THEORY OF MATRICES AND COMPLEX NUMBERS

by

LARS OLSEN

Course: MT1002 Mathematics

THEORY OF MATRICES

by

LARS OLSEN

1997

Course: MT1002 Mathematics

Table of Contents

Section 0 – Introduction

A linear equation in n unknows $x_1, \ldots, x_n$ is an equation of the form

$$a_1x_1 \quad + \quad a_2x_2 \quad + \quad \cdots \quad + \quad a_nx_n \quad = \quad b\,.$$

The word "linear" indicates that each unknown occurs just in its first power; there are no squares, cubes, square roots or fancies as $\sin x_i$.
Solving linear equations is not difficult when they come in ones and twos, but efficient methods become necessary when you have a collection of the things. A system of linear equations is a collections of linear equations:

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
a_{31}x_1 & + & a_{32}x_2 & + & \cdots & + & a_{3n}x_n & = & b_3 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m
\end{array} \tag{0.1}$$

All the a_{ij}'s and the b_i's are fixed numbers. So we have m equations in the unknowns $x_1, \ldots, x_n$. The object is to find all values for these unknows which satisfy all the equations simultaneously.
The standard way to tackle systems of simultaneous equations is to "eliminate unknowns". You use one of the equations to write one of the unknowns in terms of the others, and you then subsitute this expression into the remaining equations so as to get a smaller system in fewer unknowns. You then iterate this process by applying it to the newer system so as to get an even smaller one. And so on until you reach something you can solve for the surviving unknowns. Substituting this information back into the earlier computations will then lead to a full solution. Or at least that is the hope.
What we shall be doing with these systems of linear equations is just a streamlined version of this process, streamlined so as to cut down the amount of work and writing involved.
The central notion in achieving this streamlining is the idea of a matrix. Before you turn the page and read the definition of a matrix, take a good look at the system of linear equations in (0.1) – after having looked carefully at (0.1), the definition of a matrix should seem quite natural.

Section 1 – Definition of a Matrix

1.1. Definition of a Matrix.
Let n, m be positive integers. An $m \times n$ matrix A is a rectangular array of numbers with m rows and n columns. Thus, if A is an $m \times n$ matrix, then we denote it by

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

or

$$A = (a_{ij})_{1 \le i \le m, 1 \le j \le n},$$

or (simply)

$$A = (a_{ij}).$$

The numbers in a matrix are called the entries. Thus, a_{ij} is the entry in the ith row and jth column. The entry in the ith row and jth column of a matrix A is also commonly denoted by A_{ij}. We will denote the set of all $m \times n$ matrices with real valued entries by $M_{m,n}(\mathbb{R})$.

Examples.

$$\begin{pmatrix} 1 & 2 \\ 3 & 0 \\ -1 & 4 \end{pmatrix} \text{ is a } 3 \times 2 \text{ matrix with } a_{11} = 1,\ a_{12} = 2,\ a_{21} = 3,\ a_{22} = 0,\ a_{31} = -1,\ a_{32} = 4.$$

$$\begin{pmatrix} 2 & 1 & 0 & -3 \end{pmatrix} \text{ is a } 1 \times 4 \text{ matrix with } a_{11} = 2,\ a_{12} = 1,\ a_{13} = 0,\ a_{14} = -3.$$

$$\begin{pmatrix} -\sqrt{2} & \pi & e \\ 3 & \frac{1}{2} & 0 \\ 0 & -7 & 0 \end{pmatrix} \text{ is a } 3 \times 3 \text{ matrix with e.g. } a_{13} = e.$$

$$\begin{pmatrix} 1 \\ 3 \end{pmatrix} \text{ is a } 2 \times 1 \text{ matrix.}$$

$$\begin{pmatrix} 4 \end{pmatrix} \text{ is a } 1 \times 1 \text{ matrix.}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 0 \\ 0 & -5 \end{pmatrix} \in M_{3,2}(\mathbb{R}).$$

Section 2 – Special Types of Matrices

2.1. Column Matrix.
An $m \times 1$ matrix

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

is called a column matrix (or a columb vector).

2.2. Row Matrix.
An $1 \times n$ matrix

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{pmatrix}$$

is called a row matrix (or a row vector).

2.3. Square Matrix.
An $n \times n$ matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

is called a square matrix (of order n), and the entries $a_{11}, a_{22}, \dots, a_{nn}$ are called the diagonal entries of A.

2.4. Zero Matrix.
The $m \times n$ matrix

$$\begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix},$$

all of whose entries are zero, is called the $m \times n$ zero matrix, and is denoted by $O_{m,n}$ or O. If $n = m$ then we will write $O_n = O_{n,n}$.

2.5. Identity Matrix.
The $n \times n$ square matrix

$$\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix},$$

with 1's on the diagonal and 0's elsewhere, is called the identity matix of order n, and is denoted by I_n or I.

Examples.

$\begin{pmatrix} 3 & 1 & 0 \\ -1 & \pi & 3 \\ \sqrt{2} & 0 & 0 \end{pmatrix}$ is a square matrix with diagonal elements 3,π and 0.

$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is the zero matrix of order 3×3.

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is the identity matrix of order 3.

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is the identity matrix of order 2.

Section 3 – Operations on Matrices, Part I:

Addition of Matrices and Scalar Multiplication of Matrices

3.1. Addition of Matrices.
If A and B are matrices of the *same size*, then the sum $A+B$ is the matrix obtained by adding the entries of A to the corresponding entries of B. Thus, if

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}$$

are two matrices of the same size, then $A+B$ is defined by

$$A+B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix},$$

i.e.

$$(A+B)_{ij} = A_{ij} + B_{ij} \,.$$

Warning: Matrices of different size *cannot* be added.

Examples.

$$\begin{pmatrix} 2 & 1 & 0 & 3 \\ -1 & 0 & 2 & 4 \\ 4 & -2 & 7 & 0 \end{pmatrix} + \begin{pmatrix} -4 & 3 & 5 & 1 \\ 2 & 2 & 0 & -1 \\ 3 & 2 & -4 & 5 \end{pmatrix} = \begin{pmatrix} -2 & 4 & 5 & 4 \\ 1 & 2 & 2 & 3 \\ 7 & 0 & 3 & 5 \end{pmatrix}.$$

$$\begin{pmatrix} -2 & 4 & 5 & 4 \\ 1 & 2 & 2 & 3 \\ 7 & 0 & 3 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \quad \text{is not defined.}$$

3.2. Scalar Multiplication of Matrices.
For a $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

and a real number λ, we define the scalar product λA by

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{pmatrix}$$

i.e.

$$(\lambda A)_{ij} = \lambda A_{ij} \,.$$

In particular,

$$(-1)A = \begin{pmatrix} -a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & -a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{m1} & -a_{m2} & \dots & -a_{mn} \end{pmatrix}.$$

Example.

$$3 \begin{pmatrix} 2 & 1 & 0 & 3 \\ -1 & 0 & 2 & 4 \\ 4 & -2 & 7 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 3 & 0 & 9 \\ -3 & 0 & 6 & 12 \\ 12 & -6 & 21 & 0 \end{pmatrix}.$$

3.3. Theorem. Rules for scalar multiplication of matrices and matrix addition.
Let A, B and C by matrices of the same size, and let $\lambda, \mu \in \mathbb{R}$. Then the following statements hold.

(i) Matrix addition is commutative. i.e.

$$A + B = B + A\,.$$

(ii) Matrix addition is associative, i.e.

$$(A + B) + C = A + (B + C)\,.$$

(iii) Scalar multiplication and matrix addition are distributive, i.e.

$$\lambda(A + B) = \lambda A + \lambda B\,.$$

(iv) Scalar multiplication is associative, i.e.

$$\lambda(\mu A) = (\lambda\mu)A\,.$$

(v) $\lambda O = O$.

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

Proof

To prove these you just have to show that the matrix on the left has the same size as the one on the right, and that for all i, j, the (i, j)th entry of the one equals the (i, j)th entry of the other. We will now do this in case (ii). We have

$$\begin{aligned} ((A + B) + C)_{ij} &= (A + B)_{ij} + C_{ij} \\ &= A_{ij} + B_{ij} + C_{ij} \\ &= A_{ij} + (B + C)_{ij} \\ &= (A + (B + C))_{ij}\,. \quad \square \end{aligned}$$

3.4. Notation.
Let A, B and C by matrices of the same size, and let $\lambda, \mu \in \mathbb{R}$. Since (by Theorem 3.3) $A + (B + C) + (A + B) + C$, we will simply omit the brackets and write

$$A + B + C = A + (B + C) + (A + B) + C\,.$$

Similary, since $\lambda(\mu A) = (\lambda\mu)A$, we will omit the brackets and write

$$\lambda\mu A = \lambda(\mu A) = (\lambda\mu)A\,.$$

We will write

$$-A = (-1)A\,.$$

Finally, we shall write

$$A - B = A + (-B)\,.$$

Section 4 – Operations on Matrices, Part II:

Matrix Multiplication

4.0. Introduction.
It would be nice to define the product of two matrices on an entry by entry basis, much as we did for addition. It would be a natural way to proceed and would make for easy computation. Unfortunately, though it would make for an easy life, it would not be of much use. So instead we opt for a definition that works well with the applications we are hoping to get out of the theory of matrices.

We begin with a special case which will motivate the following definitions. Consider the system of linear equations with unknowns $x_1, \ldots, x_n$:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & \cdots & + & a_{3n}x_n & = & b_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

There is a natural $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

associated with the coefficiants in the equations. There are also natural column matrices associated with the x_i's and the b_i's, viz.

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

We are going to set up the definition so that we can write the system of equations as

$$AX = B\,.$$

Whether or not this is useful remains to be seen (but it would certainly save writing). Having decided that is what we want to do, the definition is forced.

4.1. Definition of matrix multiplication.
Let A and B be matrices such that

$$\text{“the number of coulmns in } A\text{”} \quad = \quad \text{“the number of rows in } B\text{”},$$

i.e. we let

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mp} \end{pmatrix}$$

be a $m \times p$ matrix, and we let

$$B = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \ldots & b_{pn} \end{pmatrix}$$

be a $p \times n$ matrix. The product AB of A and B is now defined as the following $m \times n$ matric

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1p} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & a_{i3} & \dots & a_{ip} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mp} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1n} \\ b_{21} & \dots & b_{2j} & \dots & b_{2n} \\ b_{31} & \dots & b_{3j} & \dots & b_{3n} \\ \vdots & & \vdots & & \vdots \\ b_{p1} & \dots & b_{pj} & \dots & b_{pn} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1j} & \dots & c_{1n} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \dots & c_{ij} & \dots & c_{in} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \dots & c_{mj} & \dots & c_{mn} \end{pmatrix},$$

where

$$\begin{aligned} (AB)_{ij} = c_{ij} &= a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \dots a_{ip}b_{pj} \\ &= \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{ip} \end{pmatrix} \bullet \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{pmatrix} \\ &= \text{the usual "dot product" of the } i\text{th row of } A \text{ and the } j\text{th column of } B. \end{aligned}$$

Warning: If A and B are matrices such that

$$\text{"the number of coulmns in } A\text{"} \quad \neq \quad \text{"the number of rows in } B\text{"},$$

then the product AB is *not* defined.

Examples.

$$\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 \cdot 3 + 2 \cdot 2 & 1 \cdot 1 + 2 \cdot 0 \\ -1 \cdot 3 + 1 \cdot 2 & -1 \cdot 1 + 1 \cdot 0 \end{pmatrix} = \begin{pmatrix} 7 & 1 \\ -1 & -1 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{pmatrix} \begin{pmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{pmatrix} = \begin{pmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{pmatrix}.$$

$$\begin{pmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{pmatrix} \quad \text{is not defined.}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 11 \\ 29 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ -1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 11 \\ 9 & 29 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} -6 & 13 \\ -7 & 11 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} \quad \text{is not defined.}$$

$$\begin{pmatrix} 2 & 3 \\ -1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 14 & 19 & 24 \\ 3 & 3 & 3 \\ 9 & 12 & 15 \end{pmatrix}.$$

Example. System of linear equations.
Consider the system of linear equations with unknowns $x_1, \dots, x_n$:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & \cdots & + & a_{3n}x_n & = & b_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \tag{4.1}$$

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

and

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Then (4.1) can be rewritten as

$$AX = B\,.$$

4.2. Points to be careful about.
This is a list of things which you might expect to be true, but which are not.

(1) AB will usually not be equal to BA, even when both products are defined.
Example:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ 11 & -3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} -2 & -2 \\ 2 & 4 \end{pmatrix}.$$

(2) It can happen that $AB = O$, even though neither A nor B are O.
Example:

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

(3) You cannot always cancel in matrix equations. In other words, $AB = AC$ and $A \neq O$ do not imply that $B = C$.
Example:

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} 2 & -4 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

4.3. Theorem. Rules for matrix multiplication.
Let A, B and C by matrices, and let $\lambda \in \mathbb{R}$. Then the following statements hold.

(i) If A is $m \times n$, B is $n \times p$ and C is $p \times q$, then

$$A(BC) = (AB)C\,.$$

(ii) If A is $m \times n$ and B, C are $n \times p$, then

$$A(B + C) = AB + AC\,.$$

(iii) If A, B are $m \times n$ and C is $n \times p$, then.

$$(A+B)C = AC + BC\,.$$

(iv) If A is $m \times n$ and B is $n \times p$, then

$$(\lambda A)B = \lambda(AB) = A(\lambda B)\,.$$

(v) $OA = AO = O$.
(vi) If A is $n \times n$ and I is the $n \times n$ identity matrix, then

$$AI = IA = A\,.$$

Proof
To prove these you just have to show that the matrix on the left has the same size as the one on the right, and that for all i, j, the (i,j)th entry of the one equals the (i,j)th entry of the other. We will now do this in case (ii). We have

$$\begin{aligned}
(A(B+C))_{ij} &= \sum_k A_{ik}(B+C)_{kj} \\
&= \sum_k A_{ik}(B_{kj}+C_{kj}) \\
&= \sum_k (A_{ik}B_{kj}+A_{ik}C_{kj}) \\
&= \sum_k A_{ik}B_{kj} + \sum_k A_{ik}C_{kj} \\
&= (AB)_{ij} + (AC)_{ij} \\
&= (AB+AC)_{ij}\,. \quad \square
\end{aligned}$$

Section 5 – Inverse Matrices

It follows from 4.3.(vi) that the indentity matrix works with respect to matrix multiplication in the same way as 1 does with respect to ordinary multiplication. The natural question to ask once you have something that functions as a 1 is "Can we divide?" Given a matrix A, can we get an A^{-1}? The answer is "sort of and sometimes".

5.1. Definition of inverse.
Let A be an $n \times n$ square matrix. If there exists an $n \times n$ square matrix B such that

$$AB = I_n = BA\,,$$

then A is called invertible, and B is called *an* inverse of A.
Observe that if A is an invertible $n \times n$ matrix and B, C are inverses of A, then

$$B = C\,.$$

Indeed, since B and C are inverse of A we have

$$BA = I_n \tag{5.1}$$

and

$$AC = I_n\,, \tag{5.2}$$

whence

$$\begin{aligned}
B &= BI_n && \text{(by 4.3.(vi))}\\
&= B(AC) && \text{(by (5.2))}\\
&= (BA)C && \text{(by 4.3.(i))}\\
&= I_nC && \text{(by (5.1))}\\
&= C\,. && \text{(by 4.3.(vi))}
\end{aligned}$$

Hence, if A is invertible , then A has exactly one inverse, and so there is no ambiguity when we speak of *the* inverse of A. The inverse of A is denoted by

$$A^{-1}\,.$$

5.2. Notes.

(1) For A^{-1} to exist, A must be square.
(2) If A is square, A^{-1} sometimes exists, but not always.

5.3. Potential use for inverses.
Consider the system of n linear equations with n unknowns $x_1, \dots, x_n$:

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1\\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2\\
a_{31}x_1 & + & a_{32}x_2 & + & \cdots & + & a_{3n}x_n & = & b_3\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & b_n
\end{array} \tag{5.3}$$

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

and

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Then (5.3) can be rewritten as

$$AX = B. \tag{5.4}$$

Suppose further that we know that A is invertible, and that we know the inverse A^{-1}. Then to solve (5.3) (i.e. to find X) we would simply multiply both sides of the matrix equation (5.4) by A^{-1},

$$\begin{aligned} & AX = B \\ \Downarrow & \\ & A^{-1}(AX) = A^{-1}B \\ \Downarrow & \\ & (A^{-1}A)X = A^{-1}B \\ \Downarrow & \\ & I_n X = A^{-1}B \\ \Downarrow & \\ & X = A^{-1}B. \end{aligned}$$

Example.
The inverse of $A = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$ is $A^{-1} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$ since

$$\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We will see how to find inverses in Section 6.

Example.
Let a, b, c and d be real numbers and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then

$$A^{-1} \text{ exists} \quad \Leftrightarrow \quad ad - bc \neq 0,$$

and $A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. This follows from the fact that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We will see a natural generalization of this result in Section 8.

Example.
The matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \end{pmatrix}$ does not have an inverse. To see why, let $B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$ be any 3×3 matrix. Then

$$AB = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & 0 \end{pmatrix}$$

where the $*$'s represent some real numbers. Hence, for any 3×3 matrix B, the $(3,3)$ entry of AB is 0, whence $AB \neq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (because the $(3,3)$ entry of $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is 1). This shows that A is not invertible.

5.4. Theorem. Rules for inverse matrices.
Let A and B be $n \times n$ matrices. Then the following statements hold.

(i) If A and B are invertible, then AB is invertible and

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

(ii) If A is invertible, then A^{-1} is invertible and

$$(A^{-1})^{-1} = A\,.$$

Proof
It is not difficult to verify these statements, and all students are encouraged to attempt to prove this theorem. $\square$

Section 6 – Finding Inverse Matrices by Row and Column Operations

There are several methods for calculating the inverse A^{-1} of a given matrix A (provided the inverse exists). We shall concentrate on the one that you will always reach for when you are dealing with a matrix with numerical entries.

6.1. Definition of row operations.
Consider a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

The three elementary row operations are:

(1) Multiplying a row by a non-zero number λ:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \xrightarrow[R_i \to \lambda R_i]{} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{i1} & \lambda a_{i2} & \dots & \lambda a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

(2) Add a multiple of one row to another row:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \xrightarrow[R_j \to R_j + \lambda R_i]{} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} + \lambda a_{i1} & a_{j2} \lambda a_{i2} & \dots & a_{jn} \lambda a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

(3) Interchanging two rows:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \xrightarrow[R_i \leftrightarrow R_j]{} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

6.2. Definition of column operations.
Consider a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

The three elementary column operations are:

(1) Multiplying a column by a non-zero number λ:

$$\begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & \dots & a_{2i} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{m2} & \dots & a_{mn} \end{pmatrix} \underset{C_i \to \lambda C_i}{\longrightarrow} \begin{pmatrix} a_{11} & \dots & \lambda a_{1i} & \dots & a_{1n} \\ a_{21} & \dots & \lambda a_{2i} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & \lambda a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

(2) Add a multiple of one column to another column:

$$\begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & \dots & a_{2i} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix} \underset{C_j \to C_j + \lambda C_i}{\longrightarrow} \begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} + \lambda a_{1i} & \dots & a_{1n} \\ a_{21} & \dots & a_{2i} & \dots & a_{2j} + \lambda a_{2i} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mj} + \lambda a_{mi} & \dots & a_{mn} \end{pmatrix}.$$

(3) Interchanging two columns:

$$\begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & \dots & a_{2i} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix} \underset{C_i \leftrightarrow C_j}{\longrightarrow} \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & \dots & a_{2j} & \dots & a_{2i} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mi} & \dots & a_{mn} \end{pmatrix}.$$

6.3. Theorem. Method for finding the inverse of a matrix by row operations.
Consider the following $n \times n$ square matrix,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

(1) Construct the augmented matrix

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right),$$

i.e. the $n \times 2n$ matrix formed by joining A and the $n \times n$ identity matrix.

(2) Perform row operations on the left hand side of the augmented matrix to reduce it to the $n \times n$ identity matrix I_n while performing the *same* row operations on the right hand side of the augmented matrix. This process will give,

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & 1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{array}\right).$$

The inverse of A is now given by,

$$A^{-1} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}.$$

6.4. Theorem. Method for finding the inverse of a matrix by column operations.
Consider the following $n \times n$ square matrix,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

(1) Construct the augmented matrix

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right),$$

i.e. the $n \times 2n$ matrix formed by joining A and the $n \times n$ identity matrix.

(2) Perform column operations on the left hand side of the augmented matrix to reduce it to the $n \times n$ identity matrix I_n while performing the *same* column operations on the right hand side of the augmented matrix. This process will give,

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & 1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{array}\right).$$

The inverse of A is now given by,

$$A^{-1} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}.$$

6.5. Note.
Observe that when you compute inverse matrices using 6.3 or 6.4, then you must *either* use row operations all the time, *or* you must use column operations all the time. You *cannot* mix row and column operations.

Example.
Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix}$. We will now compute A^{-1} using row operations. Forming the augmented matrix and performing row operations yields.

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right) \underset{\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - R_1}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right)$$

$$\underset{R_3 \to R_3 + 2R_2}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & 5 & 2 & 1 \end{array}\right)$$

$$\underset{R_3 \to -R_3}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & -5 & -2 & -1 \end{array}\right)$$

$$\underset{\substack{R_1 \to R_1 - 3R_3 \\ R_2 \to R_2 + 3R_3}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right)$$

$$\underset{R_1 \to R_1 - 2R_2}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right).$$

Hence

$$A^{-1} = \begin{pmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix}.$$

Example.

Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. We will now compute A^{-1} using row operations. Forming the augmented matrix and performing row operations yields.

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \underset{\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - R_1}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 0 & -2 & -2 & -1 & 0 & 1 \end{array}\right)$$

$$\underset{\substack{R_2 \to -\frac{1}{3}R_2 \\ R_3 \to -\frac{1}{2}R_3}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 1 & 1 & \frac{1}{2} & 0 & -\frac{1}{2} \end{array}\right)$$

$$\underset{R_3 \to R_3 - R_2}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & -1 & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{2} \end{array}\right)$$

$$\underset{R_3 \to -R_3}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & \frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \end{array}\right)$$

$$\underset{R_2 \to R_2 - 2R_3}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & -1 \\ 0 & 0 & 1 & \frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \end{array}\right)$$

$$\underset{R_1 \to R_1 - 2R_2 - 3R_3}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & -1 \\ 0 & 0 & 1 & \frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \end{array}\right).$$

Hence

$$A^{-1} = \begin{pmatrix} -\frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & -1 \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \end{pmatrix} = \frac{1}{6} \begin{pmatrix} -1 & 2 & 3 \\ 2 & 2 & -6 \\ 1 & -2 & 3 \end{pmatrix}.$$

6.6. Remark.

Question.
Why does the row (column) operation method for computing A^{-1} work?

Answer.
We will only consider the 3×3 case. Each of the elementary row (column) operations corresponds to multiplying A with a matrix from the left (right). For example:

(1) Multiplying $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ with $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}$ from the left, multiplies 2nd row in A by λ,
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} .$$
Hence, multiplying A with $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}$ from the left corresponds to the row operation: $R_2 \to \lambda R_2$.

(2) Multiplying $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ with $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix}$ from the left, multiplies the 3rd row in A by λ and adds it to the 2nd row,
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} + \lambda a_{31} & a_{22} + \lambda a_{32} & a_{23} + \lambda a_{33} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} .$$
Hence, multiplying A with $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix}$ from the left corresponds to the row operation: $R_2 \to R_2 + \lambda R_3$.

(3) Multiplying $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ with $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ from the left, interchanges 1st and 2nd row in A,
$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} .$$
Hence, multiplying A with $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ from the left corresponds to the row operation: $R_2 \leftrightarrow R_3$.

The row operation method therefore consists in multiplying matrices $E_1, E_2, \ldots, E_k$ to A from the left until we get the identity I:
$$E_k E_{k-1} \cdots E_3 E_2 E_1 A = I .$$
Hence
$$\begin{aligned} A^{-1} &= E_k E_{k-1} \cdots E_3 E_2 E_1 \\ &= E_k E_{k-1} \cdots E_3 E_2 E_1 I . \end{aligned} \tag{6.1}$$
Equation (6.1) tells us that A^{-1} can be found by applying the *same* row operations to I, but this was exactly what we did on the right hand side of the augmented matrix.
Similarly, the column operation method consists in multiplying matrices $F_1, F_2, \ldots, F_k$ to A from the right until we get the identity I:
$$A F_1 F_2 F_3 \cdots F_{k-1} F_k = I .$$

Hence

$$\begin{aligned} A^{-1} &= F_1F_2F_3\cdots F_{k-1}F_k \\ &= IF_1F_2F_3\cdots F_{k-1}F_k\,. \end{aligned} \tag{6.2}$$

Equation (6.2) tells us that A^{-1} can be found by applying the *same* column operations to I, but this was exactly what we did on the right hand side of the augmented matrix.

Section 7 – Systems of Linear Equations

Consider the system of n linear equations with n unknowns $x_1, \dots, x_n$:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & \cdots & + & a_{3n}x_n & = & b_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & b_n \end{array} \tag{7.1}$$

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

and

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} .$$

Then (7.1) can be rewritten as

$$AX = B \, . \tag{7.2}$$

We will discuss two methods of solving (7.1) (or equivalently: solving (7.2)).

7.1. Solving systems of linear equations by inverse matrices.
Suppose that we know that A is invertible with inverse A^{-1}. Then to solve (7.1) simply multiply both sides of the matrix equation (7.2) by A^{-1},

$$\begin{array}{l} AX = B \\ \Downarrow \\ A^{-1}(AX) = A^{-1}B \\ \Downarrow \\ (A^{-1}A)X = A^{-1}B \\ \Downarrow \\ I_nX = A^{-1}B \\ \Downarrow \\ X = A^{-1}B \, . \end{array}$$

Example in solving systems of linear equations by inverse matrices.
Solve the system of linear equations,

$$\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 5 \\ 2x & + & 5y & + & 3z & = & -2 \\ x & & & + & 8z & = & 1 \end{array} \tag{7.3}$$

Letting

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix} ,$$

then (7.3) can be rewritten as

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix}.$$

We know from an example in Section 6 that A is invertible with $A^{-1} = \begin{pmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix}$. The solution to (7.3) is therefore given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1}B = \begin{pmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix} \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} -200-32+9 \\ 65+10-3 \\ 25+4-1 \end{pmatrix} = \begin{pmatrix} -233 \\ 72 \\ 28 \end{pmatrix}.$$

7.2. Solving systems of linear equations by Gaussian elimination.

(1) Construct the augmented matrix associated with the system (7.1),

$$\left(\begin{array}{ccccccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1,n-2} & a_{1,n-1} & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2,n-2} & a_{2,n-1} & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \dots & a_{3,n-2} & a_{3,n-1} & a_{3n} & b_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ a_{n-2,1} & a_{n-2,2} & a_{n-2,3} & \dots & a_{n-2,n-2} & a_{n-2,n-1} & a_{n-2,n} & b_{n-2} \\ a_{n-1,1} & a_{n-1,2} & a_{n-1,3} & \dots & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} & b_{n-1} \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{n,n-2} & a_{n,n-1} & a_{nn} & b_n \end{array}\right),$$

i.e. the $n \times (n+1)$ matrix formed by joining A and B.

(2) Perform row operations on the augmented matrix until the $n \times n$ left hand side is in upper triangular form,

$$\left(\begin{array}{ccccccc|c} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1,n-2} & \alpha_{1,n-1} & \alpha_{1n} & \beta_1 \\ 0 & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2,n-2} & \alpha_{2,n-1} & \alpha_{2n} & \beta_2 \\ 0 & 0 & \alpha_{33} & \dots & \alpha_{3,n-2} & \alpha_{3,n-1} & \alpha_{3n} & \beta_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{n-2,n-2} & \alpha_{n-2,n-1} & \alpha_{n-2,n} & \beta_{n-2} \\ 0 & 0 & 0 & \dots & 0 & \alpha_{n-1,n-1} & \alpha_{n-1,n} & \beta_{n-1} \\ 0 & 0 & 0 & \dots & 0 & 0 & \alpha_{n,n} & \beta_n \end{array}\right)$$

(i.e. all entries below the main diagonal are 0). If possible try to get 1's in the main diagonal (warning: this is not always possible),

$$\left(\begin{array}{ccccccc|c} 1 & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1,n-2} & \alpha_{1,n-1} & \alpha_{1n} & \beta_1 \\ 0 & 1 & \alpha_{23} & \dots & \alpha_{2,n-2} & \alpha_{2,n-1} & \alpha_{2n} & \beta_2 \\ 0 & 0 & 1 & \dots & \alpha_{3,n-2} & \alpha_{3,n-1} & \alpha_{3n} & \beta_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & \alpha_{n-2,n-1} & \alpha_{n-2,n} & \beta_{n-2} \\ 0 & 0 & 0 & \dots & 0 & 1 & \alpha_{n-1,n} & \beta_{n-1} \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 & \beta_n \end{array}\right).$$

(3) Once in upper triangular form, translate the new augmented matrix back into the equations it represents and solve, working from the bottom up,

$$\begin{array}{ccccccccccccc} \alpha_{11}x_1 & + & \alpha_{12}x_2 & + & \alpha_{13}x_3 & + & \cdots & + & \alpha_{1,n-2}x_{n-2} & + & \alpha_{1,n-1}x_{n-1} & + & \alpha_{1n}x_n & = & \beta_1 \\ & & \alpha_{22}x_2 & + & \alpha_{23}x_3 & + & \cdots & + & \alpha_{2,n-2}x_{n-2} & + & \alpha_{2,n-1}x_{n-1} & + & \alpha_{2n}x_n & = & \beta_2 \\ & & & & \alpha_{33}x_3 & + & \cdots & + & \alpha_{3,n-2}x_{n-2} & + & \alpha_{3,n-1}x_{n-1} & + & \alpha_{3n}x_n & = & \beta_3 \\ & & & & & & \vdots & & & & & & & & \\ & & & & & & & & \alpha_{n-2,n-2}x_{n-2} & + & \alpha_{n-2,n-1}x_{n-1} & + & \alpha_{n-2,n}x_n & = & \beta_{n-2} \\ & & & & & & & & & & \alpha_{n-1,n-1}x_{n-1} & + & \alpha_{n-1,n}x_n & = & \beta_{n-1} \\ & & & & & & & & & & & & \alpha_{n,n}x_n & = & \beta_n \end{array}.$$

Example.
We solve the following system of linear equations using Gaussian elimination,

$$\begin{array}{ccccccc} x & + & 2y & - & z & = & 2 \\ x & + & 3y & + & z & = & 10 \\ x & - & y & + & 2z & = & 5 \end{array}.$$

Forming the augmented matrix and performing row operations yields.

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 1 & 3 & 1 & 10 \\ 1 & -1 & 2 & 5 \end{array}\right) \underset{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & 1 & 2 & 8 \\ 0 & -3 & 3 & 3 \end{array}\right)$$

$$\underset{R_3 \to R_3 + 3R_2}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 9 & 27 \end{array}\right)$$

$$\underset{R_3 \to \frac{1}{9}R_3}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 3 \end{array}\right).$$

This translates back as equations to

$$x+2y- \ z = 2 \tag{1}$$
$$y+2z = 8 \tag{2}$$
$$z = 3 \tag{3}$$

Equation (3) gives: $z = 3$.

Equation (2) gives: $y + 2z = 8 \quad \Rightarrow \quad y = 8 - 2z = 8 - 2(3) = 2.$

Equation (1) gives: $x + 2y - z = 2 \quad \Rightarrow \quad x = 2 - 2y + z = 2 - 2(2) + 3 = 1.$

Hence the solution is:

$$(x, y, z) = (1, 2, 3).$$

Example.
We solve the following system of linear equations using Gaussian elimination,

$$\begin{array}{rcrcrcr} x & + & y & + & 2z & = & 9 \\ 2x & + & 4y & - & 3z & = & 1 \\ 3x & + & 6y & - & 5z & = & 0 \end{array}.$$

Forming the augmented matrix and performing row operations yields.

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right) \underset{\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 3R_1}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 3 & -11 & -27 \end{array}\right)$$

$$\underset{\substack{R_2 \to \frac{1}{2}R_2 \\ R_3 \to \frac{1}{3}R_3}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 1 & -\frac{11}{3} & -9 \end{array}\right)$$

$$\underset{R_3 \to R_3 - R_2}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & -\frac{1}{6} & -\frac{1}{2} \end{array}\right)$$

$$\underset{R_3 \to -6R_3}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{array}\right).$$

This translates back as equations to

$$x+y+2z = 9 \tag{1}$$

$$y-\frac{7}{2}z = -\frac{17}{2} \tag{2}$$

$$z = 3 \tag{3}$$

Equation (3) gives: $z = 3$.

Equation (2) gives: $y - \frac{7}{2}z = -\frac{17}{2} \quad \Rightarrow \quad y = -\frac{17}{2} + \frac{7}{2}z = -\frac{17}{2} + \frac{7}{2}3 = \frac{21-17}{2} = 2.$

Equation (1) gives: $x + y + 2z = 9 \quad \Rightarrow \quad x = 9 - y - 2z = 9 - 2 - 2(3) = 1.$

Hence the solution is:

$$(x, y, z) = (1, 2, 3).$$

Example.
Let a be a real number. We solve the following system of linear equations using Gaussian elimination,

$$\begin{array}{rcrcrcl} x & - & y & + & z & = & 1 \\ x & & & + & az & = & 3\,. \\ -x & + & ay & & & = & 1 \end{array}$$

Forming the augmented matrix and performing row operations yields.

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 1 & 0 & a & 3 \\ -1 & a & 0 & 1 \end{array}\right) \underset{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 + R_1}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 1 & a-1 & 2 \\ 0 & a-1 & 1 & 2 \end{array}\right)$$

$$\underset{R_3 \to R_3 - (a-1)R_2}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 1 & a-1 & 2 \\ 0 & 0 & 1-(a-1)^2 & 2-2(a-1) \end{array}\right)$$

$$= \left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 1 & a-1 & 2 \\ 0 & 0 & a(2-a) & 2(2-a) \end{array}\right).$$

This translates back as equations to

$$x - y + \quad z = 1 \tag{1}$$
$$y + \ (a-1)z = 2 \tag{2}$$
$$a(2-a)z = 2(2-a) \tag{3}$$

The critical equation is the last one.

If $a(2-a) \neq 0$, we can divide by it to get z. Back substitution will then deliver values for x and y. We therefore have to divide our analysis into the following cases.

Case 1: $a(2-a) \neq 0$**, i.e.** $a \neq 0, 2$**.**

Since $a(2-a) \neq 0$, we can divide by it in equation (3). Equation (3) therefore gives: $\quad z = \frac{2(2-a)}{a(2-a)} = \frac{2}{a}$.

Equation (2) gives: $\quad y + (a-1)z = 2 \quad \Rightarrow \quad y = 2 - (a-1)z = 2 - (a-1)\frac{2}{a} = \frac{2}{a}$.

Equation (1) gives: $\quad x - y + z = 1 \quad \Rightarrow \quad z = 1 + y - z = 1 + \frac{2}{a} - \frac{2}{a} = 1$.

Hence in Case 1 the solution is:

$$(x, y, z) = (1, \frac{2}{a}, \frac{2}{a})\,.$$

Case 2: $a = 0$**.**

In this case equation (3) becomes $0 = 4$, which is impossible. So there is no solution in this case.

Case 3: $a = 2$**.**

In this case equation (3) becomes $0 = 0$, which is true for all valus of z. Hence z can be any real number t,

$$z = t\,.$$

Equation (2) gives (remember that $a = 2$): $\quad y+(a-1)z = 2 \quad \Rightarrow \quad y = 2-(a-1)z = 2-(2-1)t = 2-t$.

Equation (1) gives (remember that $a = 2$): $\quad x-y+z = 1 \quad \Rightarrow \quad z = 1+y-z = 1+2-t-t = 3-2t$.

Hence in Case 3 the solution is

$$(x, y, z) = (3-2t, 2-t, t) \quad \text{for any } t \in \mathbb{R}.$$

Since there are an infinite numer of possibilities for t, this means that there are an infinite number of solutions in this case.

Section 8 – Determinants

Deternimants are numbers got by doing certain calculations with the entries of a matrix. They have nice properties and are sometimes helpful when dealing with inverses of a matrix. As with inverses, they are only defined for square matrices.

8.1. Definition of determinants.
Consider the following $n \times n$ square matrix,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} .$$

The determinant of A, denoted by

$$\det A \,,$$

or

$$|A| \,,$$

or

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

(so curved brackets for a matrix and straight lines for a determinant) is defined inductively as follows:

For $n = 1$,

$$|a_{11}| = a_{11} \,.$$

For $n = 2$,

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

For $n = 3$,

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

In general, if we let

$$\begin{aligned} A_{ij} = {} & \text{the } (n-1) \times (n-1) \text{ matrix obtainen from } A \text{ by} \\ & \text{deleting the } i\text{th row and the } j\text{th column} \\ = {} & \begin{pmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{pmatrix} , \end{aligned}$$

then

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \begin{vmatrix} a_{1,2} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i-1,2} & \dots & a_{i-1,n} \\ \\ a_{i+1,2} & \dots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n,2} & \dots & a_{nn} \end{vmatrix}$$

$$= \sum_{i=1}^{n} (-1)^{i+1} a_{i1} |A_{i1}| \,. \tag{8.1}$$

With this definition we can evaluate $n \times n$ determinats provided we can evaluate $(n-1) \times (n-1)$ determinants. We know that we can deal with low values of n, and so we have an inductive definition. The definition in (8.1) is for obvious reasons called "expansion of $|A|$ along the 1st column". We can also expand along the ith row or the jth column – this is explained in the next theorem.

8.2. Theorem. Cramer's rule.
Consider the following $n \times n$ square matrix,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} .$$

(1) The determinant of A can be computed by expansion along the jth column,

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}$$

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

(2) The determinant of A can be computed by expansion along the ith row,

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}$$

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

Example.
Let $D = \begin{vmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{vmatrix}$. We first compute D by expanding along the 1st column,

$$\begin{aligned} \begin{vmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{vmatrix} &= (-1)^{1+1}3 \begin{vmatrix} -4 & 3 \\ 4 & -2 \end{vmatrix} + (-1)^{2+1}(-2) \begin{vmatrix} 1 & 0 \\ 4 & -2 \end{vmatrix} + (-1)^{3+1}5 \begin{vmatrix} 1 & 0 \\ -4 & 3 \end{vmatrix} \\ &= 3(8-12) + 2(-2-0) + 5(3-0) \\ &= -1\,. \end{aligned}$$

Next we compute D by expanding along the 1st row,

$$\begin{aligned} \begin{vmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{vmatrix} &= (-1)^{1+1}3 \begin{vmatrix} -4 & 3 \\ 4 & -2 \end{vmatrix} + (-1)^{1+2}1 \begin{vmatrix} -2 & 3 \\ 5 & -2 \end{vmatrix} + (-1)^{1+3}0 \begin{vmatrix} -2 & -4 \\ 5 & 4 \end{vmatrix} \\ &= 3(-8-12) - (4-15) \\ &= -1\,. \end{aligned}$$

Finally, we compute D by expanding along the 2nd column,

$$\begin{aligned} \begin{vmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{vmatrix} &= (-1)^{1+2}1 \begin{vmatrix} -2 & 3 \\ 5 & -2 \end{vmatrix} + (-1)^{2+2}(-4) \begin{vmatrix} 3 & 0 \\ 5 & -2 \end{vmatrix} + (-1)^{3+2}4 \begin{vmatrix} 3 & 0 \\ -2 & 3 \end{vmatrix} \\ &= -(4-15) - 4(-6) - 4(9) \\ &= -1\,. \end{aligned}$$

Example.
We compute $D = \begin{vmatrix} 7 & -3 & 0 \\ 4 & 5 & 9 \\ -1 & 2 & 0 \end{vmatrix}$. Since the 3rd column contains a large number of 0's we decide to compute D by expanding along the 3rd column,

$$\begin{aligned} \begin{vmatrix} 7 & -3 & 0 \\ 4 & 5 & 9 \\ -1 & 2 & 0 \end{vmatrix} &= (-1)^{1+3}0 \begin{vmatrix} 4 & 5 \\ -1 & 2 \end{vmatrix} + (-1)^{2+3}9 \begin{vmatrix} 7 & -3 \\ -1 & 2 \end{vmatrix} + (-1)^{3+3}0 \begin{vmatrix} 7 & -3 \\ 4 & 5 \end{vmatrix} \\ &= -9(14-3) \\ &= -99\,. \end{aligned}$$

Example.
The determinant of a 2×2 matrix is given by

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc\,.$$

It follows from the example just before 5.4 that:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ is invertible } \left(\text{with } \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \right) \quad \Leftrightarrow \quad ad - bc \neq 0\,.$$

Hence:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ is invertible } \quad \Leftrightarrow \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0\,.$$

This result is not a coincidence, as the next theorem shows.

8.3. Theorem.
Let A by a square matrix. Then:

$$A \text{ is invertible} \quad \Leftrightarrow \quad \det A \neq 0\,.$$

Proof
This is a very deep result and the proof is far beyond the scope of this course. For the case where A is a 2×2 matrix, the result is proved in the previous Example. $\square$

Example.
Find all λ for which the following matrix

$$A = \begin{pmatrix} \lambda & 0 & 1 \\ 1 & 3 & 0 \\ 2 & 5 & 2 \end{pmatrix}$$

is invertible. Since

$$\begin{aligned} \det A &= \begin{vmatrix} \lambda & 0 & 1 \\ 1 & 3 & 0 \\ 2 & 5 & 2 \end{vmatrix} \\ &= (-1)^{1+3} 1 \begin{vmatrix} 1 & 3 \\ 2 & 5 \end{vmatrix} + (-1)^{2+3} 0 \begin{vmatrix} \lambda & 0 \\ 2 & 5 \end{vmatrix} + (-1)^{3+3} 2 \begin{vmatrix} \lambda & 0 \\ 1 & 3 \end{vmatrix} \\ &= (5-6) + 2(3\lambda) \\ &= 6\lambda - 1\,, \end{aligned}$$

we obtain

$$A \text{ is invertible} \quad \Leftrightarrow \quad \det A \neq 0 \quad \Leftrightarrow \quad 6\lambda - 1 \neq 0 \quad \Leftrightarrow \quad \lambda \neq \frac{1}{6}\,.$$

8.4. Theorem. Properties of determinants.
Let A and B be $n \times n$ square matrices.

(1) $\det I_n = 1$.
(2) $\det(AB) = \det(A)\det(B)$.
(3) If A is invertible, then (by (1) and (2))

$$\det(A^{-1})\det(A) = \det(A^{-1}A) = \det I_n = 1\,,$$

and so

$$\det(A^{-1}) = \frac{1}{\det(A)}\,.$$

Proof
Statement (1) is not too difficult to verify (Try it!). Statement (2), on the other hand, is a very deep result and the proof is far beyond the scope of this course. (Much to the regret of the author of these notes, the proof is not presented in any undergraduate course at this University.) However, it is easy to verify (2) by a direct calculation for 2×2 matrices (Do it!). $\square$

Section 9 – Computation of Determinants by Row and Column Operations

The results in 9.1 and 9.2 are often useful in computing determinants.

9.1. The row and column operation theorem.

Let $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$ be a $n \times n$ square matrix.

(1) If B is the matrix that results when a single row (column) of A is multiplied by a scaler λ, then

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

i.e.

$$\begin{vmatrix} a_{11} & \dots & \lambda a_{1i} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & \lambda a_{ni} & \dots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nn} \end{vmatrix},$$

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{i1} & \dots & \lambda a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix},$$

for all i.

(2) If B is the matrix that results when two rows (columns) of A are interchanged, then

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

i.e.

$$\begin{vmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1i} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{ni} & \dots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix},$$

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix},$$

for all i, j with $i \neq j$.

(3) If B is the matrix that results when a multiple of one row of a is added to another row, then

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

i.e.

$$\begin{vmatrix} a_{11} & \dots & a_{1i}+\lambda a_{1j} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{ni}+\lambda a_{nj} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix},$$

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+\lambda a_{j1} & \dots & a_{in}+\lambda a_{jn} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix},$$

for all i, j.

Proof.
All these rules follow from the fact that the determinant of a product is the product of the determinant, cf. 8.4.(2). $\square$

9.2. Theorem. The determinant of a triangular matrix.
If A is an upper triangular square matrix, i.e. if

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1,n-2} & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2,n-2} & a_{2,n-1} & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3,n-2} & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_{n-2,n-2} & a_{n-2,n-1} & a_{n-2,n} \\ 0 & 0 & 0 & \dots & 0 & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & 0 & a_{nn} \end{pmatrix},$$

or if A is a lower triangular square matrix. i.e. if

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ a_{n-2,1} & a_{n-2,2} & a_{n-2,3} & \dots & a_{n-2,n-2} & 0 & 0 \\ a_{n-1,1} & a_{n-1,2} & a_{n-1,3} & \dots & a_{n-1,n-2} & a_{n-1,n-1} & 0 \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{n,n-2} & a_{n,n-1} & a_{nn} \end{pmatrix},$$

then

$$\det(A) = a_{11}a_{22}a_{33}\cdots a_{nn}\,.$$

Proof.

We will only consider the upper triangular case. This follows by n succesive expansions along the 1st column,

$$\begin{aligned}
\det A &= \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1,n-2} & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2,n-2} & a_{2,n-1} & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3,n-2} & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_{n-2,n-2} & a_{n-2,n-1} & a_{n-2,n} \\ 0 & 0 & 0 & \dots & 0 & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & 0 & a_{nn} \end{vmatrix} \\
&= a_{11} \begin{vmatrix} a_{22} & a_{23} & \dots & a_{2,n-2} & a_{2,n-1} & a_{2n} \\ 0 & a_{33} & \dots & a_{3,n-2} & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{n-2,n-2} & a_{n-2,n-1} & a_{n-2,n} \\ 0 & 0 & \dots & 0 & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \dots & 0 & 0 & a_{nn} \end{vmatrix} \\
&= a_{11}a_{22} \begin{vmatrix} a_{33} & \dots & a_{3,n-2} & a_{3,n-1} & a_{3n} \\ \vdots & & \vdots & \vdots & \vdots \\ 0 & \dots & a_{n-2,n-2} & a_{n-2,n-1} & a_{n-2,n} \\ 0 & \dots & 0 & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \dots & 0 & 0 & a_{nn} \end{vmatrix} \\
&\;\vdots \\
&= a_{11}a_{22}a_{33}\cdots a_{nn}\,. \quad \square
\end{aligned}$$

Example.

We will now compute the 3×3 Vandermonde determinant, $\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{vmatrix}$. We have,

$$\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ x & y-x & z-x \\ x^2 & y^2-x^2 & z^2-x^2 \end{vmatrix} \qquad \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$= \begin{vmatrix} y-x & z-x \\ y^2-x^2 & z^2-x^2 \end{vmatrix}$$

$$= \begin{vmatrix} y-x & z-x \\ (y-x)(y+x) & (z-x)(z+x) \end{vmatrix}$$

$$= (y-x) \begin{vmatrix} 1 & z-x \\ y+x & (z-x)(z+x) \end{vmatrix}$$

$$= (y-x)(z-x) \begin{vmatrix} 1 & 1 \\ y+x & z+x \end{vmatrix}$$

$$= (y-x)(z-x)(z+x-y-x)$$

$$= (x-y)(y-z)(z-x)\,.$$

Can you compute the general $n \times n$ Vandermonde determinant,

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-2} & x_2^{n-2} & \dots & x_n^{n-2} \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}?$$

Example.

$$\begin{vmatrix} b+c & c+a & b+a \\ a & b & c \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} b+c & c+a-b-c & b+a-b-c \\ a & b-a & c-a \\ 1 & 0 & 0 \end{vmatrix} \qquad \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$= \begin{vmatrix} b+c & a-b & a-c \\ a & -(a-b) & -(a-c) \\ 1 & 0 & 0 \end{vmatrix}$$

$$= (-1)^{1+3} \begin{vmatrix} a-b & a-c \\ -(a-b) & -(a-c) \end{vmatrix}$$

$$= (a-b) \begin{vmatrix} 1 & a-c \\ -1 & -(a-c) \end{vmatrix}$$

$$= (a-b)(a-c) \begin{vmatrix} 1 & 1 \\ -1 & -1 \end{vmatrix}$$

$$= (a-b)(a-c)(-1-(-1))$$

$$= 0\,.$$

Example.

$$\begin{vmatrix} 1 & 2 & -1 & 3 \\ 2 & 6 & 2 & 2 \\ -1 & 3 & 3 & 3 \\ -1 & 1 & 2 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 2 & 4 & -4 \\ -1 & 5 & 2 & 6 \\ -1 & 3 & 1 & 5 \end{vmatrix} \qquad \begin{matrix} R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 + R_1 \\ R_4 \to R_4 - 3R_1 \end{matrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ -1 & 5 & -8 & 16 \\ -1 & 3 & -5 & 11 \end{vmatrix} \qquad \begin{matrix} R_3 \to R_3 - 2R_2 \\ R_4 \to R_4 + 2R_2 \end{matrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ -1 & 5 & -8 & 0 \\ -1 & 3 & -5 & 1 \end{vmatrix} \qquad R_4 \to R_4 + 2R_3$$

$$= 1 \cdot 2 \cdot (-8) \cdot 1$$

$$= -16\,.$$

Example.

$$\begin{vmatrix} 2 & 1 & 3 & 2 \\ 3 & 0 & 1 & -2 \\ 1 & -1 & 4 & 3 \\ 2 & 2 & -1 & 1 \end{vmatrix} = -\begin{vmatrix} 1 & 2 & 3 & 3 \\ 0 & 3 & 1 & -2 \\ -1 & 1 & 4 & 3 \\ 2 & 2 & -1 & 1 \end{vmatrix} \qquad C_1 \leftrightarrow C_2$$

$$= -\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & -2 \\ -1 & 3 & 7 & 5 \\ 2 & -2 & -7 & -3 \end{vmatrix} \qquad \begin{array}{l} C_2 \to C_2 - 2C_1 \\ C_3 \to C_3 - 3C_1 \\ C_4 \to C_4 - 2C_1 \end{array}$$

$$= -\begin{vmatrix} 3 & 1 & -2 \\ 3 & 7 & 5 \\ -2 & -7 & -3 \end{vmatrix}$$

$$= (-1)^2 \begin{vmatrix} 1 & 3 & -2 \\ 7 & 3 & 5 \\ -7 & -2 & -3 \end{vmatrix} \qquad C_1 \leftrightarrow C_1$$

$$= \begin{vmatrix} 1 & 0 & 0 \\ 7 & -18 & 19 \\ -7 & 19 & -17 \end{vmatrix} \qquad \begin{array}{l} C_2 \to C_2 - 3C_1 \\ C_3 \to C_3 + 2C_1 \end{array}$$

$$= \begin{vmatrix} -18 & 19 \\ 19 & -17 \end{vmatrix}$$

$$= (-18)(-17) - 19 \cdot 19$$

$$= 306 - 361$$

$$= -55\,.$$

Example.

$$\begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} = -\begin{vmatrix} a+b+c & a+b+c & a+b+c \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} \qquad R_1 \to R_1 + R_2 + R_3$$

$$= (a+b+c)\begin{vmatrix} 1 & 1 & 1 \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix}$$

$$= (a+b+c)\begin{vmatrix} 1 & 0 & 0 \\ 2b & -b-c-a & 0 \\ 2c & 0 & -c-a-b \end{vmatrix} \qquad \begin{array}{l} C_2 \to C_2 - C_1 \\ C_3 \to C_3 - C_1 \end{array}$$

$$= (a+b+c)\begin{vmatrix} -b-c-a & 0 \\ 0 & -c-a-b \end{vmatrix}$$

$$= (a+b+c)(-b-c-a)(-c-a-b)$$

$$= (a+b+c)^3\,.$$

Section 10 – Matrix Representation of Geometric Transformations

10.1. Definition of a geometric transformation induced by a matrix.
We will denote the Euclidean plane $\{(x,y) \mid x, y \in \mathbb{R}^2\}$ by

$$\mathbb{R}^2 = \{(x,y) \mid x, y \in \mathbb{R}^2\}\,.$$

Let $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ be a 2×2 matrix. The matrix A corresponds in a natural way to a map from the plane into the place, namely the following map

$$\begin{pmatrix} x \\ y \end{pmatrix} \to A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \end{pmatrix}\,. \tag{10.1}$$

The map in (10.1) is called the geometric transformation induced by A and is also denoted by A. We will now look a various examples of transformations induced by matrices.

10.2. Homotheties.

The map that stretches all vectors with the amount λ is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\,.$$

Hence, the matrix

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

corresponds to the map that stretches all vectors with the amount λ. This map is called a homothety with ratio λ.
If $\lambda = 1$, then $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and the map induced by A is the identity map,

$$\begin{pmatrix} x \\ y \end{pmatrix} \to A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}\,.$$

If $\lambda = 0$, then $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, and the map induced by A is the zero map,

$$\begin{pmatrix} x \\ y \end{pmatrix} \to A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\,.$$

10.3. Reflections.

Reflection in the x-axis is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Hence, the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

represents a reflection in the x-axis.

Reflection in the y-axis is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} -x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Hence, the matrix

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

represents a reflection in the y-axis.

Reflection in the diagonal is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Hence, the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

represents a reflection in the diagonal.

10.4. Projections.

Projection onto the x-axis is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ 0 \end{pmatrix} .$$

Hence, the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

represents a projection onto the x-axis.

Projection onto the y-axis is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 \\ y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ y \end{pmatrix} .$$

Hence, the matrix

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

represents a projection onto the y-axis.

We will now find the matrix that represents the projection onto a general line ℓ. Let the directional cosine of the line ℓ equal $\cos(\theta)$. The equation for ℓ is therefore

$$\ell : x \sin(\theta) = y \cos(\theta) .$$

Let us temperarily use the notation

$$\begin{pmatrix} u \\ v \end{pmatrix} = \text{the projection of } \begin{pmatrix} x \\ y \end{pmatrix} \text{ onto the line } \ell \,.$$

We now have,

$$\begin{pmatrix} u \\ v \end{pmatrix} \in \ell \;\Rightarrow\; u\sin(\theta) = v\cos(\theta)\,, \tag{10.2}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} \perp \begin{pmatrix} u-x \\ v-y \end{pmatrix} \;\Rightarrow\; 0 = \begin{pmatrix} u \\ v \end{pmatrix} \bullet \begin{pmatrix} u-x \\ v-y \end{pmatrix} = u(u-x) + v(v-y) \;\Rightarrow\; ux + vy = u^2 + v^2\,. \tag{10.3}$$

We now have 2 equations, (10.2) and (10.3), for the two unknowns u and v. We now simply solve (10.2) and (10.3) for u and v. Substituting (10.2) into (10.3) gives:

$$v\frac{\cos(\theta)}{\sin(\theta)}x + vy = \left(\frac{\cos(\theta)}{\sin(\theta)}v\right)^2 + v^2$$

$$\Downarrow$$

$$\frac{\cos(\theta)}{\sin(\theta)}x + y = \frac{\cos^2(\theta)}{\sin^2(\theta)}v + v$$

$$\Downarrow$$

$$\cos(\theta)\sin(\theta)x + \sin^2(\theta)y = \cos^2(\theta)v + \sin^2(\theta)v = v\,,$$

and similarly

$$\cos^2(\theta)x + \cos(\theta)\sin(\theta)y = u\,.$$

The projection onto ℓ is therefore given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos^2(\theta)x + \cos(\theta)\sin(\theta)y \\ \cos(\theta)\sin(\theta)x + \sin^2(\theta)y \end{pmatrix} = \begin{pmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

i.e. the matrix

$$A = \begin{pmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{pmatrix}$$

represents the projection onto the line ℓ.

10.5. Rotations.

We will now find the matrix that represents counterclockwise rotation through the angle θ. Let us temperarily use the notation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos(\varphi) \\ r\sin(\varphi) \end{pmatrix}$$

Rotation counterclockwise through the angle θ is now given by

$$\begin{aligned}
\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} r\cos(\varphi+\theta) \\ r\cos(\varphi+\theta) \end{pmatrix} \\
&= \begin{pmatrix} r(\cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi)) \\ r(\sin(\theta)\cos(\varphi) + \cos(\theta)\sin(\varphi)) \end{pmatrix} \\
&= \begin{pmatrix} x\cos(\theta) - y\sin(\theta) \\ x\sin(\theta) + y\cos(\theta) \end{pmatrix} \\
&= \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},
\end{aligned}$$

i.e. the matrix

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

represents rotation counterclockwise through the angle θ.

Exercises

1. Let

$$A = \begin{pmatrix} 1 & 2 \\ -3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 0 & 3 \\ -1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}.$$

Decide if the following matrices are defined, and evaluate those that are:

$$A + B\,,\ A + C\,,\ 2A + 3C\,,\ AB\,,\ BA\,,\ AC\,,\ CA\,.$$

2. Let

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 4 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & -2 \\ 3 & 2 \end{pmatrix}.$$

Verify that $(A + B)C = AC + BC$ and that $(AB)C = A(BC)$.

3. Let

$$A = \begin{pmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & -1 \\ 0 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 5 & 2 \\ -1 & 0 & 1 \\ 3 & 2 & 4 \end{pmatrix}, \quad E = \begin{pmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{pmatrix}.$$

Compute the following matrices,

$$D + E\,,\ D - E\,,\ 5A\,,\ -7C\,,\ 2B - C\,,\ 4E - 2D\,,\ -3(D + 2E)\,,\ A - A\,.$$

4. Let

$$A = \begin{pmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & -1 \\ 0 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 5 & 2 \\ -1 & 0 & 1 \\ 3 & 2 & 4 \end{pmatrix}, \quad E = \begin{pmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{pmatrix}.$$

Decide if the following matrices are defined, and evaluate those that are:

$$AB\,,\ BA\,,\ (3E)D\,,\ (AB)C\,,\ A(BC)\,.$$

5. For a real number a define the matrix $\hat{a}$ by

$$\hat{a} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}.$$

Also define the matrix i by

$$i = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

For real numbers a, b and c show that

$$i^2 = \widehat{-1}\,, \tag{1}$$

$$(\hat{a} + i\hat{b}) + (\hat{c} + i\hat{d}) = (\hat{a} + \hat{c}) + i(\hat{b} + \hat{d})\,, \tag{2}$$

$$(\hat{a} + i\hat{b})(\hat{c} + i\hat{d}) = (\hat{a}\hat{c} - \hat{b}\hat{d}) + i(\hat{a}\hat{d} + \hat{b}\hat{c})\,. \tag{3}$$

Do equations (1), (2) and (3) remind you of something?

6. Let

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}.$$

Let $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ denote the 2×2 identity matrix and let $0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ denote the 2×2 zero matrix. Show that, if λ is a real number, then

$$\det(A - \lambda I) = \lambda^2 - 6\lambda + 1\,.$$

Verify that

$$A^2 - 6A + I = 0\,.$$

(This is *not* a coincidence. The interested student who wants to pursue this topic further is adviser to look for the Cayley-Hamilton Theorem in any textbook in linear algebra.)

7. Let A and B be $m \times n$ matrices and let C be a $n \times p$ matrix. Show that

$$(A + B)C = AC + BC\,.$$

8. Let A be a $m \times n$ matrix, let B be a $n \times p$ matrix and let C be a $p \times q$ matrix. Show that

$$(AB)C = A(BC)\,.$$

9. Evaluate the determinants

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}, \quad \begin{vmatrix} 5 & 1 \\ 1 & 0 \end{vmatrix}, \quad \begin{vmatrix} 6 & 2 \\ -1 & -1 \end{vmatrix}, \quad \begin{vmatrix} 2 & 7 \\ 0 & -2 \end{vmatrix}.$$

10. Evaluate the determinants

$$\begin{vmatrix} 1 & 2 \\ -3 & 5 \end{vmatrix}, \quad \begin{vmatrix} 3 & 1 & 1 \\ 1 & 5 & 1 \\ 3 & 0 & 2 \end{vmatrix}.$$

11. Evaluate the determinants

$$\begin{vmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ -1 & 1 & 3 \end{vmatrix}, \quad \begin{vmatrix} 2 & 2 & 1 \\ 5 & 3 & 1 \\ 0 & 0 & 1 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & 3 \\ 0 & 2 & 3 \\ 1 & -1 & 0 \end{vmatrix}.$$

12. Evaluate the determinants

$$\begin{vmatrix} a-5 & 5 \\ -3 & a-2 \end{vmatrix}, \quad \begin{vmatrix} \sqrt{2} & \sqrt{6} \\ 4 & \sqrt{3} \end{vmatrix}, \quad \begin{vmatrix} c & -4 & 3 \\ 2 & 1 & c^2 \\ 4 & c-1 & 2 \end{vmatrix}.$$

13. Find all values of λ for which $\det A = 0$ when

(a) $A = \begin{pmatrix} \lambda - 5 & 1 \\ -5 & \lambda + 4 \end{pmatrix}$.

(b) $A = \begin{pmatrix} \lambda - 4 & 0 & 0 \\ 0 & \lambda & 2 \\ 0 & 3 & \lambda - 1 \end{pmatrix}$.

14. Solve the following equation for x,

$$\begin{vmatrix} x & -1 \\ 3 & 1-x \end{vmatrix} = \begin{vmatrix} 1 & 0 & -3 \\ 2 & x & -6 \\ 1 & 3 & x-5 \end{vmatrix}.$$

15. Let A and B be 2×2 matrices. Show that

$$\det(AB) = \det(A)\det(B).$$

16. Let x, y and z be real numbers. Evaluate the following determinants

(a) $\begin{vmatrix} 1+x & y & z \\ x & 1+y & z \\ x & y & 1+z \end{vmatrix}$.

(b) $\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^3 & y^3 & z^3 \end{vmatrix}$.

(Hint: The result is $(x-y)(y-z)(z-x)(x+y+z)$. Use appropiate row or column operations together with the following algebraic identity: $u^3 - v^3 = (u^2 + v^2 + uv)(u - v)$ for all $u, v \in \mathbb{R}$.)

17. Solve the following equation for x,

$$\begin{vmatrix} 1 & 0 & 2 \\ 2 & x & 2 \\ 2 & 1 & x+1 \end{vmatrix} = 0\,.$$

For which x is the matrix $\begin{pmatrix} 1 & 0 & 2 \\ 2 & x & 2 \\ 2 & 1 & x+1 \end{pmatrix}$ invertible?

18. For which a is the matrix $\begin{pmatrix} a-5 & 5 \\ -3 & a-2 \end{pmatrix}$ invertible?

19. For which c is the matrix $\begin{pmatrix} c & -4 & 3 \\ 2 & 1 & c^2 \\ 4 & c-1 & 2 \end{pmatrix}$ invertible?

20. Show that the matrix $\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 5 \\ 2 & 1 & 1 \end{pmatrix}$ does not have an inverse.

21. Find the inverse of the following matrices:

$$\begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 5 & 2 \\ 1 & 3 \end{pmatrix}.$$

22. Find the inverse of $A = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$. Hence solve the systems of equations

(a) $\begin{array}{ccccc} 3x & + & 2y & = & 1 \\ x & + & y & = & 2 \end{array}.$

(b) $\begin{array}{ccccc} 3x & + & 2y & = & 7 \\ x & + & y & = & -3 \end{array}.$

23. Find the inverse of the following matrices:

$$\begin{pmatrix} 2 & -1 & 3 \\ 0 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 4 & 2 & -1 \\ -5 & -3 & 1 \\ 3 & 2 & 0 \end{pmatrix}.$$

24. Solve the systems of equations

$$\begin{array}{rcrcrcr} x & + & 4y & + & 3z & = & 10 \\ 2x & + & y & - & 5z & = & -1 \\ 3x & - & y & + & z & = & 11 \end{array} .$$

25. Solve the systems of equations

$$\begin{array}{rcrcrcr} x & + & y & + & z & = & 2 \\ 2x & - & y & - & 2z & = & -1 \\ 2x & - & 2y & + & z & = & 6 \end{array} .$$

26. What geometrical transformations are represented by the matrices

(a) $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$.

(b) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

(c) $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.