ ,0), (7\ ,1), (4\ ,-1), (3\ ,2)$$

when all the heads are shifted by an amount $X = (1,3)$

$$\begin{vmatrix} 1 & 2 & 7 & 4 & 3 \\ 3 & 0 & 1 & -1 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 8 & 5 & 4 \\ 6 & 3 & 4 & 2 & 5 \end{vmatrix} = -3$$

area of a polygon = 3





Theorem 4.2.5 [13]. Let $A_1 \dots A_n$ be a polygon in R^2 with odd n.

Then

$$|A_1, \dots, A_n| = |A_n, A_1, \dots, A_{n-1}|$$

Proof: Since $m = 2$, if n is odd, then $m + n$ is odd**.**

by Cyclic theorem, $(-1)^{(i+1)m} det[A_i , \dots, A_n , \dots, A_{i-1}] = det[A_1 , \dots, A_n]$

When $i = n$, then $(i + 1)m$ is even

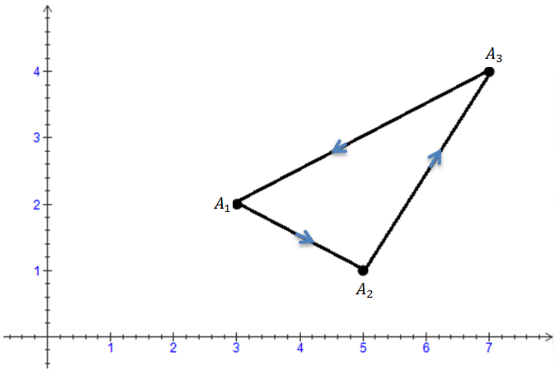
So, $|A_1, \dots, A_n| = |A_n, A_1, \dots, A_{n-1}|$.

Example 4.2.6 Use Theorem 4.2.5 to find the area of a polygon with vertices:

$A_1 = (3,2)$, $A_2 = (5,1)$, $A_3 = (7,4)$

Then, area of a polygon $A_1 A_2 A_3 =$

$$\begin{vmatrix} 3 & 5 & 7 \\ 2 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 7 & 3 & 5 \\ 4 & 2 & 1 \end{vmatrix} = 8$$



Note, we can calculate the area by starting from any vertex of the polygon paying attention to the direction and arrangement of the vertices.

Area of a polygon $A_1 A_2 A_3 = \det(A_1, A_2, A_3) = \det(A_2, A_3, A_1) = \det(A_3, A_1, A_2)$

Theorem 4.2.7 [13]. Let $A_1 \dots A_n$ be a polygon in $R^2$and let n be an even integer. Then for any point X in R^2 it holds

$$|A_1 + X, \dots, A_n + X| = |A_1, \dots, A_n|$$

Only if $\sum_{i=1}^{n} (-1)^i A_i = 0$.

Proof: This theorem is incompatible with the theorem 3.1.10 and the reason is $m = 2$, if n is even, then $m + n$ is even, This does not meet the theorem 3.1.10 requirement

Now, by Theorem 2.5.7, it is clear that for any point p in R^2 holds

$$|A_1, A_2, \dots, A_n, p| = |A_1, A_2, \dots, A_n| + (-1)^{n+1}|A_1 - A_2 + \cdots + (-1)^{n+1} A_n, p|$$
$$= |A_1, A_2, \dots, A_n| \qquad (1)$$

Only if $A_1 - A_2 + \cdots - A_n = 0$.

Now, by theorem 3.1.10, since $m + n + 1$ is odd, taking $X = -p$, we can write

$$|A_1, A_2, \dots, A_n, p| = |A_1 + (-p), A_2 + (-p), \dots, A_n + (-p), p + (-p)|$$

$$= |A_1 + (-p), A_2 + (-p), \dots, A_n + (-p), 0_2|$$

By Lemma 3.1.8

$$= |A_1 + (-p), A_2 + (-p), \dots, A_n + (-p)|$$

Putting $X = -p$, we get

$$|A_1, A_2, \dots, A_n, p| = |A_1 + X, A_2 + X, \dots, A_n + X|$$

since $|A_1, A_2, \dots, A_n, p| = |A_1, A_2, \dots, A_n|$ (from 1)

We get,

$$|A_1, A_2, \dots, A_n| = |A_1 + X, A_2 + X, \dots, A_n + X|$$ ■

That is, the area of the polygon $A_1 A_2 \dots A_n$ will not be changed when a fixed vector $X = (x, y)$ is subtracted from all the vertices (heads) of the polygon.

When for example $X = -A_1$, we obtain

$$|A_1, A_2, \dots, A_n| = |0, A_2 - A_1, \dots, A_n - A_1|$$

$$= |A_1 - A_2, \dots, A_1 - A_n|$$

$$= |B_1, B_2, \dots, B_{n-1}|$$

$$= \frac{1}{2} |B_1, B_2| + |B_2, B_3| + \cdots + |B_{n-1}, B_1|$$

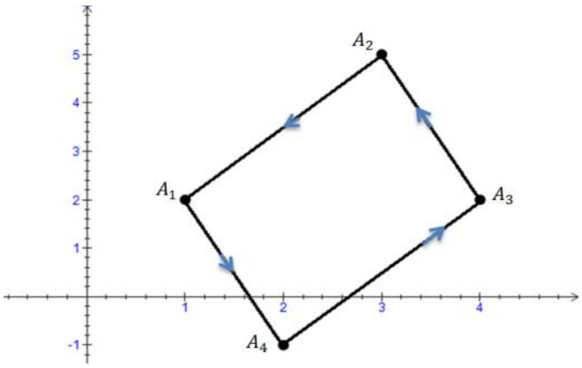
Example 4.2.8 Use Theorem 4.2.7 to find the area of a polygon with vertices:

$$A_1 = (1,2),\ \ A_2 = (3,5),\ \ A_3 = (4,2),\ \ A_4 = (2,-1)$$

Since $-A_1 + A_2 - A_3 + A_4 = 0$

area of a polygon

$$A_1A_2\,A_3A_4 = \left\|\begin{matrix}1 & 3 & 4 & 2\\ 2 & 5 & 2 & -1\end{matrix}\right\| = |-9| = 9$$



We note that when we move on the vertices of the polygon in opposite direction but in the same order, we obtain the same area but opposite value of the determinant.

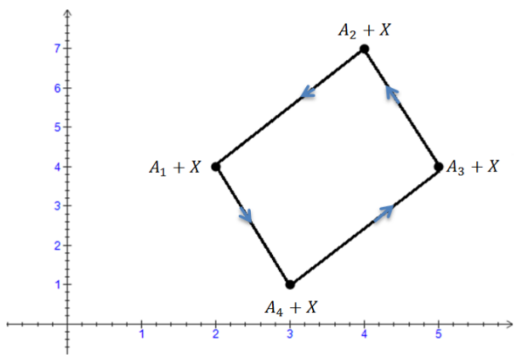
While the original space recedes itself, meaning

area of a polygon $A_1A_4\,A_3A_2 = \left|\begin{matrix}1 & 2 & 4 & 3\\ 2 & -1 & 2 & 5\end{matrix}\right| = 9$

Now, by an amount $X = (1,2)$

area of a polygon $(A_1 + X)(A_2 + X)\,(A_3 + X)(A_4 + X)$

$$= \left\|\begin{matrix}2 & 4 & 5 & 3\\ 4 & 7 & 4 & 1\end{matrix}\right\| = |-9| = 9$$



Theorem 4.2.9 [13]. Let $A_1 \dots A_n$ be a polygon in R^2 with even n and let $\sum_{i=1}^{n}(-1)^i A_i = 0$. Then

$$|A_1, \dots, A_n| = |A_n, A_1, \dots, A_{n-1}|$$

Proof: If n is even, by theorem 4.1.10, taking $X = -A_n$, we can write

$|A_1, \dots, A_n| = |A_1 - A_n, \dots, A_{n-1} - A_n, A_n - A_n|$

$= |A_1 - A_n, \dots, A_{n-1} - A_n, 0_2|$ (see lemma 3.1.8)

$= |0_2, A_n - A_1, \dots, A_n - A_{n-1}|$

Adding $-A_n$ to each column

$$det(A_1, \dots, A_n) = det(-A_n, -A_1, \dots, -A_{n-1})$$

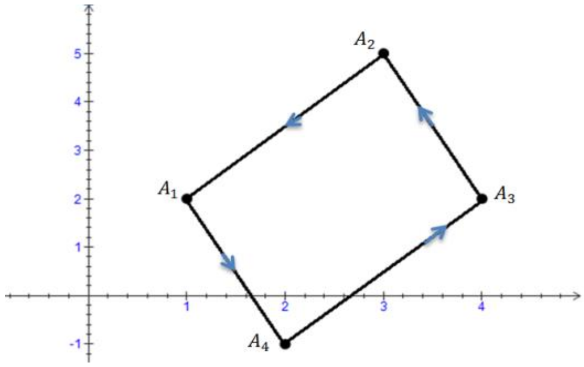
Take out a common negative signal factor from the two rows

$$det(A_1, \dots, A_n) = det(A_n, A_1, \dots, A_{n-1}) \qquad \blacksquare$$

Note that, this says that we can calculate the area by starting from any vertex of the polygon paying attention to the direction and arrangement of the vertices.

Example 4.2.10 Use Theorem 4.2.9 to find the area of a polygon with vertices:

$A_1 = (1,2),\ A_2 = (3,5),\ A_3 = (4,2),\ A_4 = (2,-1)$



Since $-A_1 + A_2 - A_3 + A_4 = 0$

$$\left|\begin{vmatrix} 1 & 3 & 4 & 2 \\ 2 & 5 & 2 & -1 \end{vmatrix}\right| = \left|\begin{vmatrix} 2 & 1 & 3 & 4 \\ -1 & 2 & 5 & 2 \end{vmatrix}\right| = 9$$

We note that when we move on the vertices of the polygon in opposite direction but in the same order, we obtain the same area but opposite value of the determinant.

Theorem 4.2.11 [13]. Let $A_1 \dots A_n$ be a polygon in R^2 with even n and let $\sum_{i=1}^{n}(-1)^i A_i = 0$. Then

$$|A_1, \dots, A_n| = |A_1, \dots, A_{n-1}|$$

Proof: By Theorem 2.5.7**,**

$$\begin{aligned}|A_1, A_2, \dots, A_n| &= |A_1, A_2, \dots, A_{n-1}| + (-1)^n |A_1 - A_2 + \cdots + A_{n-1}, A_n| \\ &= |A_1, A_2, \dots, A_{n-1}|\end{aligned}$$

since $\sum_{i=1}^{n}(-1)^i A_i = 0$, we get $A_1 - A_2 + \cdots + A_{n-1} = A_n$.

That means, $|A_1, A_2, \dots, A_n| = |A_1, A_2, \dots, A_{n-1}|$ $\blacksquare$

We notice that if we displace the shape to the original point and calculate the area by dividing the shape into such based on the origin point, the area does not change.

Example 4.2.12 If $A_1A_2A_3A_4$ is parallelogram, then

$$det[\, A_1, A_2, A_3, A_4] = det[\, A_1, A_2, A_3]$$

Now, $|\, A_1, A_2, A_3| = |\, A_1, A_2| - |\, A_1, A_3| + |\, A_2, A_3|$

$$= |\, A_1, A_2| + |-A_1 + A_2, A_3|$$

Since n is even, we get $-A_1 + A_2 - A_3 + A_4 = 0$

$|\, A_1, A_2, A_3| = |\, A_1, A_2| + |A_3 - A_4, A_3|$

$$= |\, A_1, A_2| + |-A_4, A_3| = |\, A_1, A_2| + |\, A_3, A_4|$$

$|\, A_1, A_2, A_3, A_4| = |\, A_1, A_2| + |\, A_3, A_4|$

Example 4.2.13 Use Theorem 4.2.11 to find the area of a polygon with vertices:

$$A_1 = (1,2),\ \ A_2 = (3,5),\ \ A_3 = (4,2),\ \ A_4 = (2,-1)$$

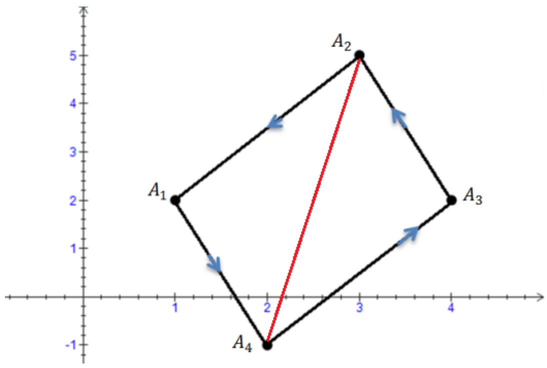
Since $-A_1 + A_2 - A_3 + A_4 = 0$

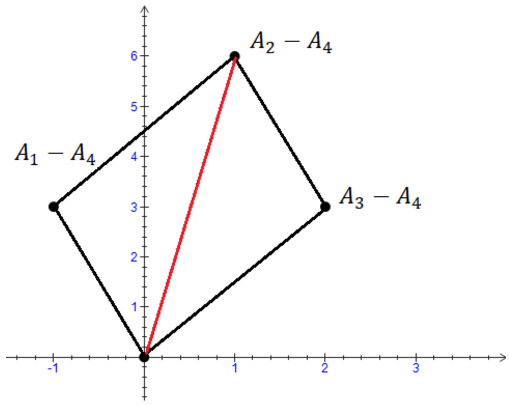
$|A_1, A_2, A_3,\ A_4| = |A_1 - A_4, A_2 - A_4, A_3 - A_4,\ A_4 - A_4|$

$$= |A_1 - A_4, A_2 - A_4, A_3 - A_4, 0| = |A_1 - A_4, A_2 - A_4, A_3 - A_4|$$

$$= \begin{vmatrix} -1 & 1 & 2 & 0 \\ 3 & 6 & 3 & 0 \end{vmatrix} = \begin{vmatrix} -1 & 1 & 2 \\ 3 & 6 & 3 \end{vmatrix} = -9$$

area of a polygon $A_1A_2\,A_3A_4 = 9$



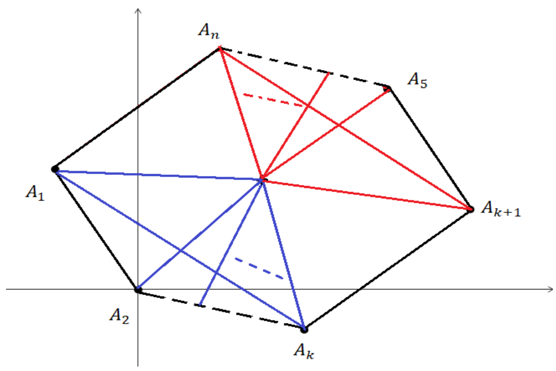


Note that the point can be considered as the center on which to divide triangles to calculate the area

then,

area of a polygon $A_1A_2\,A_3A_4$ = $\left|\begin{vmatrix} 1 & 3 & 4 & 2 \\ 2 & 5 & 2 & -1 \end{vmatrix}\right| = \left|\begin{vmatrix} 1 & 3 & 4 \\ 2 & 5 & 2 \end{vmatrix}\right| = |-9| = 9$

Theorem 4.2.14 [13]. Let $A_1 \dots A_n$ be a polygon in R^2 with even n and let $\sum_{i=1}^{n}(-1)^i A_i = 0$. Then



$$|A_1, \dots, A_n| = |A_1, \dots, A_k| + |A_{k+1}, \dots, A_n|,$$

Where k may be any integer such that $1 < k < n$.

Proof:

area of $A_1 \dots A_n$

$= \frac{1}{2}(|A_1, A_2| + |A_2, A_3| + \dots + |A_{k-1}, A_k| + |A_k, A_{k+1}| + |A_{k+1}, A_{k+2}| + \dots +$
$|A_{n-1}, A_n| + |A_n, A_1|)$

$= \frac{1}{2}(|A_1, A_2| + |A_2, A_3| + \dots + |A_{k-1}, A_k| + |A_{k+1}, A_{k+2}| + \dots + |A_{n-1}, A_n|) +$
$\frac{1}{2}(|A_k, A_{k+1}| + |A_n, A_1|)$

$|A_1, \dots, A_n| = |A_1, \dots, A_k| + |A_{k+1}, \dots, A_n| + \frac{1}{2}(|A_k, A_{k+1}| + |A_n, A_1|) -$
$\frac{1}{2}(|A_k, A_1| + |A_n, A_{k+1}|)$

let $\Delta = \frac{1}{2}(|A_k, A_{k+1}| + |A_n, A_1| - |A_k, A_1| - |A_n, A_{k+1}|)$

$|A_1, \dots, A_n| = |A_1, \dots, A_k| + |A_{k+1}, \dots, A_n| + \Delta$

But, if $\sum_{i=1}^{n}(-1)^i A_i = 0$, then $\Delta = 0$ (see figure)

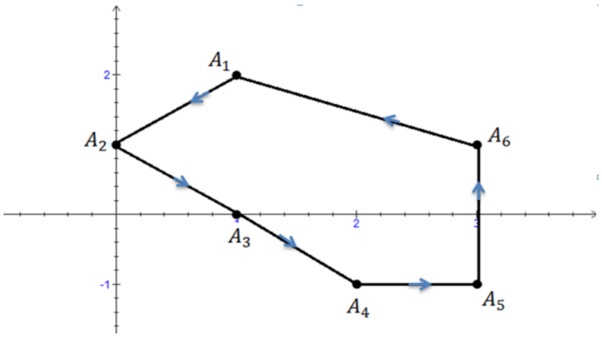
Then, $|A_1, \dots, A_n| = |A_1, \dots, A_k| + |A_{k+1}, \dots, A_n|$ ∎

Example 4.2.15 Use Theorem 4.2.14 to find the area of a polygon with vertices:

$A_1 = (1,2),\ A_2 = (0,1),\ A_3 = (1,0),\ A_4 = (2,-1),\ A_5 = (3,-1),\ A_6 = (3,1)$

Since

$$-A_1 + A_2 - A_3 + A_4 + A_5 - A_6 = 0$$

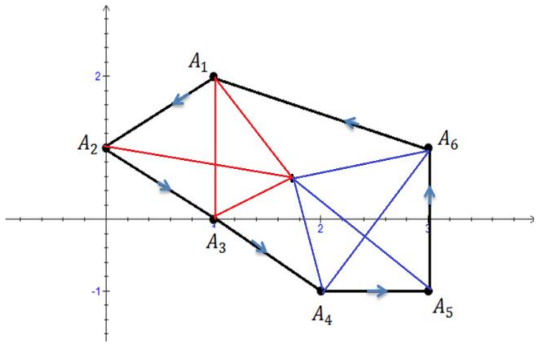


Then,

area of a polygon $A_1A_2\,A_3A_4\,A_5\,A_6 =$

$$\left\|\begin{matrix}1 & 0 & 1 & 2 & 3 & 3\\ 2 & 1 & 0 & -1 & -1 & 1\end{matrix}\right\| = 4$$

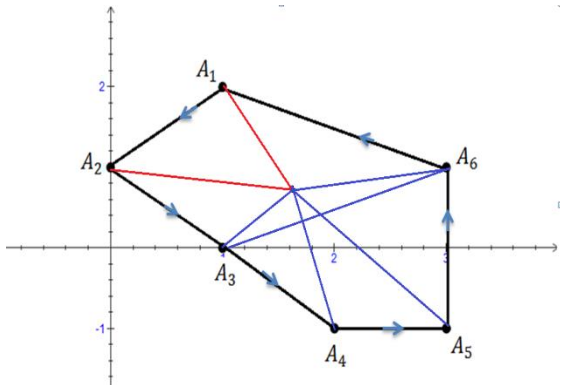
take $k = 3$ then



area of a polygon $A_1A_2\,A_3$ + area of a polygon $A_4\,A_5\,A_6$

$$= \left\|\begin{matrix}1 & 0 & 1\\ 2 & 1 & 0\end{matrix}\right\| + \left\|\begin{matrix}2 & 3 & 3\\ -1 & -1 & 1\end{matrix}\right\| = 2 + 2 = 4$$

take $k = 2$ then

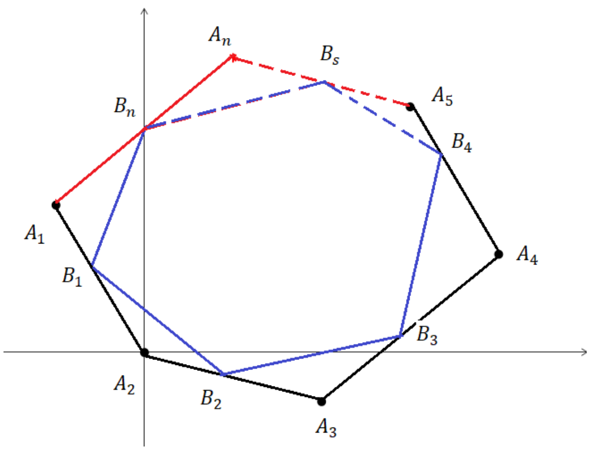


area of a polygon A_1A_2 + area of a polygon $A_3A_4\,A_5\,A_6$

$$= \left\|\begin{matrix}1 & 0\\ 2 & 1\end{matrix}\right\| + \left\|\begin{matrix}1 & 2 & 3 & 3\\ 0 & -1 & -1 & 1\end{matrix}\right\| = 1 + 3 = 4$$

Corollary 4.2.16 [13]. Let $B_1, \ldots, B_n$ be given by

$$B_1 = \frac{A_1+A_2}{2}\ ,\ B_2 = \frac{A_2+A_3}{2}\ ,\ \ldots.\ ,\ B_n = \frac{A_n+A_1}{2}.$$



Then

$$4|B_1, \ldots, B_n| = |A_1, A_2| + |A_2, A_3| + \cdots + |A_n, A_1|$$

Proof: $4|B_1, \dots, B_n| = 4\left|\frac{A_1+A_2}{2}, \frac{A_2+A_3}{2}, \dots, \frac{A_n+A_1}{2}\right|$.

By theorem 4.2.1

$$4|B_1, \dots, B_n| = 4\left(\frac{1}{4}|A_1, A_2| + \frac{1}{4}|A_2, A_3| + \cdots + \frac{1}{4}|A_n, A_1|\right)$$

$$= |A_1, A_2| + |A_2, A_3| + \cdots + |A_n, A_1| \qquad \blacksquare$$

Example 4.2.17 From the adjacent figure we find that

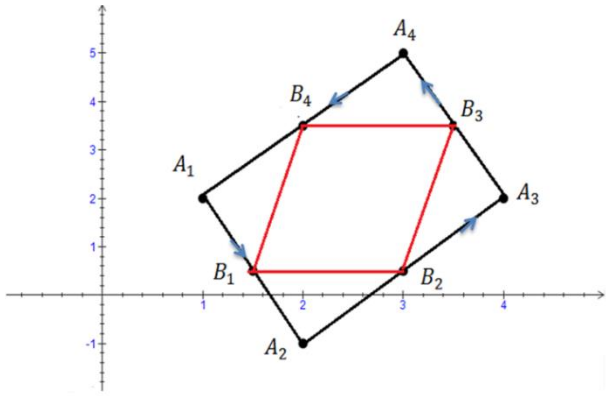
area of a polygon $B_1B_2\,B_3B_4 = \frac{1}{2}$ area of a polygon $A_1\,A_2\,A_3\,A_4$

Use Theorem 4.1.3 to find the area of a polygon with vertices:

$$A_1 = (1,2),\ A_2 = (3,5),\ A_3 = (4,2),\ A_4 = (2,-1)$$

Since $-A_1 + A_2 - A_3 + A_4 = 0$

Then, area of a polygon $A_1A_2\,A_3A_4 = 9$



area of a polygon with vertices

$B_1 = (2,3.5),\ B_2 = (3.5,3.5),\ B_3 = (3,0.5),\ B_4 = (1.5,0.5)$

Since $-B_1 + B_2 - B_3 + B_4 = 0$

Then, area of a polygon $B_1B_2\,B_3B_4 = \left\|\begin{matrix} 2 & 3.5 & 3 & 1.5 \\ 3.5 & 3.5 & 0.5 & 0.5 \end{matrix}\right\| = 4.5$

That means, by Corollary 4.2.16

area of a polygon $B_1B_2\,B_3B_4 = \frac{1}{2}$ area of a polygon $A_1\,A_2\,A_3\,A_4$.

Chapter Five

Inverse for non-square matrix

In linear algebra, the inverse of a matrix is defined only for square matrices, and if a matrix is singular, it does not have an inverse.

The aim of this chapter is the discussion of existence of inverses for non-square matrices. We know that the fundamental idea for existence of inverse of matrix it must be nonsingular. (it has non-zero determinant).

5.1 Right inverse or left inverse of a matrix

Although non-square matrices do not have inverses (both sides inverse), some of them have one side inverses. For this reason we introduce the concepts of "left inverse" and "right inverse"

Definition 5.1.1 $[3\ , p. 397\ ]$.

(i) A non-singular non-square matrix A has a left inverse if there exists a matrix A_L^{-1} such that $A_L^{-1}A = I,$ where I denote the identity matrix.

(ii) A non-singular non-square matrix A has a right inverse if there exists a matrix A_R^{-1} such that $A\, A_R^{-1} = I,$ where I denote the identity matrix.

We note here a right inverse of $A_{m\times n}$ $(m < n)$ is an A_R^{-1} matrix, where a left inverse of $A_{m\times n}$ $(m > n)$ is an A_L^{-1} matrix.

Example 5.1.2 Let A = $\begin{bmatrix} 1 & 3 & 0 \\ 2 & 2 & 1 \end{bmatrix}$ Find a right inverse of A

Solution: Let A_R^{-1} be a a right inverse of A, then

$$A\, A_R^{-1} = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} x_1 + 3x_2 & y_1 + 3y_2 \\ 2x_1 + 2x_2 + x_3 & 2y_1 + 2y_2 + y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Gauss-Jordan reduction Using we obtain the following

$$x_1 + \frac{3}{4}\, x_3 = \frac{-1}{2} \quad , \quad x_2 - \frac{1}{4}\, x_3 = \frac{1}{2} \quad , \quad y_1 + \frac{3}{4}\, y_3 = \frac{3}{4} \quad , \ y_2 - \frac{1}{4}\, y_3 = \frac{-1}{4} \qquad (*)$$

This system has infinitely many solutions, one solution gives the following

$$A_R^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & -1/3 \\ -2 & -1/3 \end{bmatrix}, \text{ another solution is } A_R^{-1} = \begin{bmatrix} 0 & 0 \\ 13/32 & 0 \\ -3/8 & 1 \end{bmatrix}$$

Computing a right inverse of horizontal matrix always can be transformed to finding a solution of a linear system with m equations and n variable $(m < n)$.

Note: When a right inverse or a left inverse for a non-singular non-square matrix exists, it is not unique.

Remark 5.1.3 For any non-singular square matrix A, left inverse and right inverse exists and it is equal to inverse of A, that is

$$A_R^{-1} = A_L^{-1} = A^{-1} = \frac{1}{\det(A)} adj(A).$$

The following theorem is from [3], but we give here another proof

Theorem 5.1.4 Every non-singular horizontal matrix A has a right inverse A_R^{-1} given by

$$A_R^{-1} = \frac{1}{\det(A)} adj(A)$$

Proof:

Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad adj(A) = \begin{bmatrix} M_{11} & M_{21} & \cdots & M_{j1} & \cdots & M_{m1} \\ M_{12} & M_{22} & \dots & M_{j2} & \cdots & M_{m2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ M_{1n} & M_{2n} & \dots & M_{jn} & \cdots & M_{mn} \end{bmatrix}$

Since A is non-singular, $det(A)$ is non-zero the matrix

$$B = \frac{1}{\det(A)} adj(A) = \frac{1}{\det(A)} \begin{bmatrix} M_{11} & M_{21} & \cdots & M_{j1} & \cdots & M_{m1} \\ M_{12} & M_{22} & \dots & M_{j2} & \cdots & M_{m2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ M_{1n} & M_{2n} & \dots & M_{jn} & \cdots & M_{mn} \end{bmatrix}$$

satisfies

$$A\,B = \frac{1}{\det(A)} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} M_{11} & M_{21} & \cdots & M_{j1} & \cdots & M_{m1} \\ M_{12} & M_{22} & \dots & M_{j2} & \cdots & M_{m2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ M_{1n} & M_{2n} & \dots & M_{jn} & \cdots & M_{mn} \end{bmatrix}$$

The $(i\,,j\,)^{th}$ element in the product matrix $A\,B$ is (**)

$$if \quad i = j, \qquad a_{i1}\,M_{j1} + a_{i2}M_{j2} + \cdots + a_{in}M_{jn} = \det(A) \qquad \text{(theorem 2.3.2)}$$

and, $if \quad i \neq j,$ we need to show that $a_{i1}\,M_{j1} + a_{i2}M_{j2} + \cdots + a_{in}M_{jn} = 0$

Consider the matrix B obtained from A by replacing the k^{th} row of A by its i^{th} row. Thus B is a matrix having two identical rows the i^{th} and k^{th} rows. Then $det(B) = 0$. Now expand $det(B)$ about the k^{th} row, the elements of the k^{th} row of B are $a_{i1}\,,\ a_{i2}\,, \dots,\ a_{in}$. The cofactors of the k^{th} row are $M_{k1}, M_{k2}\,, \dots, M_{kn}$.

We have, $0 = \det(B) = \ a_{i1}\,M_{k1} + a_{i2}M_{k2} + \cdots + a_{in}M_{kn}$

This means that

$$A\,B = \frac{1}{\det(A)} \begin{bmatrix} \det(A) & 0 & \cdots & 0 \\ 0 & \det(A) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \det(A) \end{bmatrix} = I_m$$

Hence, the matrix B is a right inverse of A. ■

In example 5.1.2 we have seen that a right inverse for $A = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 2 & 1 \end{bmatrix}$ is $A_R^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & -1/3 \\ -2 & -1/3 \end{bmatrix}$ here we compute an inverse of the same A by theorem 5.1.4 which is $A_R^{-1} = \begin{bmatrix} -1/2 & 3/2 \\ 1/2 & -1/2 \\ 0 & 1 \end{bmatrix}$. it is easy to see that the resulting inverse is another solution to the system (*).

Example 5.1.5 Let $A = \begin{bmatrix} 1 & 1 & 2 & 0 \\ 1 & 2 & 1 & 2 \\ 3 & 4 & 1 & 2 \end{bmatrix}$

to find a right inverse of A, A_R^{-1}, by applying theorem 5.1.4 we first compute cofactors of A , $\det(A) = -4$ and the result is

$$A_R^{-1} = \frac{1}{\det(A)} adj(A) = \frac{1}{-4} \begin{bmatrix} 2 & 5 & -1 \\ -2 & -3 & -1 \\ -2 & -1 & 1 \\ 2 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -1/2 & -5/4 & 1/4 \\ 1/2 & 3/4 & 1/4 \\ 1/2 & 1/4 & -1/4 \\ -1/2 & 1/4 & -1/4 \end{bmatrix}$$

Theorem 5.1.6 [3] Every non-singular vertical matrix A has a left inverse A_L^{-1}, such that

$$A_L^{-1} = \frac{1}{\det(A)} adj(A)$$

Proof: similar to the proof of theorem 5.1.4 ■

The following theorem describes the case that A has both a right inverse and a left inverse.

Theorem 5.1.7 If A is an $m \times n$ matrix such that both A_R^{-1} and A_L^{-1} exist, then $m = n$ (so A is square). Moreover, A is invertible and $A^{-1} = A_R^{-1} = A_L^{-1}$.

Proof: Let A be an $m \times n$ matrix

If $A\,A_R^{-1} = I_m$, then the equation $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution for every possible $\boldsymbol{b}$ in R^m

(given **b**, just let $\boldsymbol{x} = A_R^{-1}\boldsymbol{b}$, then $\boldsymbol{Ax} = \boldsymbol{A}(A_R^{-1}\boldsymbol{b}) = I_m\boldsymbol{b} = \boldsymbol{b}$.

Since **b** is arbitrary, in particular, for all $i = 1, \dots, m$, the system $\boldsymbol{Ax} = e_i$ (where $e_i = (0, \dots, 1, \dots, 0)^T$, the 1 in the ith component) has a solution, say r_i for every i, that is,

$$Ar_1 = e_1 \;, Ar_2 = e_2 \,, \dots, Ar_m = e_m.$$

Let $A_R^{-1} = [r_1 \;\; \cdots \;\; r_m]$. Then

$$AA_R^{-1} = A \cdot [r_1 \;\; \cdots \;\; r_m] = [A\,r_1 \;\; \cdots \;\; A\,r_m] = [e_1 \cdots e_m] = I_m.$$

Therefore A has a pivot position in every row. This force $m \leq n$, since every pivot position must be in a different column.

If $A_L^{-1}A = I_n$, consider the equation $\boldsymbol{Ax} = 0$. Then $A_L^{-1}\boldsymbol{A\,x} = A_L^{-1}\,0 = 0$.

But $A_L^{-1}\boldsymbol{A\,x} = I_n\,\boldsymbol{x} = \boldsymbol{x}$, so $\boldsymbol{x} = \boldsymbol{0}$. In other words, $\boldsymbol{Ax} = \boldsymbol{0}$ has a unique solution and therefore the columns of A must be linearly independent (see definition 1.9.3) and therefore each column must be a pivot position column. Since each pivot position must be in a different row, these forces $n \leq m$.

So, combining the two paragraphs gives that $m = n$. Since A is now known to be square, A is invertible and $A^{-1} = A_R^{-1} = A_L^{-1}$. ■

5.2 Properties for inverse and adjoint of non- square matrices.

In Chapter one we have known that taking inverse and transposing a matrix commute. (theorem 1.5.2). Here we write an example that asserts this fact for a non-square matrix.

Example 5.2.1 Let $\mathrm{A} = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 2 & 1 \end{bmatrix}$, a right inverse of A is $\mathrm{A}_R^{-1} = \begin{bmatrix} -1/2 & 3/2 \\ 1/2 & -1/2 \\ 0 & -1 \end{bmatrix}$,

$$(\mathrm{A}_R^{-1})^T = \begin{bmatrix} -1/2 & 1/2 & 0 \\ 3/2 & -1/2 & -1 \end{bmatrix}, \quad \text{also} \quad \mathrm{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 0 & 1 \end{bmatrix},$$

$$(\mathrm{A}^{\mathrm{T}})_L^{-1} = \begin{bmatrix} -1/2 & 1/2 & 0 \\ 3/2 & -1/2 & -1 \end{bmatrix} = (\mathrm{A}_R^{-1})^T$$

The following Corollaries says that this is true in general.

Corollary 5.2.2 If $A = (a_{ij})$ is non-singular $m \times n$ matrix, then

$$(\mathrm{A}^{\mathrm{T}})_L^{-1} = (\mathrm{A}_R^{-1})^T \quad \text{and} \quad (\mathrm{A}^{\mathrm{T}})_R^{-1} = (\mathrm{A}_L^{-1})^T$$

Proof: If $m < n$, then A^T is an $n \times m$

A_R^{-1} is a right inverse of A, we have $A\,A_R^{-1} = \mathrm{I}_m$

Taking transposes, we obtain

$(A\,A_R^{-1})^T = (\mathrm{I}_m)^T$

$(\mathrm{A}_R^{-1})^T\,A^T = \mathrm{I}_m$

These equations imply that

$(\mathrm{A}^{\mathrm{T}})_L^{-1} = (\mathrm{A}_R^{-1})^T$

The proof for the case $n < m$ is similar. ■

Corollary 5.2.3 Let A be a non-singular $m \times m$ matrix, and B be non-singular $m \times n$ matrix, then

$$(\mathrm{A\,B})_R^{-1} = \mathrm{B}_R^{-1}A^{-1} \quad \text{and} \quad (\mathrm{A\,B})_L^{-1} = \mathrm{B}_L^{-1}A^{-1}$$

Proof: If $m < n$, since A is an $m \times m$ matrix and B is an $m \times n$ matrix**,** we have A^{-1} is an $m \times m$ matrix and B_R^{-1} is an $n \times m$ matrix

Now,

$(\mathrm{A\,B})(\mathrm{B}_R^{-1}A^{-1}) = \mathrm{A}(\mathrm{B\,B}_R^{-1})\mathrm{A}^{-1} = \mathrm{A\,I}_m\ \mathrm{A}^{-1} = \mathrm{A}\ \ \mathrm{A}^{-1} = \mathrm{I}_m$

When $n < m$ we obtain

$(\mathrm{B}_L^{-1}A^{-1})(\mathrm{A\,B}) = \mathrm{B}_L^{-1}(\,\mathrm{A}^{-1}\,A)\mathrm{B} = \mathrm{B}_L^{-1}\,\mathrm{I}_m\mathrm{B} = \mathrm{B}_L^{-1}\,\mathrm{B} = \mathrm{I}_n$

Therefore, $(\mathrm{A\,B})_R^{-1} = \mathrm{B}_R^{-1}A^{-1}$ and $(\mathrm{A\,B})_L^{-1} = \mathrm{B}_L^{-1}A^{-1}$ ■

Example 5.2.4 Prove theorem 5.2.3 for the matrices $A = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 2 & 1 \end{bmatrix}$

Then, $B = \begin{bmatrix} 7 & 9 & 3 \\ 4 & 8 & 1 \end{bmatrix}$, $\quad adj(\,A\,B\,) = \begin{bmatrix} 7 & -6 \\ -3 & 4 \\ -4 & 2 \end{bmatrix}$

$$(\mathrm{A\,B})_R^{-1} = \begin{bmatrix} 7/10 & -6/10 \\ -3/10 & 4/10 \\ -4/10 & 2/10 \end{bmatrix} \quad \text{and} \quad \mathrm{A}^{-1} = \begin{bmatrix} -1/5 & 3/5 \\ 2/5 & -1/5 \end{bmatrix}$$

$$adj(B) = \begin{bmatrix} 1 & -3 \\ -1 & 1 \\ 0 & 2 \end{bmatrix}, \qquad \mathrm{B}_R^{-1} = \begin{bmatrix} -1/2 & 3/2 \\ 1/2 & -1/2 \\ 0 & -1 \end{bmatrix}$$

$$\mathrm{B}_R^{-1}\ A^{-1} = \begin{bmatrix} 7/10 & -6/10 \\ -3/10 & 4/10 \\ -4/10 & 2/10 \end{bmatrix} = (\mathrm{A\ B})_R^{-1}$$

Corollary 5.2.5 If k is nonzero real number and $A = (a_{ij})$ is a non-singular $m \times n$ matrix, then

$$(\mathrm{k\ A})^{-1} = \frac{1}{k}\ \mathrm{A}^{-1}$$

where the inverse is a right or a left inverse according to $m \le n$ or $m > n$

Proof: If $m \le n$, since A is an $m \times n$ matrix, A_R^{-1} is an $n \times m$ matrix and we have

$(\mathrm{k\ A})\left(\frac{1}{k}\ A_R^{-1}\right) = \mathrm{A}\ A_R^{-1} = I_m$, which gives

$$(\mathrm{k\ A})_R^{-1} = \frac{1}{k}\ A_R^{-1}$$

Also, if $n < m$, $\left(\frac{1}{k}\ A_L^{-1}\right)(\mathrm{k\ A}) = A_L^{-1}\ \mathrm{A} = I_n$ which gives

$$(\mathrm{k\ A})_L^{-1} = \frac{1}{k}\ A_L^{-1}$$ ■

We notice here that the properties of the inverse that are satisfied in square matrices (see theorem 1.5.2.b-c-d) are also satisfied in non-square matrices (see theorem 5.2.2, 5.2.3, 5.2.5)

The following theorem is given in [8], but we are providing our own proof.

Theorem 5.2.6 If $A = (a_{ij})$ is an $m \times n$ ($m < n$) matrix, then

$$\mathrm{A\ adj(A)} = \det(A)\ I_m$$

Proof: Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, $dj(A) = \begin{bmatrix} M_{11} & M_{21} & \cdots & M_{j1} & \cdots & M_{m1} \\ M_{12} & M_{22} & \dots & M_{j2} & \cdots & M_{m2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ M_{1n} & M_{2n} & \dots & M_{jn} & \cdots & M_{mn} \end{bmatrix}$

$$A\,adj(A) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} M_{11} & M_{21} & \cdots & M_{j1} & \cdots & M_{m1} \\ M_{12} & M_{22} & \dots & M_{j2} & \cdots & M_{m2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ M_{1n} & M_{2n} & \dots & M_{jn} & \cdots & M_{mn} \end{bmatrix}$$

The $(i\,,j\,)^{th}$ element in the product matrix $A\,adj(A)$ is

$$if \quad i = j, \qquad a_{i1}\,M_{j1} + a_{i2}M_{j2} + \cdots + a_{in}M_{jn} = \det(A)$$

and, $if \quad i \neq j, \qquad a_{i1}\,M_{j1} + a_{i2}M_{j2} + \cdots + a_{in}M_{jn} = 0$ (It was pre-established in theorem 5.1.4 **)

Hence, $A\;adj(A) = \begin{bmatrix} \det(A) & 0 & \cdots & 0 \\ 0 & \det(A) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \det(A) \end{bmatrix} = \det(A)\,I_m$ ■

We note that this property is valid for square matrices except that the switch here is not permissible. (see theorem 1.4.4.a).

i.e, $A\;adj(A) = \det(A)\,I_m \neq \big(adj(A)\big)A$

Example 5.2.7 Prove theorem 5.2.6 for the matrices $A = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 1 & -1 \end{bmatrix}$

Then, $adj(A) = \begin{bmatrix} 2 & 3 \\ 3 & 1 \\ 1 & -2 \end{bmatrix}$, and $det(A) = -7$

so, $A\;adj(A) = -7\,I_2$ but $\big(adj(A)\big)A \neq -7\,I_2$

All of the following corollaries in this section are proved by using theorem 5.2.6 as follows**:**

Corollary 5.2.8 If $A = (a_{ij})$ is singular $m \times n$ matrix, then $adj(A)$ is singular

Proof: by theorem 5.2.6**,** $\mathrm{A\,adj(A)} = \det(A)\,I_m$

Since, A is singular, then $\det(A) = 0$

So, $\mathrm{A\,adj(A)} = 0,$

If $A = 0$ (null and singular), then $adj(A) = 0$ and hence $adj(A)$ is singular too.

If $A \neq 0$ (non-null and singular), then A contains a non-null row, say the i^{th} row a_i' it follows that $a_i' \, adj(A) = 0$

Which implies that the rows of $adj(A)$ are linearly dependent (see definition 1.9.3), and hence $adj(A)$ is singular.

■

This property is satisfied in the case of square matrices.

Example 5.2.9 Prove theorem 5.2.8 for the matrices $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \end{bmatrix}$

then $adj(A) = \begin{bmatrix} -1 & 0 \\ 2 & 0 \\ -1 & 0 \end{bmatrix}$, and $det(A) = 0,$ $\det(adj(A)) = 0$

Corollary 5.2.10 If $A = (a_{ij})$ is an $m \times n$ $(m < n)$ non-singular matrix, then

$$(\mathrm{adj(A)})^{\mathrm{T}} = \mathrm{adj(A^T)}$$

Proof: By theorem 5.1.4**,** if $m < n,$ **,** then $\mathrm{adj(A)} = \det(A) \; A_R^{-1}$

Taking transposes, we obtain

$(\mathrm{adj(A)})^T = (\det(A) \; A_R^{-1})^T$

$(\mathrm{adj(A)})^{\mathrm{T}} = \det(A) \, (A_R^{-1})^T$ (1)

And, $\mathrm{adj(A}^T) = \det(A) \, (A^T)_L^{-1}$

$\mathrm{adj(A}^T) = \det(A) \; (A_R^{-1})^T$ (2)

From (1) and (2), we obtain

$(\mathrm{adj(A)})^{\mathrm{T}} = \mathrm{adj(A^T)}$ ■

This property is valid for singular matrices of all size (square or non-square)

Example 5.2.11 Let $A = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 1 & -1 \end{bmatrix}$

then $adj(A) = \begin{bmatrix} 2 & 3 \\ 3 & 1 \\ 1 & -2 \end{bmatrix}$, $(adj(A))^T = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 1 & -2 \end{bmatrix}$

$$A^T = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & -1 \end{bmatrix}, \quad adj(A^T) = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 1 & -2 \end{bmatrix} = (adj(A))^T$$

Corollary 5.2.12 If k is a scalar and $A = (a_{ij})$ is non-singular $m \times n$ matrix, then

$$\text{adj}(\text{k A}) = \text{k}^{m-1}\, \text{adj(A)}$$

Proof: By theorem 5.1.4, if $m \leq n$

(from corollaries 2.4.4, 5.2.5)

$$\text{adj}(\text{k A}) = det(\text{k A})(k\,A)_R^{-1} = k^m \det(A)\, \frac{1}{k}\, A_R^{-1}$$

$$= k^{m-1} \det(A)\; A_R^{-1} = k^{m-1} \text{adj(A)}$$

Similarly, if $n < m$. ■

This property is valid for singular matrices of all size (square or non-square)

Corollary 5.2.13 Let A be an $m \times m$ non-singular matrix, and B be an $m \times n$ non-singular matrix, then

$$adj(A\,B) = \text{adj(B) adj(A)}$$

Proof: By theorem 5.1.4 , if $m \leq n$

$adj(A\,B) = \det(\text{A B})(\text{A B})_R^{-1}$

$= \det(A) . \det(B) .\ B_R^{-1} A^{-1}$ (from theorem 3.1.2 , corollary 5.2.3)

$= \det(B) .\ B_R^{-1} . \det(A) . A^{-1}$

$= \text{adj(B)} . \text{adj(A)}$

Similarly, if $n < m$ ■

Example 5.2.14 Let $A = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 2 & 1 \end{bmatrix}$

then $A\,B = \begin{bmatrix} 7 & 9 & 3 \\ 4 & 8 & 1 \end{bmatrix}$, $\quad adj(\,A\,B\,) = \begin{bmatrix} 7 & -6 \\ -3 & 4 \\ -4 & 2 \end{bmatrix}$

$adj(B) = \begin{bmatrix} 1 & -3 \\ -1 & 1 \\ 0 & 2 \end{bmatrix}$, $\quad adj(A) = \begin{bmatrix} 1 & -3 \\ -2 & 1 \end{bmatrix}$

$$adj(B)\ \ adj(A) = \begin{bmatrix} 1 & -3 \\ -1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 7 & -6 \\ -3 & 4 \\ -4 & 2 \end{bmatrix} = adj(\,A\,B)$$

Example 5.2.15 Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 5 & 8 \\ 3 & 2 & 1 \end{bmatrix}$

Then, $\det(A) = 0$, $\quad det(B) = 0$, $\quad A\,B = \begin{bmatrix} 8 & 9 & 10 \\ 16 & 18 & 20 \end{bmatrix}$, $\quad \det(AB) = -4$,

$adj(\,A\,B\,) = \begin{bmatrix} -2 & 1 \\ 4 & -2 \\ -2 & -1 \end{bmatrix}$, $\quad adj(B) = \begin{bmatrix} 1 & 3 \\ -2 & -6 \\ 1 & 3 \end{bmatrix}$, $\quad adj(A) = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}$

$\det\big(adj(B)\big) = 0$, $\quad det(adj(A)) = 0$, $\quad det\big(adj(AB)\big) = -12$,

$$adj(B)\ \ adj(A) = \begin{bmatrix} -2 & 1 \\ 4 & -2 \\ -2 & -1 \end{bmatrix} = adj(\,A\,B)$$

We notice here that the properties of the adjoint of a matrix that are satisfied in square matrices (see theorem 1.4.4.d-e-f) are also satisfied in non-square matrices (see theorem 5.2.10, 5.2.12, 5.2.13).

5.3 Pseudo inverse of non-square matrices.

In section 5.2 we compute an inverse of a rectangular matrix using solution of a linear system and an adjoint of A, here we discuss another method which gives an inverse of A .

Theorem 5.3.1 $[11, p. 216\,]$ Let A be an $m \times n$ matrix, the null space of A is denoted by $\mathrm{N}(A)$. The dimension of the null space of A is called the nullity of A.

(1) $\mathrm{N}(A) = \mathrm{N}(A^T A)$

(2) $rank(A) = rank(A^T A)$

Proof: (1) show $\mathrm{N}(A) \subset \mathrm{N}(A^T A)$

Consider any $x \in \mathrm{N}(A)$, Then we have $\boldsymbol{Ax} = \boldsymbol{0}$ (see theorem 1.9.6)

Multiplying it by A^T from the left, we obtain

$A^T \boldsymbol{Ax} = A^T 0 = 0$

Thus $x \in \mathrm{N}(A^T A)$

and hence $\mathrm{N}(A) \subset \mathrm{N}(A^T A)$ $\quad (i)$

show $\mathrm{N}(A^T A) \subset \mathrm{N}(A)$

let $x \in \mathrm{N}(A^T A)$, thus we have $A^T \boldsymbol{Ax} = 0$

Multiplying it by x^T from the left, we obtain

$x^T A^T \boldsymbol{Ax} = x^T 0 = 0$

This implies that we have $0 = (\boldsymbol{Ax})^T \boldsymbol{Ax} = \|\boldsymbol{Ax}\|^2$

and the length of the vector $\boldsymbol{Ax}$ is zero, thus the vector $\boldsymbol{Ax} = \boldsymbol{0}$. Hence $x \in \mathrm{N}(A)$

hence $\mathrm{N}(A^T A) \subset \mathrm{N}(A)$ $\quad (ii)$

from (i) and (ii)

Hence, $\mathrm{N}(A) = \mathrm{N}(A^T A)$

(2) We use the rank-nullity theorem and obtain (see theorem 1.9.8)

$$rank(A) = n - dim\big(\mathrm{N}(A)\big)$$

$$= n - dim\big(\mathrm{N}(A^T A)\big)$$

$$= rank(A^T A)$$

(Note that the size of the matrix $A^T A$ is $n \times n$)

■

Definition 5.3.2 [4] The matrix $A_R^+ = A^T(A\,A^T\,)^{-1}$, when A is $m \times n\ (m \le n)$ and $(A\,A^T\,)^{-1}$ exists, and $rank\ (A) = m,$ is called the right Pseudo inverse of A.

Definition 5.3.3 [4] The matrix $A_L^+ = (\,A^T\,A)^{-1}A^T,$ when A is $m \times n\ (m > n)$ and $(\,A^T\,A)^{-1}$ exists, and $rank\ (A) = n,$ is called the left Pseudo inverse of A

Remark 5.3.4 The matrix $A_{m\times n}\ \ (m > n)$ has $rank = n,$ and therefore A^T also has $rank = m,$ A^TA is also of $rank = n$,(see theorem 5.3.1) since A^TA is a $n \times n$ matrix. it therefore has full rank and its inverse exists. (see theorem 1.9.14)

Note that $A\,A^T$ is a $m \times m$ matrix but its inverse does not exist.

Example 5.3.5 Let A = $\begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix},$ $\quad rank\ (A) = 2 = \ m = rank\ (AA^T)$

AA^T is invertible, $|A\,A^T| = 33 - 9 = 24,$ and $(A\,A^T)^{-1} = \frac{1}{24}\begin{bmatrix} 11 & -3 \\ -3 & 3 \end{bmatrix}$

$$A_R^{-1} = A_R^+ = A^T(A\,A^T\,)^{-1} = \frac{1}{24}\begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 11 & -3 \\ -3 & 3 \end{bmatrix} = \frac{1}{24}\begin{bmatrix} 2 & 6 \\ 14 & -6 \\ 8 & 0 \end{bmatrix}$$

$$A\,A_R^{-1} = \frac{1}{24}\begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 6 \\ 14 & -6 \\ 8 & 0 \end{bmatrix} = \ I_2$$

Since $\det(A\,) = 0,$ we cannot use adjoint method

Example 5.3.6 Let A = $\begin{bmatrix} 1 & -2 \\ -2 & 1 \\ 1 & 1 \end{bmatrix},$ $\quad rank\ (A) = 2 = \ n = rank\ (A^T\,A)$

$A^T\,A$ is invertible, $|\,A^TA\,| = 36 - 9 = 27,$ and $(\,A^TA)^{-1} = \frac{1}{27}\begin{bmatrix} 6 & 3 \\ 3 & 6 \end{bmatrix}$

$$A_L^{-1} = A_L^+ = (\,A^T\,A)^{-1}A^T = \frac{1}{27}\begin{bmatrix} 6 & 3 \\ 3 & 6 \end{bmatrix}\begin{bmatrix} 1 & -2 & 1 \\ -2 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & \frac{-1}{3} & \frac{1}{3} \\ \frac{-1}{3} & 0 & \frac{1}{3} \end{bmatrix}$$

$$A_L^{-1}A = \begin{bmatrix} 0 & \frac{-1}{3} & \frac{1}{3} \\ \frac{-1}{3} & 0 & \frac{1}{3} \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -2 & 1 \\ 1 & 1 \end{bmatrix} = I_2$$

Chapter Six

Applications to linear systems of equations $\boldsymbol{A\,X = B}$

In this chapter we discuss some results concerning the solutions of a linear system $\boldsymbol{Ax = b}$ using inverses as well as the pseudo-inverse and adjoint of a rectangular $m \times n$ matrix A.

The solution of the system can be expressed as $\boldsymbol{x} = A^{-1}\,\boldsymbol{b}$ where A^{-1} is the inverse of A. when matrix A is of order $m \times n\,(m < n)$ because if A_R^{-1} is the right inverse of A then we have $\boldsymbol{A}A_R^{-1}\boldsymbol{x} = \boldsymbol{b}A_R^{-1}$ which yields $I_m\boldsymbol{x} = \boldsymbol{b}A_R^{-1}$ implies that $\boldsymbol{x} = \boldsymbol{b}A_R^{-1}$ but $\boldsymbol{b}A_R^{-1}$ is not defined. Since we can't find an actual solution to the system, we will now try to find solution to the system.

6.1 Solving a linear system Using pseudo inverse

If A is an $m \times n$ matrix, then the linear system $\boldsymbol{A\,x = b}$ is a system of m equations in n unknowns.

Theorem 6.1.1[5] Let $\boldsymbol{A\,x = b}$ be a linear system with A an $m \times n\,(m > n)$ matrix. If A^TA is invertible, then the solution can be given by $x = \boldsymbol{A_L^{-1}b}$ where A_L^{-1} is a left pseudo-inverse of A

Proof: Let A be an $m \times n\,(m > n)$ and A^TA is invertible, then $(A^TA)^{-1}$ exists, and we can multiply $\boldsymbol{A\,x = b}$ by A^T on both sides. Obtaining $A^T\boldsymbol{A\,x} = A^T\boldsymbol{b}$

So, multiply $\mathrm{A^T}\mathbf{A\,x} = \mathrm{A^T}\mathbf{b}$ by $(\mathrm{A^T\,A})^{-1}$ on the both sides,

$$\begin{aligned}(A^TA)^{-1}\,(A^T\boldsymbol{A\,x}) &= (A^TA)^{-1}\,(A^T\,\boldsymbol{b})\\ I_n\,\boldsymbol{x} &= (A^TA)^{-1}\,(A^T\,\boldsymbol{b})\end{aligned}$$

$$\begin{aligned}\boldsymbol{x} &= A_L^{-1}(\,A^T)^{-1}A^T\,\boldsymbol{b}\\ \boldsymbol{x} &= A_L^{-1}\,\boldsymbol{b}\end{aligned}$$

■

Example 6.1.2 Find a solution to the system

$$\begin{aligned} x_1 + 3x_2 &= -2 \\ 3x_1 - x_2 &= 4 \\ 2x_1 + 2x_2 &= 0 \end{aligned}$$

Solution: Let $\boldsymbol{x}$ be a solution to $\boldsymbol{Ax} = \boldsymbol{b}$, (if exists)

$\boldsymbol{A} = \begin{bmatrix} 1 & 3 \\ 3 & -1 \\ 2 & 2 \end{bmatrix}$ is the coefficient matrix, and $\boldsymbol{b} = \begin{bmatrix} -2 \\ 4 \\ 0 \end{bmatrix}$ is the constant vector

$$\boldsymbol{x} = (A^T A)^{-1} A^T \mathbf{b}$$

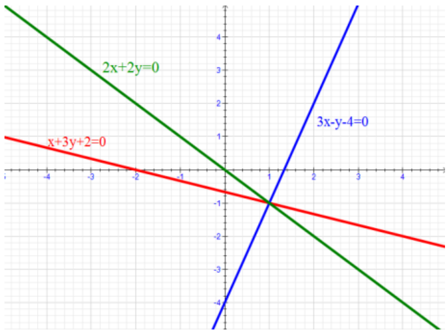
$$A^T A = \begin{bmatrix} 1 & 3 & 2 \\ 3 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 3 & -1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 14 & 4 \\ 4 & 14 \end{bmatrix}$$

So, $(A^T A)^{-1} = \frac{1}{90} \begin{bmatrix} 7 & -2 \\ -2 & 7 \end{bmatrix}$

$$(A^T A)^{-1} A^T = \frac{1}{90} \begin{bmatrix} 7 & -2 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 3 & -1 & 2 \end{bmatrix} = \frac{1}{90} \begin{bmatrix} 2 & 23 & 10 \\ 19 & -13 & 10 \end{bmatrix}$$

$$\boldsymbol{x} = (A^T A)^{-1} A^T \mathbf{b} = \frac{1}{90} \begin{bmatrix} 2 & 23 & 10 \\ 19 & -13 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 4 \\ 0 \end{bmatrix} = \frac{1}{90} \begin{bmatrix} 90 \\ -90 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Note that $\mathrm{x} = (1, -1)$ represents the intersection point of the three straight lines shown in the graph.



We note that $rank(A) = 2$, which means that there is a unique solution to the original system occurs at the intersection of these three lines

Note that : if $A^T A$ is invertble, then the only possibility for $\boldsymbol{Ax} = \boldsymbol{b}$ are either unique solution or no solution.

Example 6.1.3 Find a solution to the system

$$\begin{aligned} x_1 + 3\,x_2 &= 5 \\ x_1 - \ x_2 &= 1 \\ x_1 + \ x_2 &= 0 \end{aligned}$$

Solution: Let x be a solution to $\boldsymbol{Ax} = \boldsymbol{b}$, (if exists)

$A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$ is the coefficient matrix and $\boldsymbol{b} = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$ is the constant vector

$$\boldsymbol{x} = (\,A^T\,A)^{-1}A^T\mathbf{b}$$

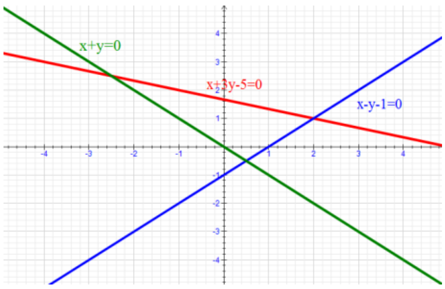
We calculate $A^T\,A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix}$

Next, $(A^TA)^{-1} = \frac{1}{24}\begin{bmatrix} 11 & -3 \\ -3 & 3 \end{bmatrix}$

$$(\,A^T\,A)^{-1}A^T = \frac{1}{24}\begin{bmatrix} 11 & -3 \\ -3 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix} = \frac{1}{24}\begin{bmatrix} 2 & 14 & 8 \\ 6 & -6 & 0 \end{bmatrix}$$

$$\boldsymbol{x} = (\,A^T\,A)^{-1}A^T\mathbf{b} = \frac{1}{24}\begin{bmatrix} 2 & 14 & 8 \\ 6 & -6 & 0 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{24}\begin{bmatrix} 24 \\ 24 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Note that x $= (1\,,1)$ doesn't satisfy the three equations, that is, it doesn't represent the intersection point of the three straight lines. This coincides with the second case in the note.



Theorem 6.1.4 [5] Let $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$ be a linear system with A an $m \times n\,(\,m < n)$ matrix. If $A\,A^T$ is invertible, then the solution can be given by $\boldsymbol{x} = A_R^{-1}\boldsymbol{b}$ where A_R^{-1} is aright pseudo-inverse of A

Proof: Let A be an $m \times n\,(\,m < n)$ and $A\,A^T$ is invertible, then $(\,A\,A^T)^{-1}$ exists, and we can multiply $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$ by $(A\,A^T)^{-1}$ on both sides, obtaining

$$\begin{aligned}(A\,A^T)^{-1}\boldsymbol{A}\,\boldsymbol{x} &= (A\,A^T)^{-1}\,\boldsymbol{b}\\ (A^T)_R^{-1}\,A_L^{-1}\boldsymbol{A}\,\boldsymbol{x} &= (A\,A^T)^{-1}\,\boldsymbol{b}\\ (A^T)_R^{-1}\,I_n\,\boldsymbol{x} &= (A\,A^T)^{-1}\,\boldsymbol{b} \qquad (1)\end{aligned}$$

So, multiply (1) by A^T on both sides, we obtain

$$\boldsymbol{x} = A^T(A\,A^T)^{-1}\,\mathbf{b}$$ ■

Example 6.1.5 Find a solution to the system

$$\begin{aligned}2x_1 + 3\,x_2 - 2x_3 &= 4\\ -6\,x_1 - 8\,x_2 + 6x_3 &= 1\end{aligned}$$

Solution: Let **x** be a solution to $\boldsymbol{Ax} = \boldsymbol{b}$, (if exists)

$$\boldsymbol{x} = (\,A^T\,A)^{-1}A^T\mathbf{b}$$

We calculate $A\,A^T = \begin{bmatrix} 2 & 3 & -2\\ -6 & -8 & 6\end{bmatrix}\begin{bmatrix} 2 & -6\\ 3 & -8\\ -2 & 6\end{bmatrix} = \begin{bmatrix} 17 & -48\\ -48 & 136\end{bmatrix}$

Next, $(A^TA)^{-1} = \frac{1}{8}\begin{bmatrix}136 & 48\\ 48 & 17\end{bmatrix}$

$$A^T(A\,A^T)^{-1} = \frac{1}{8}\begin{bmatrix} 2 & -6\\ 3 & -8\\ -2 & 6\end{bmatrix}\begin{bmatrix}136 & 48\\ 48 & 17\end{bmatrix} = \frac{1}{8}\begin{bmatrix}-16 & -6\\ 24 & 8\\ 16 & 6\end{bmatrix}$$

$\boldsymbol{x} = A^T(A\,A^T)^{-1}\,\mathbf{b} = \frac{1}{8}\begin{bmatrix}-16 & -6\\ 24 & 8\\ 16 & 6\end{bmatrix}\begin{bmatrix}4\\1\end{bmatrix} = \frac{1}{8}\begin{bmatrix}-70\\ 104\\ 70\end{bmatrix} = \begin{bmatrix}\frac{-70}{8}\\ 13\\ \frac{70}{8}\end{bmatrix}$ is a solution

The general solution in vector form as $\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix}\frac{-35}{4}\\ 13\\ \frac{35}{4}\end{bmatrix} + t\begin{bmatrix}1\\0\\1\end{bmatrix}$

This is a parameter vector equation of the line of intersection L of the three lines. The coordinates of each of $L's$ points make one of the infinitely many solutions of the system.

Theorem 6.1.6 [6]. (**General solution**)

Let $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$ be a full rank underdetermined system (A an $m \times n\,(\,m < n)$ matrix.). then the solution set is given by

$$\boldsymbol{x} = A^T(A\,A^T)^{-1}\boldsymbol{b} \;+\; (I - A^T(A\,A^T)^{-1}A)y$$

$$\boldsymbol{x} \;= A_R^{-1}\boldsymbol{b} + (I_n - A_R^{-1}A)y, \qquad (1)$$

where y is an arbitrary vector in $R^{n\times 1}$.

Proof: To verify that (1) is a solution, pre-multiply by A

$$\boldsymbol{Ax} \;=\; A\,A_R^{-1}\,\boldsymbol{b} \;+\; A\,(I - A_R^{-1}A)y$$

$$= I_m\,\boldsymbol{b} \;+\; (A \;-\; A\,A_R^{-1}\,A)y \qquad \text{by hypothesis}$$

$$= \quad \boldsymbol{b}\;, \qquad \text{since} \quad A\,A_R^{-1}\,A \;= I_m\,A_{m\times n} = A$$

That all solutions are of this form can be seen as follows.

Let $\boldsymbol{z}$ be an arbitrary solution of $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$, i.e., $\boldsymbol{A}\,\boldsymbol{z} \;=\; \boldsymbol{b}$. Then we can write

$$z = A_R^{-1}\boldsymbol{A}\,\boldsymbol{z} + (I - A_R^{-1}A)z = A_R^{-1}\boldsymbol{A}\,\boldsymbol{z} + z - A_R^{-1}\boldsymbol{A}\,\boldsymbol{z}$$

So that any solution $\boldsymbol{x}$ of $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$ is given by (1) with $\boldsymbol{y} = \boldsymbol{z}$ ■

Remark: When A is square and nonsingular, $A_R^{-1} = A^{-1}$ and so $(I - A_R^{-1}A) = 0$

Thus, there is no “arbitrary” component, leaving only the unique solution $x = A^{-1}b$

Note that: when A is not full rank, then the above theorem cannot be used.

6.2 Using adjoint to solve a system of linear equations

Let A be a non-singular $m \times n$ matrix. In this section we use adjoint of A is find a solution for $\boldsymbol{Ax} = \boldsymbol{b}$.

Let $A_{m\times n}$, $m \le n$, be nonsingular, then A_R^{-1} exists and we can multiply $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$ by A_R^{-1} on both sides, obtaining $\boldsymbol{A}\,A_R^{-1}\boldsymbol{x} = A_R^{-1}\boldsymbol{b}$

Then $I_m\,\boldsymbol{x} = A_R^{-1}\boldsymbol{b}$, but $A_R^{-1} = \frac{1}{\det(A)} adj(A)$ (see theorem 5.1.4)

So, $\boldsymbol{x} = \frac{1}{\det(A)} adj(A)\,\boldsymbol{b}$

assume $m > n$, and A is nonsingular, then A_L^{-1} exists and we can multiply both sides of $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$ by A_L^{-1}, obtaining $A_L^{-1}\boldsymbol{A}\,\boldsymbol{x} = A_L^{-1}\boldsymbol{b}$

That is $I_n\,\boldsymbol{x} = A_L^{-1}\boldsymbol{b}$, but $A_L^{-1} = \frac{1}{\det(A)} adj(A)$ (see theorem 5.1.6)

So, $\boldsymbol{x} = \frac{1}{\det(A)} adj(A)\,\boldsymbol{b}$

Example 6.2.1 Find a solution to the system

$$\begin{aligned} x_1 &+ 3\,x_2 = -2 \\ 3\,x_1 &- \ \ x_2 = 4 \\ 2\,x_1 &+ 2\,x_2 = 0 \end{aligned}$$

Solution: Let $\boldsymbol{x}$ be a solution to $\boldsymbol{Ax} = \boldsymbol{b}$, (if exists)

$A = \begin{bmatrix} 1 & 3 \\ 3 & -1 \\ 2 & 2 \end{bmatrix}$ is the coefficient matrix, and $\boldsymbol{b} = \begin{bmatrix} -2 \\ 4 \\ 0 \end{bmatrix}$ is the constant vector

$det(A) = 2$, since A is 3×2, A_L^{-1} exists, $\boldsymbol{x} = \frac{1}{\det(A)} adj(A)\,\boldsymbol{b}$

So, we calculate $adj(A) = \begin{bmatrix} -3 & -1 & 4 \\ -1 & -1 & 2 \end{bmatrix}$

Next, $A_L^{-1} = \frac{1}{\det(A)} adj(A) = \frac{1}{2}\begin{bmatrix} -3 & -1 & 4 \\ -1 & -1 & 2 \end{bmatrix}$

$$\boldsymbol{x} = A_L^{-1}\boldsymbol{b} = \frac{1}{2}\begin{bmatrix} -3 & -1 & 4 \\ -1 & -1 & 2 \end{bmatrix}\begin{bmatrix} -2 \\ 4 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

We note that in this example, adjoint method and pseudo method give the same solution (see example 6.1.2).

Example 6.2.2 Find a solution to the system

$$\begin{aligned} 2x_1 &+ 3\,x_2 - 2x_3 = 4 \\ -6x_1 &- 8\,x_2 + 6x_3 = 1 \end{aligned}$$

Solution: Let $\boldsymbol{x}$ be a solution to $\boldsymbol{Ax} = \boldsymbol{b}$, (if exists)

Hold by $A = \begin{bmatrix} 2 & 3 & -2 \\ -6 & -8 & 6 \end{bmatrix}$ is the coefficient matrix, and $\boldsymbol{b} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ is the constant vector. $det(A) = 4$, since A is 2×3, A_R^{-1} exists, $\boldsymbol{x} = \frac{1}{\det(A)} adj(A)\,\boldsymbol{b}$

But, $adj(A) = \begin{bmatrix} -14 & -5 \\ 12 & 4 \\ 2 & 1 \end{bmatrix}$

Next, $A_R^{-1} = \frac{1}{\det(A)} adj(A) = \frac{1}{4}\begin{bmatrix} -14 & -5 \\ 12 & 4 \\ 2 & 1 \end{bmatrix}$

$$\boldsymbol{x} = \frac{1}{\det(A)} adj(A)\,\boldsymbol{b} = \frac{1}{4}\begin{bmatrix} -14 & -5 \\ 12 & 4 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} -61 \\ 52 \\ 9 \end{bmatrix}$$

That is, $\mathrm{x} = \left(\frac{-71}{4}, \frac{52}{4}, \frac{9}{4}\right)$

Here the solution we have obtained using adjoint method is different from the pseudo method solution (see example 6.1.5) and both of them are members from the general solution given by

$$x = A^T(A\,A^T)^{-1}b + (I - A^T(A\,A^T)^{-1}A)y = A_R^{-1}b + (I - A_R^{-1}A)y$$

$$x = \begin{bmatrix} \frac{-35}{4} \\ 13 \\ \frac{35}{4} \end{bmatrix} + \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{8}\begin{bmatrix} -16 & -6 \\ 24 & 8 \\ 16 & 6 \end{bmatrix}\begin{bmatrix} 2 & 3 & -2 \\ -6 & -8 & 6 \end{bmatrix}\right)\begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

$$x = \begin{bmatrix} \frac{-35}{4} \\ 13 \\ \frac{35}{4} \end{bmatrix} + \frac{1}{2}t\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \text{Choose } t = -13, \quad \mathrm{x} = \left(\frac{-71}{4}, \frac{52}{4}, \frac{9}{4}\right).$$

6.3 Cramer's rule for nonsingular $m \times n$ matrices

In linear algebra, Cramer's rule (see theorem 1.7.4) gives an explicit formula for the solution of a system of linear equations with as many equations as unknowns. That is, for the solution of a system with a square matrix and provided that the coefficient matrix is invertible, Cramer's rule offers a simple and a convenient formula for the solution. In this section we want to generalize this method for an $m < n$ system of linear equations. As in the usual method of Cramer's, the result for rectangular matrices uses the minors of a matrix. We also use the results in order to solve a matrix equation. in the case of systems with an infinite number of solutions, get the final formula for calculating the unknowns by the minors of the augmented matrix of the system.

The key to Cramer's Rule is replacing the variable column of interest with the constant column and calculating the determinants.

Theorem 6.3.1 Consider the following linear system of m equations in n unknowns

$$\begin{array}{ccccccc} a_{11}x_1 + & a_{12}x_2 & + \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 + & a_{22}x_2 & + \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 & + \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

$A = [a_{ij}]$ be the coefficient matrix, $b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$. If $\det(A) \neq 0$. Then

$$x_1 = \frac{\det(A_1)}{\det(A)} \quad , \quad x_2 = \frac{\det(A_2)}{\det(A)} \quad , \quad .... \quad , \quad x_n = \frac{\det(A_n)}{\det(A)}.$$

Is a solution for the system $\boldsymbol{Ax} = \boldsymbol{b}$, where A_j is the matrix obtained from A by replacing the j^{th} column of A by $\boldsymbol{b}$.

proof: We look at the linear system $\boldsymbol{Ax} = \boldsymbol{b}$.

$$\boldsymbol{x} = A^{-1}\,\boldsymbol{b}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{\det(A)}\,(adj\ A)\,\boldsymbol{b}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{m1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{mn} \end{bmatrix} \ \boldsymbol{b}$$

$$\begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \frac{A_{11}}{\det(A)} & \frac{A_{21}}{\det(A)} & \cdots & \frac{A_{m1}}{\det(A)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_{1j}}{\det(A)} & \frac{A_{2j}}{\det(A)} & \cdots & \frac{A_{mj}}{\det(A)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_{1n}}{\det(A)} & \frac{A_{2n}}{\det(A)} & \cdots & \frac{A_{mn}}{\det(A)} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

This means that

$$x_j = \frac{A_{1j}}{\det(A)} b_1 + \frac{A_{2j}}{\det(A)} b_2 + \cdots + \frac{A_{mj}}{\det(A)} b_m \quad (1 \le j \le \mathrm{n})$$

Where

$$A_j = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j-1} & b_1 & a_{1j+1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j-1} & b_2 & a_{2j+1} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj-1} & b_m & a_{nj+1} & \dots & a_{mn} \end{bmatrix}$$

If we evaluate $\det(A_j)$ be expanding about the j^{th} column, we find that

$$\det(A_j) = A_{1j}b_1 + A_{2j}b_2 + \cdots + A_{mj}b_m$$

Hence

$$x_j = \frac{\det(A_j)}{\det(A)}.$$

For $j = 1, 2, \dots, n$. In this expression for x_j, the determinant of A_j, $\det(A_j)$, can be calculated by any method. It was only in the derivation of the expression for x_j that we had to evaluate it by expanding about the j^{th} column. ■

Example 6.3.2. Find a solution to the system

$$\begin{aligned} 2x_1 + 3x_2 - 2x_3 &= 4 \\ -6x_1 - 8x_2 + 6x_3 &= 1 \end{aligned}$$

Solution: Let $\boldsymbol{x}$ be a solution to $\boldsymbol{Ax} = \boldsymbol{b}$, (if exists)

Here the coefficient matrix is $A = \begin{bmatrix} 2 & 3 & -2 \\ -6 & -8 & 6 \end{bmatrix}$, $det(A) = 4$, and $\boldsymbol{b} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ is the constant vector.

$$A_1 = \begin{bmatrix} 4 & 3 & -2 \\ 1 & -8 & 6 \end{bmatrix},\ \det(A_1) = -59 \quad \Rightarrow \quad x_1 = \frac{\det(A_1)}{\det(A)} = {-59}/{4}$$

$$A_2 = \begin{bmatrix} 2 & 4 & -2 \\ -6 & 1 & 6 \end{bmatrix},\ \det(A_2) = 52 \quad \Rightarrow \quad x_2 = \frac{\det(A_2)}{\det(A)} = {52}/{4}$$

$$A_3 = \begin{bmatrix} 2 & 3 & 4 \\ -6 & -8 & 1 \end{bmatrix},\ \det(A_3) = 11 \quad \Rightarrow \quad x_3 = \frac{\det(A_3)}{\det(A)} = {11}/{4}$$

A solution given by Cramer's rule is $x = \left({-59}/{4},\ 13,\ {11}/{4}\right)$

We note that the general solution for this example given by Theorem 6.1.6 is

$$x = A^T(A\,A^T)^{-1}b + (I - A^T(A\,A^T)^{-1}A)y = A_R^{-1}b + (I - A_R^{-1}A)y$$

$$x = \begin{bmatrix} \frac{-35}{4} \\ 13 \\ \frac{35}{4} \end{bmatrix} + \frac{1}{2}t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Choose $t = -12$, $x = \left({-59}/{4}, 13, {11}/{4}\right)$

Theorem 6.3.3 [10]. **(Generalization of Cramer's rule)**

For all $b \in R^n$ and A an $m \times n$ the system $\boldsymbol{Ax = b}$ is solvable if and only if $det(AA^T) \neq 0$.

Moreover, one solution for this equation is given by $x = A^T(A\,A^T)^{-1}b$, where A^T is the transpose of A.

Also, this solution coincides with the Cramer's rule formula when $n = m$. In fact, this formula is given as follows:

$$x_j = \sum_{i=1}^{m} a_{ij} \frac{det\big((A\,A^T)\big)_i}{det(A\,A^T)} \quad, j = 1, 2, 3, \dots, n,$$

where $(A\,A^T)_i$ is the matrix obtained by replacing the entries in the j^{th} column of $A\,A^T$ by the entries in the matrix $\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$

Proof: The matrix A gives a linear transformation $T_A : R^m \to R^n$ (see definition 1.9.11) and its transpose A^T gives the linear transformation, adjoint operator, such that $T_{A^T}: R^n \to R^m$.

The system $\boldsymbol{Ax = b}$ is solvable for all $b \in R^m$, if and only if, the operator A is onto (see theorem 1.9.15).

i.e, $Range(A) = R^n$ Hence, from the lemma 1.8.15 there exists $\gamma > 0$ such that $\|A^T z\|_{R^m} \geq \gamma \|z\|_{R^n}, z \in R^n$.

Therefore,

$\langle A\,A^T z, z\rangle \geq \gamma^2 \|z\|^2{}_{R^n}, z \in R^n$.

This implies that $T_{A\,A^T}: R^n \to R^n$, and $Rang(A\,A^T) = R^n$ and $Ker\,(AA^T) = \{0\}$

From the Theorem 1.9.15, lemma 1.9.16 then $A\,A^T$ is one to one.

Since $A\,A^T$ is a $n \times n$ matrix, from the Theorem 1.9.14 then $det(A\,A^T) \neq 0$.

Suppose now that $det(A\,A^T) \neq 0$. Then $(A\,A^T)^{-1}$ exists and given $b \in R^n$ we can see that $x = A^T\,(A\,A^T)^{-1}b$ is a solution of $\boldsymbol{A\,x = b}$.

Now, since $z = (A\,A^T)^{-1}b$ is the only solution of the equation $(A\,A^T)\,z = b$, then from theorem 6.3.1 (Cramer's rule) we obtain that :

$$z_1 = \frac{\det((A\,A^T)_1)}{\det(A\,A^T)} \quad , \quad z_2 = \frac{\det((A\,A^T)_2)}{\det(A\,A^T)} \quad , \quad .... \quad , \quad z_n = \frac{\det((A\,A^T)_n)}{\det(A\,A^T)}.$$

Where $(A\,A^T)_i$ is the matrix obtained by replacing the entries in the j^{th} column of $A\,A^T$ by the entries in the matrix $\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$.

Then, the solution $x = A^T(A\,A^T)^{-1}\,b = A^T z$ of $\boldsymbol{A\,x = b}$ can be written as follows

$$x_j = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{21} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} \frac{\det((A\,A^T)_1)}{\det(A\,A^T)} \\ \frac{\det((A\,A^T)_2)}{\det(A\,A^T)} \\ \vdots \\ \frac{\det((A\,A^T)_m)}{\det(A\,A^T)} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_{j,1} \frac{det\left((A\,A^T)_j\right)}{det(A\,A^T)} \\ \sum_{j=1}^{n} a_{j,2} \frac{det\left((A\,A^T)_j\right)}{det(A\,A^T)} \\ \vdots \\ \sum_{j=1}^{n} a_{j,m} \frac{det\left((A\,A^T)_j\right)}{det(A\,A^T)} \end{bmatrix}$$

Example 6.3.4 Find a solution to the system

$$\begin{aligned} 2x_1 \; + \; 3\,x_2 \; - \; 2x_3 \; &= \; 4 \\ -6x_1 \; - \; 8\,x_2 \; + \; 6x_3 \; &= \; 1 \end{aligned}$$

Solution: Let x be a solution to $\boldsymbol{Ax = b}$, (if exists)

Here the coefficient matrix is $A = \begin{bmatrix} 2 & 3 & -2 \\ -6 & -8 & 6 \end{bmatrix}$, $det(A) = 4$, and $b = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ is the constant vector.

$$A\,A^T = \begin{bmatrix} 2 & 3 & -2 \\ -6 & -8 & 6 \end{bmatrix} \begin{bmatrix} 2 & -6 \\ 3 & -8 \\ -2 & 6 \end{bmatrix} = \begin{bmatrix} 17 & -48 \\ -48 & 136 \end{bmatrix} \quad \rightarrow \quad det(A\,A^T) = 8$$

$$(A\,A^T)_1 = \begin{bmatrix} 4 & -48 \\ 1 & 136 \end{bmatrix} \quad \rightarrow \quad det(A\,A^T)_1 = 592$$

$$(A\,A^T)_2 = \begin{bmatrix} 17 & 4 \\ -48 & 1 \end{bmatrix} \quad \rightarrow \quad det(A\,A^T)_2 = 209$$

$$x_1 = \frac{a_{11}det(A\,A^T)_1}{det(A\,A^T)} + \frac{a_{21}det(A\,A^T)_2}{det(A\,A^T)} = \frac{2 \times 592}{8} + \frac{-6 \times 209}{8} = \frac{-70}{8}$$

$$x_2 = \frac{a_{12}det(A\,A^T)_1}{det(A\,A^T)} + \frac{a_{22}det(A\,A^T)_2}{det(A\,A^T)} = \frac{3 \times 592}{8} + \frac{-8 \times 209}{8} = \frac{104}{8} = 13$$

$$x_3 = \frac{a_{13}det(A\,A^T)_1}{det(A\,A^T)} + \frac{a_{23}det(A\,A^T)_2}{det(A\,A^T)} = \frac{-2 \times 592}{8} + \frac{6 \times 209}{8} = \frac{70}{8}$$

We find that the solution of the system in this way is the same as pseudo solution given in example 6.2.2 which comes from the general solution (see theorem 6.1.6)

$$x = \begin{bmatrix} \frac{-35}{4} \\ 13 \\ \frac{35}{4} \end{bmatrix} + \frac{1}{2}t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \text{at} \quad t = 0$$

That is both a Generalization of Cramer's rule and pseudo method give the same solution.

6.4 Particular Cases and Examples

In this section we shall consider some particular cases and examples to illustrate the results of what we have done in the previous sections especially in applying pseudo inverse (theorem 6.1.1) to some certain examples.

Example 6.4.1 [10]. Consider the following particular case of the system $\boldsymbol{Ax = b}$

$$a_{11}x_1 + a_{12}x_2 + \cdots . + a_{1n}x_n = b. \qquad (1)$$

In this case $m = 1$ and $A = [a_{11}, a_{12}, \ldots ., a_{1n}]$.

Then, if we define the column vector $I_1 = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix}$,

$$A\,A^T = [a_{11} \quad a_{12} \quad \cdots \quad a_{1n}] \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix} = \|I_1\|^2. \quad \text{(see definition 1.8.2)}$$

Then, $(A\,A^T\,)^{-1}\boldsymbol{b} \;=\; \frac{1}{\|I_1\|^2}\,\boldsymbol{b}$ and

$$\boldsymbol{x} = A^T\,(A\,A^T\,)^{-1}\boldsymbol{b} = \frac{1}{\|I_1\|^2}\,\boldsymbol{b} \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix} = \begin{bmatrix} a_{11}\,b\,\|I_1\|^{-2} \\ a_{12}b\,\|I_1\|^{-2} \\ \vdots \\ a_{1n}b\,\|I_1\|^{-2} \end{bmatrix}.$$

Therefore, a solution of the system (1) is given by:

$$x_j = \frac{a_{1j}\,b}{\|I_1\|^2} = \frac{a_{1j}\,b}{\sum_{j=1}^{n} a_{1j}^2}\,, \quad j = 1,2,\dots,n$$

Example 6.4.2 [10].

Here we apply pseudo inverse to the case $m = 2$, for any natural n in $\boldsymbol{Ax} = \boldsymbol{b}$

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots. + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots. + a_{2n}x_n &= b_2 \end{aligned} \qquad (2)$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \end{bmatrix}$$

Then, let I_1 , I_2 be the column vectors

$$I_1 = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix}, \quad I_2 = \begin{bmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{bmatrix}.$$

Then,

$$A\,A^T = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ \vdots & \vdots \\ a_{1n} & a_{2n} \end{bmatrix} = \begin{bmatrix} \|I_1\|^2 & \langle I_1, I_2 \rangle \\ \langle I_2, I_1 \rangle & \|I_2\|^2 \end{bmatrix}.$$

$$(A\,A^T)^{-1} = \frac{1}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2}\begin{bmatrix} \|I_2\|^2 & -\langle I_1, I_2\rangle \\ -\langle I_2, I_1\rangle & \|I_1\|^2 \end{bmatrix}.$$

Hence, from the formula $\boldsymbol{x_j} = A^T\,(A\,A^T\,)^{-1}\boldsymbol{b}$ we obtain that:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2}\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ \vdots & \vdots \\ a_{1n} & a_{2n} \end{bmatrix}\begin{bmatrix} \|I_2\|^2 & -\langle I_1, I_2\rangle \\ -\langle I_2, I_1\rangle & \|I_1\|^2 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Therefore, a solution of the system (2) is given by :

$$x_1 = a_{11}\frac{b_1\|I_2\|^2 - b_2\langle I_1, I_2\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2} + a_{21}\frac{b_2\|I_1\|^2 - b_1\langle I_2, I_1\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2}$$

$$x_2 = a_{12}\frac{b_1\|I_2\|^2 - b_2\langle I_1, I_2\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2} + a_{22}\frac{b_2\|I_1\|^2 - b_1\langle I_2, I_1\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2}$$

$$\vdots$$

$$x_n = a_{1n}\frac{b_1\|I_2\|^2 - b_2\langle I_1, I_2\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2} + a_{2n}\frac{b_2\|I_1\|^2 - b_1\langle I_2, I_1\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2}$$

Example 6.4.3. Find a solution of the following system

$$\begin{aligned} x_1 + x_2 \qquad &= 1 \\ -x_1 + x_2 + x_3 &= -1 \end{aligned}$$

Solution: With the above notations

$$I_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad I_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Then, $\|I_1\|^2 = 1 + 1 = 2\,,\quad \|I_2\|^2 = 1 + 1 + 1 = 3,\quad |\langle I_1, I_2\rangle| = -1 + 1 + 0 = 0$

$$det(A\,A^T) = \|I_1\|^2\,\|I_2\|^2 - |\langle I_1, I_2\rangle|^2 = 2 \times 3 - 0 = 6$$

$$x_1 = a_{11}\frac{b_1\|I_2\|^2 - b_2\langle I_1, I_2\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2} + a_{21}\frac{b_2\|I_1\|^2 - b_1\langle I_2, I_1\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2}$$

$$= \frac{1(1\times 3 - -1\times 0)}{6} + \frac{(-1)(-1\times 2 - 1\times 0)}{6} = \frac{3}{6} + \frac{2}{6} = \frac{5}{6}$$

$$x_2 = a_{12}\frac{b_1\|I_2\|^2 - b_2\langle I_1, I_2\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2} + a_{22}\frac{b_2\|I_1\|^2 - b_1\langle I_2, I_1\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2}$$

$$= \frac{1(3-0)}{6} + \frac{(1)(-2-0)}{6} = \frac{3}{6} + \frac{-2}{6} = \frac{1}{6}$$

$$x_3 = a_{13}\frac{b_1\|I_2\|^2 - b_2\langle I_1, I_2\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2} + a_{23}\frac{b_2\|I_1\|^2 - b_1\langle I_2, I_1\rangle}{\|I_1\|^2\|I_2\|^2 - |\langle I_1, I_2\rangle|^2}$$

$$= \frac{(0)(3)}{6} + \frac{(1)(-2)}{6} = \frac{-2}{6}$$

We note that if $\|I_1\|^2\,\|I_2\|^2 = |\langle I_1, I_2\rangle|^2$ this means that the angle between $I_1,\ I_2$ the equals zero and this indicates that the system of equations are identical, meaning that there is an infinite number of solutions and in this case it is the same as in example 6.4.1.

But if the angle between $I_1,\ I_2$ is not equal to zero, then this means that the solution of the system has an infinite number of solutions.

We note that the general solution for this example given by Theorem 6.1.6 is

$$x = A^T(A\,A^T)^{-1}b + (I - A^T(A\,A^T)^{-1}A)y = A_R^{-1}b + (I - A_R^{-1}A)y$$

$$x = \frac{1}{6}\begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix} + \frac{1}{6}t\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

Choose $t = 0,\quad x = \left({}^{5}/_{6}, {}^{1}/_{6}, {}^{-2}/_{6}\right)$

Example 6.4.4 [10]. Consider the following general case of system $\boldsymbol{Ax} = \boldsymbol{b}$

$$\begin{array}{ccccccccc} a_{11}x_1 + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_n \end{array} \qquad (3)$$

Then, if $\{I_1, I_2, \dots, I_m\}$ is an orthogonal set in R^n, (see definition 1.9.4) we obtain

$$A\,A^T = \begin{bmatrix} \|I_1\|^2 & 0 & 0 & \cdots & 0 \\ 0 & \|I_2\|^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \|I_n\|^2 \end{bmatrix}$$

The solution of the system $\boldsymbol{A\,x} = \boldsymbol{b}$ is simple and is given by:

$$x_j = \sum_{i=1}^{m} a_{ij}\ b_{ji}\ \|I_i\|^{-2}\,, \quad j = 1,2,\dots,n$$

Example 6.4.5. Find the solution of the following system

$$\begin{array}{ccccccccc} -x_1 & - & x_2 & + & x_3 & + & x_4 & = & 1 \\ -x_1 & + & x_2 & - & x_3 & + & x_4 & = & 1 \\ x_1 & - & x_2 & - & x_3 & + & x_4 & = & 1 \end{array}$$

Solution:

$$I_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \quad I_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Then, $\|I_1\|^2 = 4\,,\quad \|I_2\|^2 = 4,\quad \|I_3\|^2 = 4$

$$det(A\,A^T) = \|I_1\|^2\ \|I_2\|^2\ \ \|I_3\|^2 = 4 \times 4 \times 4 = 64$$

$$x_1 = \frac{a_{11}b_1}{\|I_1\|^2} + \frac{a_{21}b_2}{\|I_2\|^2} + \frac{a_{31}b_3}{\|I_3\|^2} = \frac{(-1)(1)}{4} + \frac{(-1)(1)}{4} + \frac{(1)(1)}{4} = \frac{-1}{4}$$

$$x_2 = \frac{a_{12}b_1}{\|I_1\|^2} + \frac{a_{22}b_2}{\|I_2\|^2} + \frac{a_{32}b_3}{\|I_3\|^2} = \frac{(-1)(1)}{4} + \frac{(1)(1)}{4} + \frac{(-1)(1)}{4} = \frac{-1}{4}$$

$$x_3 = \frac{a_{13}b_1}{\|I_1\|^2} + \frac{a_{23}b_2}{\|I_2\|^2} + \frac{a_{33}b_3}{\|I_3\|^2} = \frac{(1)(1)}{4} + \frac{(-1)(1)}{4} + \frac{(-1)(1)}{4} = \frac{-1}{4}$$

$$x_4 = \frac{a_{14}b_1}{\|I_1\|^2} + \frac{a_{24}b_2}{\|I_2\|^2} + \frac{a_{34}b_3}{\|I_3\|^2} = \frac{(1)(1)}{4} + \frac{(1)(1)}{4} + \frac{(1)(1)}{4} = \frac{3}{4}$$

We note that the general solution for this example given by Theorem 6.1.6 is

$$x = A^T(A\,A^T)^{-1}b \;+\; (I - A^T(A\,A^T)^{-1}A)y = A_R^{-1}b + (I - A_R^{-1}A)y$$

$$x = \frac{1}{4}\begin{bmatrix} -1 \\ -1 \\ -1 \\ 3 \end{bmatrix} + \frac{1}{4}t\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Choose $t = 0$, $x = (\frac{-1}{4}, \frac{-1}{4}, \frac{-1}{4}, \frac{3}{4})$

Finally we note that if a system of linear equations has an infinite number of solutions, then we can use a suitable method discussed in this chapter to find a fixed solution and construct the general solution mentioned in Theorem 6.1.6. Certainly, it will be the same general solution given by Gauss elimination(see definition 1.7.2).

References: -

[1] Amiri, M., Fathy, M., Bayat, M., Generalization of some determinantal identities for non-square matrices based on Radics definition, TWMS J. Pure Appl. Math. 1, no 2(2010), 163-175.

[2] Anton, H., Ilementary Linear Algebra with application, Seventh edition.

[3] Arunkumar, M., Murthy, S., and Ganapath, G., Determinant for non-square matrices, IJMSEA, no 5 (2011), 389-401.

[4] Barata, J.C.A.,. Hussein, M.S; The Moore-Penrose Pseudoinverse. A Tutorial Review of the Theory. John Hopkins University Press,(2013).

[5] Jamshidvanda,S., Ghalandarzadeha,S., Amiraslanib,A., Oliaa,F., On the Maximal Solution of A Linear System over Tropical Semirings, STEM Department, The University of Hawaii-Maui College, Kahului, Hawaii, USA, (June 26, 2019).

[6] Jimmy, The Moore-Penrose Pseudoinverse, Math 33A: Laub, University of California - Los Angeles(2011).

[7] Joshi, V.N., A determinant for rectangular matrices, no. 21 (1980) , 137-146.

[8] Joyce,D., Determinants, part II , Math 130 Linear Algebra, Clark University(2015).

[9] Kolman, B., Introductory Linear Algebra with application, Six edition.

[10] Leiva, H., A Generalization of Cramer's Rule, Advances in linear Algebra and Matrix Theory, no 5 (2015), 156-166.

[11] Leon,S.,J., Linear Algebra with Applications, Fifth edition.

[12]Makarewicz,A., Pikuta, P., and Szalkowski, D., Properties of the determinant of a rectangular matrix, AUMC-S L-P, LXVIII, no 1(2014), 31-41.

[13] Radic, M.,A bout a determinant of rectangular $2 \times n$ matrix and its geometric interpretation, Beitrage Algebra Geom. 46,no.1(2005),321-349.

[14] Radić, M., A definition of determinant of rectangular matrix, Glas. Mat. Ser. III 1(21) (1966), 17–22.