Linear algebra

Teaching script for Computer Science students

Grzegorz Dymek

KUL 2024

Preliminaries

The teaching script is created from lectures of the course of Linear algebra, which author have on KUL for Computer Science students. The course consists of only 15 hours of lectures, which means that there is not time for many interesting topics. There are presented complex numbers, matrices and determinants, systems of linear equations and polynomials. Discussed notions are given in understanding form and often illustrated by examples. Author hopes that the teaching script will be helpfull for student.

1. Complex numbers

Consider the set

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

Let $z_1 = (x_1, y_1), z_2 = (x_2, y_2) \in \mathbb{R}^2$. We define the equality of z_1, z_2 as follows

$$(x_1, y_1) = (x_2, y_2) \underset{df}{\Leftrightarrow} x_1 = x_2 \text{ and } y_1 = y_2$$

and the addition and multiplication operations on $\mathbb{R}^2$ in the following way

$$z_1 + z_2 = (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

and

$$z_1 \cdot z_2 = (x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1).$$

Then $z_1 + z_2$ is the *sum* of z_1, z_2, and $z_1 \cdot z_2$ is the *product* of z_1, z_2.

Example. Let $z_1 = (-5, 6)$ and $z_2 = (1, -2)$. Then

$$z_1 + z_2 = (-5, 6) + (1, -2) = (-4, 4)$$

and

$$z_1 \cdot z_2 = (-5, 6) \cdot (1, -2) = (7, 16).$$

Definition. The set $\mathbb{R}^2$ together with the addition and multiplication operations, is called the *set of complex numbers*, denoted by $\mathbb{C}$. Any element $z = (x, y) \in \mathbb{C}$ is called a *complex number.*

Theorem. (**Properties of addition**) Let $z, z_1, z_2, z_3 \in \mathbb{C}$. Then

1) $z_1 + z_2 = z_2 + z_1$ (*commutative law*),
2) $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$ (*associative law*),
3) $z + 0 = 0 + z = z$, where $0 = (0, 0) \in \mathbb{C}$ (*additive identity*),
4) $\bigwedge_{z=(x,y)\in\mathbb{C}} \bigvee_{-z=(-x,-y)\in\mathbb{C}} z + (-z) = (-z) + z = 0.$

Proof. Points 1), 3) and 4) are easy to prove. We show the second point. Let $z_1 = (x_1, y_1), z_2 = (x_2, y_2), z_3 = (x_3, y_3) \in \mathbb{C}$. We have

$$\begin{aligned} L = (z_1 + z_2) + z_3 &= ((x_1, y_1) + (x_2, y_2)) + (x_3, y_3) \\ &= (x_1 + x_2, y_1 + y_2) + (x_3, y_3) \\ &= ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3) \end{aligned}$$

and

$$\begin{aligned} R = z_1 + (z_2 + z_3) &= (x_1, y_1) + ((x_2, y_2) + (x_3, y_3)) \\ &= (x_1, y_1) + (x_2 + x_3, y_2 + y_3) \\ &= (x_1 + (x_2 + x_3), y_1 + (y_2 + y_3)). \end{aligned}$$

Since the addition of real numbers is associative, it follows $L = R$. $\square$

Definition. Let $z_1 = (x_1, y_1), z_2 = (x_2, y_2) \in \mathbb{C}$. The *difference* of numbers z_1, z_2 is defined

$$z_1 - z_2 \underset{df}{=} z_1 + (-z_2)$$

and

$$z_1 - z_2 = (x_1, y_1) - (x_2, y_2) \underset{df}{=} (x_1 - x_2, y_1 - y_2).$$

Theorem. (**Properties of multiplication**) Let $z, z_1, z_2, z_3 \in \mathbb{C}$. Then

1) $z_1 \cdot z_2 = z_2 \cdot z_1$ (*commutative law*),
2) $(z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3)$ (*associative law*),
3) $z \cdot 1 = 1 \cdot z = z$, where $1 = (1, 0) \in \mathbb{C}$ (*multipicative identity*),
4) $\bigwedge\limits_{z=(x,y)\in\mathbb{C}\setminus\{0\}} \bigvee\limits_{z^{-1}\in\mathbb{C}} z \cdot z^{-1} = z^{-1} \cdot z = 1.$

Proof. Points 1) and 2) are easy to show. We prove points 3) and 4).

3) Let $z = (x, y) \in \mathbb{C}$. We have

$$z \cdot 1 = (x, y) \cdot (1, 0) = (x \cdot 1 - y \cdot 0, x \cdot 0 + 1 \cdot y) = (x, y) = z$$

and

$$1 \cdot z = (1, 0) \cdot (x, y) = (1 \cdot x - 0 \cdot y, 1 \cdot y + x \cdot 0) = (x, y) = z.$$

4) Let $z^{-1} = (x', y') \in \mathbb{C}$. Since $(x, y) \neq (0, 0)$, it follows $x \neq 0$ or $y \neq 0$. Hence $x^2 + y^2 \neq 0$. Thus $z \cdot z^{-1} = 1$ implies $(x, y) \cdot (x', y') = (1, 0)$, that is, $(x \cdot x' - y \cdot y', x \cdot y' + x' \cdot y) = (1, 0)$. Hence we get

$$\begin{cases} x \cdot x' - y \cdot y' = 1, \\ x \cdot y' + x' \cdot y = 0. \end{cases}$$

Solving that system with respect to x' and y' we obtain

$$x' = \frac{x}{x^2 + y^2} \quad \text{and} \quad y' = -\frac{y}{x^2 + y^2}.$$

Hence for $z = (x, y) \in \mathbb{C}\setminus\{0\}$ we have

$$z^{-1} = \frac{1}{z} = \left(\frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2} \right) \in \mathbb{C}\setminus\{0\}.$$

By the commutative law we also have $z^{-1} \cdot z = 1$. $\square$

Definition. Let $z_1 = (x_1, y_1) \in \mathbb{C}$, $z = (x, y) \in \mathbb{C}\backslash\{0\}$. Then the following is the *quotient of* numbers z_1 and z:

$$\frac{z_1}{z} = z_1 \cdot z^{-1} = (x_1, y_1) \cdot \left(\frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2} \right) = \left(\frac{x_1 x + y_1 y}{x^2 + y^2}, \frac{-x_1 y + x y_1}{x^2 + y^2} \right) \in \mathbb{C}.$$

Definition. An *integer power* of a complex number $z \in \mathbb{C}\backslash\{0\}$ is defined by

$$z^0 = 1,\ z^1 = z,\ z^2 = z \cdot z$$
$$z^n = \underbrace{z \cdot z \cdot \ldots \cdot z}_{n \text{ times}} \quad \text{for all } n = 1, 2, \ldots$$

and

$$z^n = \left(z^{-1}\right)^{-n} \quad \text{for all } n = -1, -2, \ldots$$

When $z = 0$, we define $0^n = 0$ for all $n = 1, 2, \ldots$

Theorem. Let $z, z_1, z_2 \in \mathbb{C}\backslash\{0\}$ and $m, n \in \mathbb{Z}$. Then

1) $z^m \cdot z^n = z^{m+n}$,
2) $\frac{z^m}{z^n} = z^{m-n}$,
3) $(z^m)^n = x^{mn}$,
4) $(z_1 \cdot z_2)^n = z_1^n \cdot z_2^n$,
5) $\left(\frac{z_1}{z_2}\right)^n = \frac{z_1^n}{z_2^n}$.

Proof. Easy. $\square$

Yet we have the following property of operations of complex numbers.

Theorem. Let $z_1, z_2, z_3 \in \mathbb{C}$. Then the following distributive law holds:

$$z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3.$$

Proof. Easy. $\square$

Now, consider the set $\mathbb{R} \times \{0\}$ together with the addition and multiplication operations defined on $\mathbb{R}^2$. The function

$$f : \mathbb{R} \to \mathbb{R} \times \{0\}, \quad f(x) = (x, 0)$$

is bijective (that is, it is one-to-one and onto) and moreover

$$(x, 0) + (y, 0) = (x + y, 0)$$

and

$$(x, 0) \cdot (y, 0) = (xy, 0).$$

We see that the algebraic operations on $\mathbb{R} \times \{0\}$ are similar to the operations on $\mathbb{R}$. Therefore we can identify the ordered pair $(x, 0)$ with the number x for all $x \in \mathbb{R}$. Hence we can write

$$(x, 0) = x.$$

Define

$$i \underset{df}{=} (0, 1).$$

Theorem. Any complex number $z = (x, y)$ can be uniquely represented in the form

$$z = x + yi,$$

where $x, y \in \mathbb{R}$. Moreover, $i^2 = -1$.

Proof. We have

$$z = (x, y) = (x, 0) + (0, y) = (x, 0) + (y, 0) \cdot (0, 1) = x + yi = (x, 0) + (0, 1) \cdot (y, 0) = x + iy.$$

Moreover,

$$i^2 = i \cdot i = (0, 1) \cdot (0, 1) = (-1, 0) = -1. \quad \square$$

The form $z = x + yi$ is called the *algebraic form* of a complex number $z = (x, y)$. Then $x = \mathrm{Re}(z)$ is the *real part* of z, and $y = \mathrm{Im}(z)$ is the *imaginary part* of z. The complex number i is called the *imaginary unit*.

Theorem. Let $z, z_1, z_2 \in \mathbb{C}$. Then

1) $z_1 = z_2 \;\Leftrightarrow\; \mathrm{Re}(z_1) = \mathrm{Re}(z_2) \;\wedge\; \mathrm{Im}(z_1) = \mathrm{Im}(z_2)$,
2) $z \in \mathbb{R} \;\Leftrightarrow\; \mathrm{Im}(z) = 0$,
3) $z \in \mathbb{C} \backslash \mathbb{R} \;\Leftrightarrow\; \mathrm{Im}(z) \neq 0$.

Proof. Easy. $\square$

Now we look at operations on complex numbers in the algebraic form.

1) Addition:

$$z_1 + z_2 = (x_1 + y_1 i) + (x_2 + y_2 i) = (x_1 + x_2) + (y_1 + y_2)i.$$

So,

$$\mathrm{Re}(z_1 + z_2) = \mathrm{Re}(z_1) + \mathrm{Re}(z_2)$$

and

$$\mathrm{Im}(z_1 + z_2) = \mathrm{Im}(z_1) + \mathrm{Im}(z_2).$$

2) Multiplication:

$$z_1z_2 = (x_1 + y_1i)(x_2 + y_2i) = (x_1x_2 - y_1y_2) + (x_1y_2 + x_2y_1)i.$$

So,

$$\mathrm{Re}(z_1 \cdot z_2) = \mathrm{Re}(z_1)\mathrm{Re}(z_2) - \mathrm{Im}(z_1)\mathrm{Im}(z_2)$$

and

$$\mathrm{Im}(z_1z_2) = \mathrm{Re}(z_1)\mathrm{Im}(z_2) + \mathrm{Re}(z_2)\mathrm{Im}(z_1).$$

In particular,

$$\lambda z = \lambda(x + yi) = \lambda x + \lambda yi$$

for $\lambda \in \mathbb{R}$ and $z \in \mathbb{C}$.

Theorem. Let $z, z_1, z_2 \in \mathbb{C}$ and $\lambda, \lambda_1, \lambda_2 \in \mathbb{R}$. Then

1) $\lambda(z_1 + z_2) = \lambda z_1 + \lambda z_2$,
2) $\lambda_1(\lambda_2 z) = (\lambda_1\lambda_2)z$,
3) $(\lambda_1 + \lambda_2)z = \lambda_1 z + \lambda_2 z$.

Proof. Easy. $\square$

Now, we define the notion of a conjugate of a complex number.

Definition. Let $z = x + yi \in \mathbb{C}$. The *conjugate* of z is a complex number:

$$\overline{z} \underset{df}{=} x - yi \in \mathbb{C}.$$

Theorem. (**Properties of conjugate**) Let $z, z_1, z_2 \in \mathbb{C}$. Then

1) $z = \overline{z} \iff z \in \mathbb{R}$,
2) $z = \overline{\overline{z}}$,
3) $z\overline{z} \geq 0$,
4) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$,
5) $\overline{z_1z_2} = \overline{z_1} \cdot \overline{z_2}$,
6) $\overline{z^{-1}} = \overline{z}^{-1}, \quad z \neq 0$,
7) $\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\overline{z_1}}{\overline{z_2}}, \quad z_2 \neq 0$,
8) $\mathrm{Re}(z) = \frac{z+\overline{z}}{2}$ and $\mathrm{Im}(z) = \frac{z-\overline{z}}{2i}$.

Proof. Let $z = x + yi, z_1 = x_1 + y_1i, z_2 = x_2 + y_2i$, where $x, x_1, x_2, y, y_1, y_2 \in \mathbb{R}$. Then

1) The following steps are equivalent:

$$\begin{aligned} z &= \overline{z} \\ x + yi &= x - yi \\ 2yi &= 0 \\ y &= 0 \\ z &= x \in \mathbb{R}. \end{aligned}$$

2) Since $z = x + yi$, we get $\overline{z} = x - yi$, and consequently

$$\overline{\overline{z}} = x - (-yi) = x + yi = z.$$

3) It can be easily seen that

$$z\overline{z} = (x + yi)(x - yi) = x^2 + y^2 \geq 0.$$

4) We have

$$\begin{aligned} \overline{z_1 + z_2} &= \overline{(x_1 + y_1 i) + (x_2 + y_2 i)} \\ &= (x_1 + x_2) - (y_1 + y_2)i \\ &= (x_1 - y_1 i) + (x_2 - y_2 i) \\ &= \overline{z_1} + \overline{z_2}. \end{aligned}$$

5) Similarly,

$$\begin{aligned} \overline{z_1 z_2} &= \overline{(x_1 x_2 - y_1 y_2) + (x_1 y_2 + x_2 y_1)i} \\ &= (x_1 x_2 - y_1 y_2) - (x_1 y_2 + x_2 y_1)i \\ &= (x_1 - y_1 i)(x_2 - y_2 i) \\ &= \overline{z_1} \cdot \overline{z_2}. \end{aligned}$$

6) Let $z \neq 0$. Because $z \cdot \frac{1}{z} = 1$, we have $\overline{\left(z \cdot \frac{1}{z}\right)} = \overline{1}$. Hence $\overline{z} \cdot \overline{\left(\frac{1}{z}\right)} = 1$, that is, $\overline{z^{-1}} = \overline{z}^{-1}$.
7) Let $z_2 \neq 0$. Now,

$$\begin{aligned} \overline{\left(\frac{z_1}{z_2}\right)} = \overline{\left(z_1 \cdot \frac{1}{z_2}\right)} &= \overline{z_1} \cdot \overline{\left(\frac{1}{z_2}\right)} \\ = \overline{z_1} \cdot \left(\frac{1}{\overline{z_2}}\right) &= \frac{\overline{z_1}}{\overline{z_2}}. \end{aligned}$$

8) We have

$$z + \overline{z} = (x + yi) + (x - yi) = 2x = 2\mathrm{Re}(z)$$

and

$$z - \overline{z} = (x + yi) - (x - yi) = 2yi = 2i\mathrm{Im}(z).$$

Hence,

$$\operatorname{Re}(z) = \frac{z+\overline{z}}{2} \quad \text{and} \quad \operatorname{Im}(z) = \frac{z-\overline{z}}{2i}. \quad \square$$

Remark. The properties 4) and 5) can be easily extended to give

4') $\overline{\left(\sum\limits_{k=1}^{n} z_k\right)} = \sum\limits_{k=1}^{n} \overline{z_k}$,

5') $\overline{\left(\prod\limits_{k=1}^{n} z_k\right)} = \prod\limits_{k=1}^{n} \overline{z_k}$

for all $z_k \in \mathbb{C}$, $k = 1, 2, \ldots, n$.

As a consequence of 5') and 6) we have

5") $\overline{z^n} = \overline{z}^n$

for any $z \in \mathbb{C}$ and $n \in \mathbb{Z}$.

Remark. Let $z = x + yi, z_1 = x_1 + y_1 i, z_2 = x_2 + y_2 i \in \mathbb{C}$ and $z, z_2 \neq 0$. Then

$$\frac{1}{z} = \frac{\overline{z}}{z \cdot \overline{z}} = \frac{x - yi}{x^2 + y^2} = \frac{x}{x^2 + y^2} - \frac{y}{x^2 + y^2} i$$

and

$$\frac{z_1}{z_2} = \frac{z_1 \cdot \overline{z_2}}{z_2 \cdot \overline{z_2}} = \frac{(x_1 + y_1 i)(x_2 - y_2 i)}{x_2^2 + y_2^2} = \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} + \frac{-x_1 y_2 + x_2 y_1}{x_2^2 + y_2^2} i.$$

Definition. Let $z = x + yi \in \mathbb{C}$. The *modulus* of z is a number:

$$|z| \underset{df}{=} \sqrt{x^2 + y^2} \in \mathbb{R}.$$

Theorem. (**Properties of modulus**) Let $z, z_1, z_2 \in \mathbb{C}$. Then

1) $-|z| \leq \operatorname{Re}(z) \leq |z|$ and $-|z| \leq \operatorname{Im}(z) \leq |z|$,
2) $|z| \geq 0$ and $|z| = 0 \Leftrightarrow z = 0$,
3) $|z| = |-z| = |\overline{z}|$,
4) $z \cdot \overline{z} = |z|^2$,
5) $|z_1 z_2| = |z_1||z_2|$,
6) $|z_1| - |z_2| \leq |z_1 + z_2| \leq |z_1| + |z_2|$,
7) $|z^{-1}| = |z|^{-1}$, $\quad z \neq 0$,
8) $\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$, $\quad z_2 \neq 0$,
9) $|z_1| - |z_2| \leq |z_1 - z_2| \leq |z_1| + |z_2|$.

Proof. Points 1)–4) are easy. We prove next properties.

5) Since

$$|z_1 z_2|^2 = (z_1 z_2)(\overline{z_1 z_2}) = (z_1 \overline{z_1})(z_2 \overline{z_2}) = |z_1|^2 |z_2|^2,$$

we get, by 2), $|z_1 z_2| = |z_1||z_2|$.

6) Observe that

$$|z_1 + z_2|^2 = (z_1 + z_2)(\overline{z_1 + z_2}) = (z_1 + z_2)(\overline{z_1} + \overline{z_2}) = |z_1|^2 + z_1\overline{z_2} + \overline{z_1}z_2 + |z_2|^2.$$

Because $\overline{z_1\overline{z_2}} = \overline{z_1} \cdot \overline{\overline{z_2}} = \overline{z_1}z_2$, it follows that

$$z_1\overline{z_2} + \overline{z_1}z_2 = z_1\overline{z_2} + \overline{z_1\overline{z_2}} = 2\mathrm{Re}(z_1\overline{z_2}) \leq 2|z_1\overline{z_2}| = 2|z_1||z_2|.$$

Hence

$$|z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2,$$

that is,

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

For inequality on the left-hand side note that

$$|z_1| = |z_1 + z_2 + (-z_2)| \leq |z_1 + z_2| + |-z_2| = |z_1 + z_2| + |z_2|.$$

Hence

$$|z_1| - |z_2| \leq |z_1 + z_2|.$$

7) We know that $z \cdot \frac{1}{z} = 1$. So $|z| \cdot \left|\frac{1}{z}\right| = 1$, that is, $\left|\frac{1}{z}\right| = \frac{1}{|z|}$. Hence

$$\left|z^{-1}\right| = |z|^{-1}.$$

8) We easily get

$$\left|\frac{z_1}{z_2}\right| = \left|z_1\frac{1}{z_2}\right| = \left|z_1 z_2^{-1}\right| = |z_1|\left|z_2^{-1}\right| = |z_1|\,|z_2|^{-1} = \frac{|z_1|}{|z_2|}.$$

9) We have

$$|z_1| = |z_1 - z_2 + z_2| \leq |z_1 - z_2| + |z_2|.$$

So

$$|z_1| - |z_2| \leq |z_1 - z_2|.$$

On the other hand

$$|z_1 - z_2| = |z_1 + (-z_2)| \leq |z_1| + |-z_2| = |z_1| + |z_2|. \quad \square$$

Remarks.

1) It is not difficult to see that

$$|z_1 + z_2| = |z_1| + |z_2| \;\Leftrightarrow\; \mathrm{Re}(z_1\overline{z_2}) = |z_1||z_2| \;\Leftrightarrow\; z_1 = tz_2 \text{ for some } t \geq 0.$$

2) The properties 5) and 6) can be easily extended to give

5') $\left|\prod_{k=1}^{n} z_k\right| = \prod_{k=1}^{n} |z_k|,$

6') $\left|\sum_{k=1}^{n} z_k\right| \leq \sum_{k=1}^{n} |z_k|,$

for all $z_k \in \mathbb{C}$, $k = 1, 2, \dots, n$.

As a consequence of 5') and 7) we have

5") $|z^n| = |z|^n$

for any $z \in \mathbb{C}$ and $n \in \mathbb{Z}$.

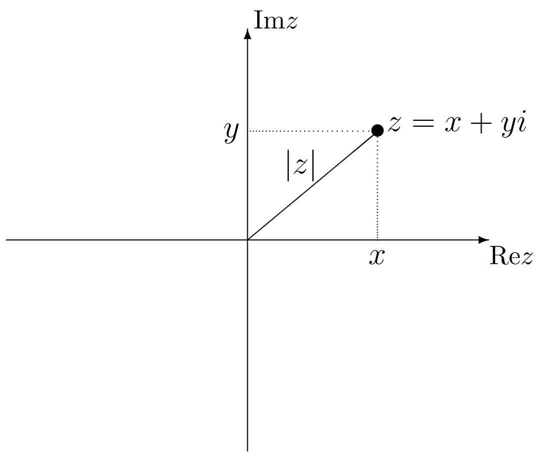
Next we give geometric interpretation of a complex number. We have defined a complex number $z = (x, y) = x + yi$ to be an ordered pair $(x, y) \in \mathbb{R}^2$. So it is natural to let a complex number $z = x + yi$ correspond to a point (x, y) on the plane $\mathbb{R}^2$.

Consider a plane equipped with the coordinate system xoy.

Definition. The point (x, y) is called the *geometric image of the complex number* $z = x + yi$.

The x-axis is called the *real axis* and the y-axis is called the *imaginary axis*. The plane whose points are identified with complex numbers is called the *complex plane*.

Thus we have:



Definition. For a complex number $z = x + yi$ we can write the following *trigonometric form*:

$$z = r(\cos\varphi + i\sin\varphi),$$

where $r \geq 0$ and $\varphi \in [0, 2\pi)$. Then φ is called the *argument* of z, denoted $\arg(z)$ and $r = |z|$. For $z \neq 0$, the modulus and argument of z are uniquely determined.

Consider $z = r(\cos\varphi + i\sin\varphi)$ and let $\varphi' = \varphi + 2k\pi$ for $k \in \mathbb{Z}$. Then

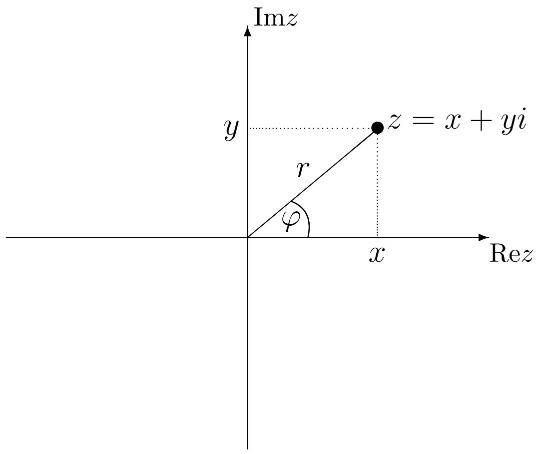
$$z = r[\cos(\varphi' - 2k\pi) + i\sin(\varphi' - 2k\pi)] = r(\cos\varphi' + i\sin\varphi'),$$

that is, any complex number z can be represented as $z = r(\cos\varphi' + i\sin\varphi')$, where $r \geq 0$ and $\varphi' \in \mathbb{R}$. The set

$$\operatorname{Arg}(z) = \{\varphi' : \varphi' = \varphi + 2k\pi,\ k \in \mathbb{Z}\}$$

is called the *extended argument* of the complex number z.

Thus we have:



and we see that

$$\cos\varphi = \frac{x}{r} \quad \text{and} \quad \sin\varphi = \frac{y}{r}.$$

Remark. Note that two complex numbers $z_1, z_2 \neq 0$ represented as $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ and $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$ are equal if and only if $r_1 = r_2$ and $\varphi_1 - \varphi_2 = 2k\pi$, gdzie $k \in \mathbb{Z}$.

Example. For $z = -1 - i$ we have

$$x = -1, \quad y = -1 \quad \text{and} \quad r = |z| = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2}.$$

Then

$$\cos\varphi = \frac{x}{r} = -\frac{1}{\sqrt{2}} = -\frac{\sqrt{2}}{2} \quad \text{and} \quad \sin\varphi = \frac{y}{r} = -\frac{1}{\sqrt{2}} = -\frac{\sqrt{2}}{2}.$$

Hence $\varphi = \frac{5}{4}\pi$ and finally the trigonometric form of z is

$$z = r(\cos\varphi + i\sin\varphi) = \sqrt{2}\left(\cos\frac{5}{4}\pi + i\sin\frac{5}{4}\pi\right).$$

Theorem. Let $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ and $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$. Then

$$z_1 z_2 = r_1 r_2 \left(\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)\right)$$

Proof. Indeed,

$$\begin{aligned} z_1 z_2 &= r_1 r_2 (\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) \\ &= r_1 r_2 (\cos\varphi_1\cos\varphi_2 + i\cos\varphi_1\sin\varphi_2 + i\sin\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) \\ &= r_1 r_2 ((\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + i(\sin\varphi_1\cos\varphi_2 + \sin\varphi_2\cos\varphi_1)) \\ &= r_1 r_2 ((\cos(\varphi_1+\varphi_2) + i\sin(\varphi_1+\varphi_2)). \quad \square \end{aligned}$$

Remarks.

1) We find again that $|z_1 z_2| = |z_1||z_2|$.

2) We have $\arg(z_1 z_2) = \arg(z_1) + \arg(z_2) - 2k\pi$, where

$$k = \begin{cases} 0 & \text{if} \quad \arg(z_1) + \arg(z_2) < 2\pi, \\ 1 & \text{if} \quad \arg(z_1) + \arg(z_2) \geq 2\pi. \end{cases}$$

3) Also we can write

$$\mathrm{Arg}(z_1 z_2) = \{\arg(z_1) + \arg(z_2) + 2k\pi : \ k \in \mathbb{Z}\}.$$

4) Formula from above Theorem can be extended to $n \geq 2$ complex numbers. If $z_k = r_k(\cos\varphi_k + i\sin\varphi_k)$, where $k = 1, 2, \ldots, n$, then

$$z_1 z_2 \cdots z_n = r_1 r_2 \cdots r_n (\cos(\varphi_1 + \varphi_2 + \ldots + \varphi_n) + i\sin(\varphi_1 + \varphi_2 + \ldots + \varphi_n)).$$

Theorem. (**De Moivre**) Let $z = r(\cos\varphi + i\sin\varphi)$ and $n \in \mathbb{N}$. Then

$$z^n = r^n(\cos n\varphi + i\sin n\varphi).$$

Proof. Apply formula from Remark 4) for $z = z_1 = z_2 = \ldots = z_n$ to obtain

$$\begin{aligned} z_n &= \underbrace{r \cdot r \cdot \ldots \cdot r}_{n \text{ times}} (\cos(\underbrace{\varphi + \varphi + \ldots + \varphi}_{n \text{ times}}) + i\sin(\underbrace{\varphi + \varphi + \ldots + \varphi}_{n \text{ times}})) \\ &= r^n(\cos n\varphi + i\sin n\varphi). \quad \square \end{aligned}$$

Remarks.

1) We find again that $|z^n| = |z|^n$.

2) If $r = 1$, then $(\cos\varphi + i\sin\varphi)^n = \cos n\varphi + i\sin n\varphi$.

3) We can write

$$\mathrm{Arg}\left(z^n\right) = \left\{n \arg z + 2k\pi : k \in \mathbb{Z}\right\}.$$

Theorem. Let $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ and $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2) \neq 0$. Then

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}\left(\cos(\varphi_1-\varphi_2) + i\sin(\varphi_1-\varphi_2)\right).$$

Proof. We have

$$\begin{aligned}
\frac{z_1}{z_2} &= \frac{r_1(\cos\varphi_1 + i\sin\varphi_1)}{r_2(\cos\varphi_2 + i\sin\varphi_2)} \\
&= \frac{r_1(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 - i\sin\varphi_2)}{r_2(\cos\varphi_2 + i\sin\varphi_2)(\cos\varphi_2 - i\sin\varphi_2)} \\
&= \frac{r_1}{r_2}\left((\cos\varphi_1\cos\varphi_2 + \sin\varphi_1\sin\varphi_2) + i(\sin\varphi_1\cos\varphi_2 - \sin\varphi_2\cos\varphi_1)\right) \\
&= \frac{r_1}{r_2}\left(\cos(\varphi_1-\varphi_2) + i\sin(\varphi_1-\varphi_2)\right). \quad \square
\end{aligned}$$

Remarks.

1) We find again that $\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$.

2) We can write

$$\mathrm{Arg}\left(\frac{z_1}{z_2}\right) = \{\arg z_1 - \arg z_2 + 2k\pi : k \in \mathbb{Z}\}.$$

3) For $z_1 = 1$ and $z_2 = z$ we get

$$\frac{1}{z} = z^{-1} = \frac{1}{r}\left(\cos(-\varphi) + i\sin(-\varphi)\right).$$

4) De Moivre's formula also holds for negative integer exponents n, that is we have

$$z^n = r^n(\cos n\varphi + i\sin n\varphi) \quad \text{for} \quad n \in \mathbb{Z}.$$

Definition. Let $z \in \mathbb{C}$. Any solution w of the equation

$$w^n = z$$

is called the *nth root of a complex number z*.

Theorem. Let $z = r(\cos\varphi + i\sin\varphi)$ be a complex number with $r > 0$ and $\varphi \in [0, 2\pi)$. The number z has n distinct nth roots given by

$$z_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}\right)$$

for $k = 0, 1, \ldots, n-1$.

Proof. We look for all solutions of the equation $w^n = z$. Let $w = \rho(\cos\alpha + i\sin\alpha)$, where $\rho > 0$ and $\alpha \in \mathbb{R}$. By De Moivre's formula

$$w^n = \rho^n(\cos n\alpha + i\sin n\alpha) = r(\cos\varphi + i\sin\varphi).$$

Hence, $\rho^n = r$ and $n\alpha = \varphi + 2k\pi$, where $k \in \mathbb{Z}$. Thus,

$$\rho = \sqrt[n]{r} \quad \text{and} \quad \alpha = \frac{\varphi + 2k\pi}{n}.$$

Finally, let $z_k = \sqrt[n]{r}\left(\cos\frac{\varphi+2k\pi}{n} + i\sin\frac{\varphi+2k\pi}{n}\right)$ be a root which corresponds with k. Since trigonometric functions are periodic, roots z_k and z_{k+n} do cover. So, there are n distinct roots $z_0, z_1, \ldots, z_{n-1}$ of z. $\square$

Remark. The geometric images of the nth roots of a complex number $z \neq 0$ are the vertices of a regular n-gon inscribed in a circle with center at the origin and radius $\sqrt[n]{r}$.

Example. Let us find the fourth roots of the number $z = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and represent them in the complex plane. We have $r = 1$, $\varphi = \frac{2}{3}\pi$ and

$$z = 1 \cdot \left(\cos\frac{2}{3}\pi + i\sin\frac{2}{3}\pi\right).$$

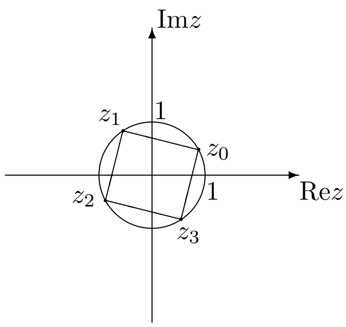
Hence the fourth roots of z are

$$z_k = \sqrt[4]{1}\left(\cos\frac{\frac{2}{3}\pi + 2k\pi}{4} + i\sin\frac{\frac{2}{3}\pi + 2k\pi}{4}\right)$$

for $k = 0, 1, 2, 3$. Thus we have

$$\begin{aligned}
z_0 &= \cos\frac{\pi}{6} + i\sin\frac{\pi}{6} = \frac{\sqrt{3}}{2} + \frac{1}{2}i, \\
z_1 &= \cos\frac{2}{3}\pi + i\sin\frac{2}{3}\pi = -\frac{1}{2} + \frac{\sqrt{3}}{2}i, \\
z_2 &= \cos\frac{7}{6}\pi + i\sin\frac{7}{6}\pi = -\frac{\sqrt{3}}{2} - \frac{1}{2}i, \\
z_3 &= \cos\frac{5}{3}\pi + i\sin\frac{5}{3}\pi = \frac{1}{2} - \frac{\sqrt{3}}{2}i
\end{aligned}$$

and



So we get the regular 4-gon, that is, the square with vertices z_0, z_1, z_2, z_3.

2. Matrices and determinants

Definition. A *matrix* is a rectangular array of (real or complex) numbers arranged in rows and columns.

A general matrix with m rows and n columns is written as follows:

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

The above matrix consists of $m \cdot n$ elements, giving the $m \times n$ array. The symbol $m \times n$ is read "m by n" and is called the *size* of the above matrix. Notation a_{ij} denotes an element in the ith row and jth column of a matrix.

If $m = n$, then a matrix is called a *square matrix*, and the number of rows is called its *order*. A square matrix of order n is also called an $n \times n$ matrix. The diagonal containing elements $a_{11}, a_{22}, \dots, a_{nn}$ of such matrix is called the *main diagonal*. If $m \neq n$, then a matrix is called *rectangular matrix* of an $n \times n$ *matrix*.

Examples.

1) Matrix $\begin{bmatrix} 2 & 1 \\ -1 & 3 \\ 0 & -2 \end{bmatrix}$ is a 3×2 matrix.

2) Matrix $\begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}$ is a 2×2 matrix or a square matrix of order 2.

Matrices will usually be denoted by Roman capital letters: $A, B, C, \dots$, or by $A_{m \times n}$ or $A = [a_{ij}]$ or $A = [a_{ij}]_{m \times n}$.

Let $\mathbb{F} = \mathbb{R}$ (or $\mathbb{C}$). We denote by $M_{m \times n}(\mathbb{F})$ the set of all $m \times n$ matrices with elements in $\mathbb{F}$.

Definition. A square matrix is called *triangular* iff

$$(a_{ij} = 0 \ \text{ for } \ i > j) \quad \text{or} \quad (a_{ij} = 0 \ \text{ for } \ i < j),$$

so

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{nn} \end{bmatrix}$$

The first matrix is called an *upper triangular* matrix and the second is called a *lower triangular* matrix. We see that a triangular matrix is a square matrix whose elements below the main diagonal (or above the main diagonal) are all zero.

Definition. A square matrix of the form

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

is called a *diagonal matrix.* Another notation for the above diagonal matrix is

$$A = \operatorname{diag}[a_{11}a_{22}\cdots a_{nn}].$$

Notice that a diagonal matrix may have nonzero elements only on the main diagonal. It means that a diagonal matrix is a triangular matrix (an upper as well as a lower).

Definition. A square matrix of the form

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

is called a *unit matrix* or an *identity matrix.*

Definition. A matrix with all zero elements is called a *zero matrix* and is denoted by 0, so

$$0 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Notice that a zero matrix may be a rectangular or square matrix.

Definition. Matrices A and B are *equal*, written $A = B$ iff

1) A and B are of the same size,
2) all corresponding elements are equal, so $a_{ij} = b_{ij}$ for all i, j.

Definition. (**Matrix addition**) Let $A = [a_{ij}], B = [b_{ij}] \in M_{m\times n}(\mathbb{F})$. The *sum* $A + B$ is the matrix $C = [c_{ij}] \in M_{m\times n}(\mathbb{F})$ such that

$$c_{ij} = a_{ij} + b_{ij}.$$

Remark. We see that $A + B$ is obtained by adding the corresponding elements of A and B.

Definition. (**Scalar multiplication**) Let $A \in M_{m\times n}(\mathbb{F})$ and $\lambda \in \mathbb{F}$. The *product of* λ *and* A, written λA, is defined to be

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

Theorem. (**Properties of matrix addition and scalar multiplication**) Let $A, B, C \in M_{m\times n}(\mathbb{F})$ and $\lambda, \beta \in \mathbb{F}$. Then

1) $A + B = B + A$,
2) $(A + B) + C = A + (B + C)$,
3) $A + 0 = A$,
4) $1A = A$,
5) $-A = -1A$,
6) $A + (-A) = 0$,
7) $\lambda(A + B) = \lambda A + \lambda B$,
8) $(\lambda + \beta)A = \lambda A + \beta A$,
9) $\lambda(\beta A) = (\lambda\beta)A$.

Proof. Easy. □

Definition. (**Matrix multiplication**) Let $A = [a_{ij}] \in M_{m\times n}(\mathbb{F})$ and $B = [b_{ij}] \in M_{n\times p}(\mathbb{F})$. The *product* $A \cdot B$ is the matrix $C = [c_{ij}] \in M_{m\times p}(\mathbb{F})$ such that

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj},$$

where $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, p$.

Remark. Note that element c_{ij} is obtained by multiplying ith row of A by jth column of B, so it depends on all elements in row i of A and on all elements in column j of B.

Remark. The product $A \cdot B$ of two matrices A and B is defined if number of columns in A equals number of rows in B.

Example. We calculate

$$\overset{3\times 2}{\begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 0 & 2 \end{bmatrix}} \cdot \overset{2\times 3}{\begin{bmatrix} 1 & -2 & 0 \\ 1 & 3 & -1 \end{bmatrix}} = \overset{3\times 3}{\begin{bmatrix} 0 & -5 & 1 \\ 3 & -1 & -1 \\ 2 & 6 & -2 \end{bmatrix}}.$$

Theorem. (**Properties of matrix multiplication**) Let matrices A, B, C be such that sums and products in below properties are defined. Then

1) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$,

2) $A \cdot (B + C) = A \cdot B + A \cdot C$,
3) $(B + C) \cdot A = B \cdot A + C \cdot A$,
4) $\lambda(A \cdot B) = (\lambda A) \cdot B = A \cdot (\lambda B)$, where $\lambda \in \mathbb{F}$.

Proof. Easy. $\square$

Examples.

1) Let us take matrices

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Then

$$A \cdot B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

So, we see that $A \cdot B = 0$ does not necessarily imply $A = 0$ or $B = 0$.

2) Let us take the same matrices A, B. Then

$$A \cdot B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B \cdot A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}.$$

So, we see that products $A \cdot B$ and $B \cdot A$ of matrices A and B need not be equal.

Definition. (**Transpose of a matrix**) Let $A \in M_{m \times n}(\mathbb{F})$. The *transpose* A^T of a matrix A is a matrix obtained by interchanging rows and columns in A.

Remark. Note that if the size of A is $m \times n$, then the size of A^T is $n \times m$.

Example. Let $A = \begin{bmatrix} 2 & 1 \\ -1 & 0 \\ 1 & 1 \end{bmatrix}$. Then $A^T = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

Theorem. Let matrices A, B be such that operations in below properties are defined. Then

1) $(A + B)^T = A^T + B^T$,
2) $\left(A^T\right)^T = A$,
3) $(\lambda A)^T = \lambda A^T$, where $\lambda \in \mathbb{F}$.

Proof. Let $A = [a_{ij}]_{m \times n}$, $B = [b_{ij}]_{m \times n}$, $A^T = [a'_{ij}]_{n \times m}$ and $B^T = [b'_{ij}]_{n \times m}$.

1) Let $C = A + B = [c_{ij}]$ and $C = [c'_{ij}]$. Then

$$c'_{ij} = c_{ji} = a_{ji} + b_{ji} = a'_{ij} + b'_{ij}.$$

So, 1) holds.

2) We have

$$(a'_{ij})' = (a_{ji})' = a_{ij}.$$

Thus also 2) holds.

3) We easily get

$$(\lambda a_{ij})' = \lambda a'_{ij}.$$

Hence point 3) holds too. $\square$

Definition. A square matrix A is called *symmetric* if $A^T = A$.

Remark. The symmetry in a symmetric matrix occurs about its main diagonal.

Examples.

1) Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 4 \\ 0 & 4 & 2 \end{bmatrix}$. Then $A^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 4 \\ 0 & 4 & 2 \end{bmatrix}$. So, $A^T = A$, that is, A is symmetric.

2) Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$. Then $A^T = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$. So, $A^T \neq A$, that is, A is not symmetric.

Definition. A square matrix A is called *skew-symmetric* if $A^T = -A$.

Remark. Note that in the above definition we have $a_{jj} = -a_{jj}$, which implies that elements in main diagonal of a skew-symmetric matrix are all zero.

Example. Let $A = \begin{bmatrix} 0 & 4 & -1 \\ -4 & 0 & 2 \\ 1 & -2 & 0 \end{bmatrix}$. Then $A^T = \begin{bmatrix} 0 & -4 & 1 \\ 4 & 0 & -2 \\ -1 & 2 & 0 \end{bmatrix}$. So, $A^T = -A$, that is, A is skew-symmetric.

Remark. Notice that it possible that a square matrix is neither a symmetric nor skew-symmetric matrix.

Theorem. Let A be a square matrix. Then $A + A^T$ is symmetric and $A - A^T$ is skew-symmetric.

Proof. We have

$$\left(A + A^T\right)^T = A^T + \left(A^T\right)^T = A^T + A = A + A^T,$$

that is, $A + A^T$ is symmetric, and

$$\left(A - A^T\right)^T = A^T + \left(-A^T\right)^T = A^T - \left(A^T\right)^T = A^T - A = -\left(A - A^T\right),$$

that is, $A - A^T$ is skew-symmetric. $\square$

Theorem. Any square matrix can be written as a sum of a symmetric and skew-symmetric matrix.

Proof. We have
$$A=\left(\frac{1}{2}+\frac{1}{2}\right)A=\frac{1}{2}A+\frac{1}{2}A=\frac{1}{2}A+\frac{1}{2}A^T-\frac{1}{2}A^T+\frac{1}{2}A=\frac{1}{2}\left(A+A^T\right)+\frac{1}{2}\left(A-A^T\right).$$
By previous Theorem we have that $\frac{1}{2}\left(A+A^T\right)$ is symmetric and $\frac{1}{2}\left(A-A^T\right)$ is skew symmetric. Thus the proof is finished. □

Theorem. If $A\cdot B$ is defined, then
$$(A\cdot B)^T=B^T\cdot A^T.$$

Proof. Let $A=[a_{ij}]_{m\times n}$, $B=[b_{ij}]_{n\times p}$ and $A\cdot B=C$. Then $C=[c_{ij}]_{m\times p}$ and
$$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj},$$
where $i=1,2,\ldots,m$ and $j=1,2,\ldots,p$. Let $A^T=[a'_{ij}]$, $B^T=[b'_{ij}]$ and $C^T=[c'_{ij}]$. Then
$$c'_{ij}=c_{ji}=\sum_{k=1}^{n}a_{jk}b_{ki}=\sum_{k=1}^{n}a'_{kj}b'_{ik}=\sum_{k=1}^{n}b'_{ik}a'_{kj}.$$
Hence we have shown that the ijth element of C^T is equal to the ijth element of $B^T\cdot A^T$.
□

Remark. If the product $A\cdot B\cdot C$ is defined, then
$$(A\cdot B\cdot C)^T=C^T\cdot B^T\cdot A^T.$$
The same result holds for any finite number of factors, that is,
$$(A_1\cdot A_2\cdots A_{n-1}\cdot A_n)^T=A_n^T\cdot A_{n-1}^T\cdots A_2^T\cdot A_1^T.$$

Theorem. Let A be a symmetric matrix of order n and B be an $n\times m$ matrix. Then $B^T\cdot A\cdot B$ is a symmetric matrix of order m.

Proof. Since A is symmetric, it follows $A^T=A$. We have
$$\left(B^T\cdot A\cdot B\right)^T=B^T\cdot A^T\cdot\left(B^T\right)^T=B^T\cdot A\cdot B.\ \square$$

Definition. (**Trace of a square matrix**) Let $A=[a_{ij}]\in M_{n\times n}(\mathbb{F})$. The *trace* of a matrix A is a number
$$\operatorname{tr}(A)\underset{df}{=}\sum_{i=1}^{n}a_{ii}.$$

Examples.

1) Let $A=\begin{bmatrix}1&4\\3&-3\end{bmatrix}$. Then $\operatorname{tr}(A)=1-3=-2$.

2) Let $A = \begin{bmatrix} 2 & 4 & -1 \\ 3 & -2 & 6 \\ 0 & 3 & 5 \end{bmatrix}$. Then $\mathrm{tr}(A) = 2 - 2 + 5 = 5$.

Theorem. Let $A, B \in M_{n\times n}(\mathbb{F})$. Then

1) $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$,
2) $\mathrm{tr}(A) = \mathrm{tr}(A^T)$,
3) $\mathrm{tr}(A \cdot B) = \mathrm{tr}(B \cdot A)$,
4) $\mathrm{tr}(\lambda A) = \lambda \mathrm{tr}(A)$, where $\lambda \in \mathbb{F}$.

Proof. Let $A = [a_{ij}]$ and $B = [b_{ij}]$.

1) We have

$$\mathrm{tr}(A + B) = \sum_{i=1}^{n} (a_{ii} + b_{ii}) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \mathrm{tr}(A) + \mathrm{tr}(B).$$

2) We easily get

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii} = \mathrm{tr}(A^T).$$

3) Let $A \cdot B = [c_{ij}]$ and $B \cdot A = [c'_{ij}]$. By definition of matrix multiplication we have

$$\begin{aligned} \mathrm{tr}(A \cdot B) &= \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik} b_{ki} \right) = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} b_{ki} a_{ik} \right) \\ &= \sum_{k=1}^{n} \left(\sum_{i=1}^{n} b_{ki} a_{ik} \right) = \sum_{k=1}^{n} c'_{kk} = \mathrm{tr}(B \cdot A). \end{aligned}$$

4) Again we easily get

$$\mathrm{tr}(\lambda A) = \sum_{i=1}^{n} \lambda a_{ii} = \lambda \sum_{i=1}^{n} a_{ii} = \lambda \mathrm{tr}(A). \quad \square$$

Definition. (**Determinant**) Let $A \in M_{n\times n}(\mathbb{F})$. The *determinant* of a matrix A, denoted by $\det(A)$ or $|A|$, is a single number associated with A in the following way:

1) if $n = 1$, then $A = [a_{11}]$ and $\det(A) = a_{11}$,

2) if $n = 2$, then $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

3) if $n > 1$, let M_{ij} denotes the determinant of the $(n-1)\times(n-1)$ submatrix of A obtained by deleting its ith row and jth column. The determinant M_{ij} is called the *minor* of an element a_{ij}. Then

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

where j can be a number of any column of A. The product $(-1)^{i+j}M_{ij}$ is called the *cofactor* of an element a_{ij}. The determinant of A can be also written as

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

where i can be a number of any row of A.

We try to illustrate the application of the first formula from 3) for $n = 3$, so for a matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Let $j = 1$ (column 1). Then

$$M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{23}a_{32},$$

$$M_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} = a_{12}a_{33} - a_{13}a_{32},$$

$$M_{31} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = a_{12}a_{23} - a_{13}a_{22}$$

and

$$\begin{aligned}\det(A) &= \sum_{i=1}^{3}(-1)^{i+1}a_{i1}M_{i1} \\ &= (-1)^{1+1}a_{11}M_{11} + (-1)^{2+1}a_{21}M_{21} + (-1)^{3+1}a_{31}M_{31} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}).\end{aligned}$$

Hence,

$$\det(A) = a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

The above expression need not be memorized. It can be obtained as follows:

$$\det(A) = \left|\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}\right.$$
$$\ominus \quad \ominus \quad \ominus \qquad \oplus \quad \oplus \quad \oplus$$

Example. We can calculate:

$$\left|\begin{array}{ccc|cc} 1 & 2 & 1 & 1 & 2 \\ 0 & 1 & 3 & 0 & 1 \\ 0 & 1 & 2 & 0 & 1 \end{array}\right. = 1\cdot 1\cdot 2 + 2\cdot 3\cdot 0 + 1\cdot 0\cdot 1 - 1\cdot 1\cdot 0 - 1\cdot 3\cdot 1 - 2\cdot 0\cdot 2 = 2 - 3 = -1.$$

Example. Using second formula from 3), we get (for $i = 1$):

$$\begin{vmatrix} 1 & 2 & 3 & 0 \\ 4 & 2 & 1 & 0 \\ 1 & 2 & 0 & -1 \\ 2 & 0 & -1 & -1 \end{vmatrix} = (-1)^{1+1}\cdot 1\cdot \begin{vmatrix} 2 & 1 & 0 \\ 2 & 0 & -1 \\ 0 & -1 & -1 \end{vmatrix} + (-1)^{1+2}\cdot 2\cdot \begin{vmatrix} 4 & 1 & 0 \\ 1 & 0 & -1 \\ 2 & -1 & -1 \end{vmatrix}$$
$$+ (-1)^{1+3}\cdot 3\cdot \begin{vmatrix} 4 & 2 & 0 \\ 1 & 2 & -1 \\ 2 & 0 & -1 \end{vmatrix} = 1\cdot(2-2) - 2\cdot(-2+1-4)$$
$$+ 3\cdot(-8-4+2) = -20.$$

The evaluation of determinants of high order is a lengthy process. Now we establish a number of important theorems leading us to rapid methods of evaluating determinants.

Theorem. (**Transposition**) Let A be a square matrix. Then

$$\det(A^T) = \det(A).$$

Proof. Proof is by induction. Let A be a 2×2 matrix. Then

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

and

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} = \det(A^T).$$

Thus the result holds for 2×2 matrices. Assume now that the result holds for $(n-1)\times(n-1)$ matrices. Consider an $n\times n$ matrix

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

Let A_{ij} denotes a minor of an element a_{ij} of A. By B_{ij} we denote a minor of an element a'_{ij} of A^T. Expand $\det(A^T)$ in terms of an nth row and $\det(A)$ in terms of an nth column to get

$$\begin{aligned}\det(A^T) &= \begin{vmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} \\ &= \sum_{i=1}^{n} (-1)^{n+i} a_{in} B_{in} \\ &= \sum_{i=1}^{n} (-1)^{i+n} a_{in} A_{in} \qquad \text{(since } B_{in} = A_{in}\text{)} \\ &= \det(A).\end{aligned}$$

By induction, theorem holds for all square matrices. □

Theorem. (**Interchange of rows or columns**) Let A, B be square matrices. If B is obtained from A by interchanging two rows or two columns, then

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

Proof. Proof is by induction. Indeed, theorem holds for determinants of square matrices of order 2. Assume that it holds for determinants of square matrices of order $n-1$. Let A be a square matrix of order n, and B be obtained from A by interchanging two rows. Expand $\det(A)$ and $\det(B)$ in terms of a row which is not one of those interchanged, call it ith row. Then

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

and

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

where minor B_{ij} is obtained from minor A_{ij} of a_{ij} in A by interchanging two rows. Notice that minors A_{ij} and B_{ij} are determinants of square matrices of order $n-1$, so by assumption, we have $B_{ij} = -A_{ij}$. It gives $\det(B) = -\det(A)$. By induction, theorem holds for all square matrices. Proof for columns is similar. □

Theorem. (**Proportional rows or columns**) Let A be a square matrix. If two rows or two columns of A are proportional, then

$$\det(A) = 0.$$

Proof. Interchange two proportional rows and use previous theorem to obtain

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

So, $2\det(A) = 0$, that is, $\det(A) = 0$. $\square$

Remark. Observe that a square matrix with two identical rows or two identical colums has a zero determinant.

Example. Let

$$A = \begin{bmatrix} 1 & 2 & 4 & -1 \\ 0 & 1 & 3 & 0 \\ 2 & 4 & 8 & -2 \\ 0 & 5 & 1 & 7 \end{bmatrix}.$$

Then, $\det(A) = 0$, since two rows (first and third) are proportional.

Theorem. (**Multiplication by a constant**) Let A be a square matrix. If all elements of one row or one column of A are multiplied by a constant c, then a determinant of A is multiplied by c.

Proof. Expand a determinant by that row or column whose elements are multiplied by c. $\square$

From previous theorem with $c = 0$ we get the following theorem.

Theorem. Let A be a square matrix. If all elements of one row or one column of A are zero, then a determinant of A is zero.

Example. By above theorem we have

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

Theorem. (**Addition of a row or column**) The value of a determinant is unchanged if a scalar multiple of one row (or column) is added to another row (or column).

Proof. We prove the result for rows, proof for columns is similar. Assume that $A = [a_{ij}]_{n\times n}$ and $B = [b_{ij}]_{n\times n}$. Let B be obtained from A be adding a multiple of an sth row to an ith row. Then $b_{ij} = a_{ij} + ca_{sj}$. Notice that all elements in rows other than an ith row of B are identical to corresponding elements in A. Now expand $\det(B)$ in terms of an ith row. Let B_{ij} be a minor of an element b_{ij} of B. By A_{ij} we denote a minor of an element a_{ij} of A. Then we have

$$\begin{aligned}\det(B) &= \sum_{j=1}^{n}(-1)^{i+j}b_{ij}B_{ij} = \sum_{j=1}^{n}(-1)^{i+j}(a_{ij}+ca_{sj})A_{ij}\\ &= \sum_{j=1}^{n}(-1)^{i+j}a_{ij}A_{ij} + c\sum_{j=1}^{n}(-1)^{i+j}a_{sj}A_{ij}\\ &= \det(A) + c\det(\overline{A}),\end{aligned}$$

where $\overline{A}$ is a matrix obtained from A by replacing an ith row by an sth row. So, $\det(\overline{A}) = 0$, because an ith row and an sth row of $\overline{A}$ are identical. Thus,

$$\det(B) = \det(A). \quad \square$$

Example. By above theorem we calculate

$$\begin{vmatrix} 1 & 4 & -3 & 1 & 5\\ 0 & 1 & 3 & 1 & 2\\ -1 & 0 & 4 & 2 & 0\\ 2 & 1 & 1 & 0 & 1\\ 0 & 0 & 1 & -1 & 1 \end{vmatrix} \underset{r_4-2r_2}{\overset{r_3+r_1}{=}} \begin{vmatrix} 1 & 4 & -3 & 1 & 5\\ 0 & 1 & 3 & 1 & 2\\ 0 & 4 & 1 & 3 & 5\\ 0 & -7 & 7 & -2 & -9\\ 0 & 0 & 1 & -1 & 1 \end{vmatrix} = (-1)^{1+1}\cdot 1\cdot \begin{vmatrix} 1 & 3 & 1 & 2\\ 4 & 1 & 3 & 5\\ -7 & 7 & -2 & -9\\ 0 & 1 & -1 & 1 \end{vmatrix} \underset{c_4-c_2}{\overset{c_3+c_2}{=}}$$

$$\begin{vmatrix} 1 & 3 & 4 & -1\\ 4 & 1 & 4 & 4\\ -7 & 7 & 5 & -16\\ 0 & 1 & 0 & 0 \end{vmatrix} = (-1)^{4+2}\cdot 1\cdot \begin{vmatrix} 1 & 4 & -1\\ 4 & 4 & 4\\ -7 & 5 & -16 \end{vmatrix} \underset{c_3+c_1}{\overset{c_2-4c_1}{=}} \begin{vmatrix} 1 & 0 & 0\\ 4 & -12 & 8\\ -7 & 33 & -23 \end{vmatrix} =$$

$$(-1)^{1+1}\cdot 1\cdot \begin{vmatrix} -12 & 8\\ 33 & -23 \end{vmatrix} = 276 - 264 = 12.$$

Theorem. Let A be a triangular matrix. Then a determinant of A equals product of diagonal elements.

Proof. Let A be an upper triangular matrix of order n, that is,

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

We have

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}$$

$$= a_{11}a_{22} \cdot \begin{vmatrix} a_{33} & \dots & a_{3n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{vmatrix} = \ldots = a_{11}a_{22} \cdots a_{nn}.$$

It means that theorem holds. Proof for a lower triangular matrix is similar. □

Remark. From above theorem we get

$$\det(I) = 1.$$

Theorem. If each element of a row or a column of a determinant is expressed as a binomial, then that determinant can be written as a sum of two determinants.

Proof. Expand a determinant by a row or a column whose elements are binomials. □

Theorem. (**Cauchy**) Let $A, B \in M_{n\times n}(\mathbb{F})$. Then

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

(without proof)

Remark. Let $A, B \in M_{n\times n}(\mathbb{F})$. Then

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

Remark. Let A be a square matrix. Then

$$\det(A^k) = (\det(A))^k.$$

Remark. Let $A \in M_{n\times n}(\mathbb{F})$ and $\lambda \in \mathbb{F}$. Then

$$\det(\lambda A) = \lambda^n \det(A).$$

Remark. Observe that the equality

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

does not have to be true for matrices $A, B \in M_{n\times n}(\mathbb{F})$.

Example. Let

$$A = \begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}.$$

We have $\det(A) = 5$, $\det(B) = 1$ and

$$A + B = \begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}.$$

Then

$$\det(A + B) = 9 \neq \det(A) + \det(B) = 6.$$

Definition. (**Inverse of a matrix**) Let A be a square matrix. The *inverse* of a matrix A is a matrix A^{-1} such that

$$A \cdot A^{-1} = A^{-1} \cdot A = I.$$

Definition. (**Invertible matrix**) A square matrix A is called *invertible* if an inverse A^{-1} of A exists.

Example. Matrix $\begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{1}{5} & -\frac{2}{5} \end{bmatrix}$ is an inverse of matrix $\begin{bmatrix} 2 & -3 \\ 1 & -4 \end{bmatrix}$, because

$$\begin{bmatrix} 2 & -3 \\ 1 & -4 \end{bmatrix} \cdot \begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{1}{5} & -\frac{2}{5} \end{bmatrix} = \begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{1}{5} & -\frac{2}{5} \end{bmatrix} \cdot \begin{bmatrix} 2 & -3 \\ 1 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Remark. The unit matrix I is an inverse of itself, because $I \cdot I = I$. That is, $I^{-1} = I$.

Theorem. An inverse (if it exists) of a square matrix A is unique.

Proof. Assume that B and C are inverses of A, so $A \cdot B = I$ and $C \cdot A = I$. Then we have

$$B = I \cdot B = (C \cdot A) \cdot B = C \cdot (A \cdot B) = C \cdot I = C.$$

It proves that an inverse of a matrix A is uniquely determined. $\square$

Remark. Notice that the zero matrix $0 = 0_{n \times n}$ has no inverse, because for any square matrix A of order n, we have $A \cdot 0 = 0 \cdot A = 0$.

Now we will give a formula of an inverse of a matrix. Consider a square matrix A of order n:

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

Let M_{ij} be a minor an element a_{ij}, that is, a determinant of an $(n-1)\times(n-1)$ submatrix of A obtained by deleting its ith row and jth column. Then a cofactor A_{ij} of a_{ij} is given by

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

Construct a transpose A^T of A. The *adjoint* of A, written $adj(A)$, is a matrix obtained by replacing any element of A^T by A_{ij}, so

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

Theorem. For any square matrix A the formula

$$A \cdot adj(A) = adj(A) \cdot A = \det(A) \cdot I$$

holds. If $\det(A) \neq 0$, then

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

Proof. Let $C = A \cdot adj(A)$ and $C = [c_{ij}]$. Then

$$c_{ij} = \sum_{k=1}^{n} a_{ik} A_{kj} = \sum_{k=1}^{n} (-1)^{k+j} a_{ik} M_{kj} = \begin{cases} \det(A) & \text{if } \; i = j, \\ 0 & \text{if } \; i \neq j. \end{cases}$$

Explanation: (1) if $i = j$, then M_{ki} is a minor of a_{ki} of A^T, so of a_{ik} of A, that is, $c_{ij} = \det(A)$ in that case; (2) if $i \neq j$, then $a_{ik} = a_{ki}$ and a jth column is replaced by an ith column, so there are two identical columns in a determinant, that is, $c_{ij} = 0$.

Thus the first formula is satisfied. The rest of this theorem is evident. □

Remark. Above theorem gives us an important method of obtaining an inverse of a matrix.

Example. We will find an inverse of a matrix

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 1 & -1 & 1 \\ -1 & -1 & 2 \end{bmatrix}.$$

First we have $\det(A) = -1 \neq 0$, so an inverse of A exists. Next

$$A^T = \begin{bmatrix} 1 & 1 & -1 \\ -2 & -1 & -1 \\ 3 & 1 & 2 \end{bmatrix} \quad \text{and} \quad adj(A) = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix},$$

where $A_{ij} = (-1)^{i+j} M_{ij}$ for $i, j = 1, 2, 3$. Thus we have

$$A_{11} = (-1)^{1+1} \begin{vmatrix} -1 & -1 \\ 1 & 2 \end{vmatrix} = -1, \quad A_{12} = (-1)^{1+2} \begin{vmatrix} -2 & -1 \\ 3 & 2 \end{vmatrix} = 1, \quad A_{13} = (-1)^{1+3} \begin{vmatrix} -2 & -1 \\ 3 & 1 \end{vmatrix} = 1,$$

$$A_{21} = (-1)^{2+1} \begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} = -3, \quad A_{22} = (-1)^{2+2} \begin{vmatrix} 1 & -1 \\ 3 & 2 \end{vmatrix} = 5, \quad A_{23} = (-1)^{2+3} \begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} = 2,$$

$$A_{31} = (-1)^{3+1} \begin{vmatrix} 1 & -1 \\ -1 & -1 \end{vmatrix} = -2, \quad A_{32} = (-1)^{3+2} \begin{vmatrix} 1 & -1 \\ -2 & -1 \end{vmatrix} = 3, \quad A_{33} = (-1)^{3+3} \begin{vmatrix} 1 & 1 \\ -2 & -1 \end{vmatrix} = 1$$

and

$$adj(A) = \begin{bmatrix} -1 & 1 & 1 \\ -3 & 5 & 2 \\ -2 & 3 & 1 \end{bmatrix}.$$

Finally,

$$A^{-1} = \frac{1}{\det(A)} adj(A) = \frac{1}{-1} \begin{bmatrix} -1 & 1 & 1 \\ -3 & 5 & 2 \\ -2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ 3 & -5 & -2 \\ 2 & -3 & -1 \end{bmatrix}.$$

The above calculations can be verified by computing

$$A \cdot A^{-1} = \begin{bmatrix} 1 & -2 & 3 \\ 1 & -1 & 1 \\ -1 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & -1 \\ 3 & -5 & -2 \\ 2 & -3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

and

$$A^{-1} \cdot A = \begin{bmatrix} 1 & -1 & -1 \\ 3 & -5 & -2 \\ 2 & -3 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -2 & 3 \\ 1 & -1 & 1 \\ -1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I.$$

Definition. A square matrix A is called

1) *singular* if $\det(A) = 0$,
2) *nonsingular* if $\det(A) \neq 0$.

Theorem. (**Existence of an inverse**) An inverse of a square matrix A exists if and only if A is nonsingular.

Proof. First assume that A^{-1} exists. Then $A \cdot A^{-1} = I$ and $\det(A \cdot A^{-1}) = 1$, so

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

Hence $\det(A) \neq 0$, that is, A is nonsingular.

Conversely, let A be nonsingular. Hence $\det(A) \neq 0$, and by previous Theorem we have

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

that is, A^{-1} exists. $\square$

Example. We will find all values of k for which matrix

$$A = \begin{bmatrix} 1 & -1 & -1 & 0 \\ 0 & k & 1 & 2 \\ 0 & 0 & 1 & k \\ 0 & 1 & 1 & 2 \end{bmatrix}$$

has an inverse. We know that A^{-1} exists iff $\det(A) \neq 0$. We have

$$\det(A) = \begin{vmatrix} 1 & -1 & -1 & 0 \\ 0 & k & 1 & 2 \\ 0 & 0 & 1 & k \\ 0 & 1 & 1 & 2 \end{vmatrix} = (-1)^2 \cdot 1 \cdot \begin{vmatrix} k & 1 & 2 \\ 0 & 1 & k \\ 1 & 1 & 2 \end{vmatrix} = 2k + k - 2 - k^2 = -k^2 + 3k - 2$$

and $-k^2 + 3k - 2 = 0$ for $k \in \{1, 2\}$. Hence matrix A has an inverse for $k \neq 1$ and $k \neq 2$.

Theorem. Let A, B be nonsingular matrices of the same order. Then

1) $\det(A^{-1}) = \frac{1}{\det(A)}$,
2) $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$,
3) $(A^{-1})^{-1} = A$,
4) $\left(A^T\right)^{-1} = (A^{-1})^T$,
5) $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$, where $\lambda \in \mathbb{F} \backslash \{0\}$.

Proof. Let us take two nonsingular matrices A, B of the same order.

1) Since $A^{-1} \cdot A = I$, it follows that $\det(A^{-1} \cdot A) = 1$. Hence, $\det(A^{-1}) \cdot \det(A) = 1$, that is,

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

2) We easily get

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

3) By point 2) we have

$$\left(A^{-1}\right)^{-1} = \left(A^{-1}\right)^{-1} \cdot A^{-1} \cdot A = \left(A \cdot A^{-1}\right)^{-1} \cdot A = I^{-1} \cdot A = I \cdot A = A.$$

4) Again by point 2) we get

$$\begin{aligned}\left(A^T\right)^{-1} &= \left(A^T\right)^{-1} \cdot \left(\left(A^{-1}\right)^T\right)^{-1} \cdot \left(A^{-1}\right)^T = \left(\left(A^{-1}\right)^T \cdot A^T\right)^{-1} \cdot \left(A^{-1}\right)^T \\ &= \left(\left(A \cdot A^{-1}\right)^T\right)^{-1} \cdot \left(A^{-1}\right)^T = \left(I^T\right)^{-1} \cdot \left(A^{-1}\right)^T = I \cdot \left(A^{-1}\right)^T \\ &= \left(A^{-1}\right)^T.\end{aligned}$$

5) We have

$$\begin{aligned}(\lambda A)^{-1} &= (\lambda A)^{-1} \cdot \left(\frac{1}{\lambda}A^{-1}\right)^{-1} \cdot \frac{1}{\lambda}A^{-1} = \left(\frac{1}{\lambda}A^{-1} \cdot \lambda A\right)^{-1} \cdot \frac{1}{\lambda}A^{-1} \\ &= \left(A^{-1} \cdot A\right)^{-1} \cdot \frac{1}{\lambda}A^{-1} = I^{-1} \cdot \frac{1}{\lambda}A^{-1} = I \cdot \frac{1}{\lambda}A^{-1} \\ &= \frac{1}{\lambda}A^{-1}. \ \square\end{aligned}$$

Example. Let A, B be two 4×4 nonsingular matrices. We will compute $\det(B^{-1})$, if $B = 2A$ and $AB = -I$. We have

$$\begin{aligned}\det(B) &= \det(2A) \\ \det(B) &= 2^4 \cdot \det(A) \\ \det(B) &= 16 \cdot \det(A) \\ \det(A) &= \frac{1}{16}\det(B).\end{aligned}$$

and

$$\begin{aligned}\det(AB) &= \det(-I) \\ \det(A)\det(B) &= (-1)^4\det(I) \\ \det(A)\det(B) &= 1 \\ \frac{1}{16}\left(\det(B)\right)^2 &= 1 \\ \left(\det(B)\right)^2 &= 16, \\ \det(B) = 4 \ &\vee \ \det(B) = -4.\end{aligned}$$

Finally,

$$\det(B^{-1}) = \frac{1}{4} \ \vee \ \det(B^{-1}) = -\frac{1}{4}.$$

3. Systems of linear equations

Definition. A system of m linear equations in n unknowns $x_1, x_2, \dots, x_n$ is a set of equations of the form:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

where $a_{ij} \in \mathbb{R}$ are called the *coefficients* of a system, and $b_i \in \mathbb{R}$.

Remark. A system of linear equations can be written in the form:

$$AX = B,$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

A matrix A is called the *coefficient matrix* of a system.

Definition. A system of linear equations $AX = B$ is called

1) *homogeneous* if $B = 0$,
2) *nonhomogeneous* if $B \neq 0$.

Definition. A system of linear equations $AX = B$ is called the *Cramer's system* if A is a square nonsingular matrix.

Theorem. Let A be a square matrix of order n. A Cramer's system $AX = B$ has precisely one solution. This solution is given by formulas:

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \dots, x_n = \frac{D_n}{D},$$

where $D = \det(A)$, and D_k, for $k = 1, 2, \dots, n$, is a determinant obtained from D by replacing in D a kth column by column B.

Proof. First, we have

$$AX = B, \quad A^{-1}AX = A^{-1}B, \quad X = A^{-1}B.$$

Now, let $A^{-1} = C = [c_{ij}]$. From definition of a matrix multiplication and by the formula of an inverse matrix we get

$$x_k = \sum_{i=1}^{n} c_{ki} b_i = \sum_{i=1}^{n} \frac{A_{ik}}{\det(A)} b_i = \frac{1}{\det(A)} \sum_{i=1}^{n} A_{ik} b_i = \frac{D_k}{D}$$

for $k = 1, 2, \ldots, n$. $\square$

Remark. If a Cramer's system is homogeneous, then $D_1 = 0, D_2 = 0, \ldots, D_n = 0$, that is, it has only the zero solution $x_1 = 0, x_2 = 0, \ldots, x_n = 0$.

Example. Using Cramer's formulas we solve the following system

$$\begin{cases} x_1 + x_2 + x_3 = 0, \\ 2x_1 + 5x_2 + 3x_3 = 1, \\ -x_1 + 2x_2 + x_3 = 2. \end{cases}$$

We have

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & 3 \\ -1 & 2 & 1 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

Now,

$$D = \det(A) = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 5 & 3 \\ -1 & 2 & 1 \end{vmatrix} = 3$$

and

$$D_1 = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 5 & 3 \\ 2 & 2 & 1 \end{vmatrix} = -3, \quad D_2 = \begin{vmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ -1 & 2 & 1 \end{vmatrix} = 0, \quad D_3 = \begin{vmatrix} 1 & 1 & 0 \\ 2 & 5 & 1 \\ -1 & 2 & 2 \end{vmatrix} = 3.$$

So,

$$x_1 = \frac{D_1}{D} = -1, \quad x_2 = \frac{D_2}{D} = 0, \quad x_3 = \frac{D_3}{D} = 1.$$

Theorem. (**Method of matrix inversion**) A solution of a Cramer's system $AX = B$ is given by a formula

$$X = A^{-1} B.$$

Proof. We easily have

$$AX = B, \quad A^{-1}AX = A^{-1}B, \quad X = A^{-1}B. \ \square$$

Example. Using method of matrix inversion we solve the following system

$$\begin{cases} 2x_1 - x_2 + x_3 = 3, \\ x_1 + 2x_2 - x_3 = -1, \\ 3x_1 + x_2 + 2x_3 = 2. \end{cases}$$

Now,

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & -1 \\ 3 & 1 & 2 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad B = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}.$$

We have $\det(A) = 10$ and

$$A^{-1} = \frac{1}{10} \begin{bmatrix} 5 & 3 & -1 \\ -5 & 1 & 3 \\ -5 & -5 & 5 \end{bmatrix}.$$

Thus,

$$X = A^{-1}B = \frac{1}{10} \begin{bmatrix} 5 & 3 & -1 \\ -5 & 1 & 3 \\ -5 & -5 & 5 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Method of Gaussian elimination

We explain this method on examples.

Example. Using method of Gaussian elimination we solve the following system

$$\begin{cases} x_1 + x_2 + x_3 = 0, \\ 2x_1 + 5x_2 + 3x_3 = 1, \\ -x_1 + 2x_2 + x_3 = 2. \end{cases}$$

We have

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & 3 \\ -1 & 2 & 1 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

The whole elimination procedure can be systematized by operating directly on the following matrix

$$[A|B] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 5 & 3 & 1 \\ -1 & 2 & 1 & 2 \end{array} \right].$$

This matrix is called the *augmented matrix* of the given system. Instead of performing transformations on the system of linear equations, we can perform equivalent transformations on augmented matrix. These transformations are of three types:

1) interchanging two rows of the matrix ($r_i \leftrightarrow r_j$),
2) multiplying a row by a nonzero constant ($c \cdot r_i,\ c \neq 0$),
3) adding a multiple of one row of the matrix to another row ($r_i + c \cdot r_j$).

We say that two matrices are equivalent if one is obtained from the other using the above transformations.

We have

$$[A|B] = \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 2 & 5 & 3 & 1 \\ -1 & 2 & 1 & 2 \end{array}\right] \underset{r_3+r_1}{\xrightarrow{r_2-2r_1}} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & 3 & 1 & 1 \\ 0 & 3 & 2 & 2 \end{array}\right] \xrightarrow{r_3-r_2} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & 3 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\underset{r_2-r_3}{\xrightarrow{r_1-r_3}} \left[\begin{array}{rrr|r} 1 & 1 & 0 & -1 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right] \underset{\frac{1}{3}r_2}{\xrightarrow{r_1-\frac{1}{3}r_2}} \left[\begin{array}{rrr|r} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

Thus,

$$\begin{cases} x_1 = -1, \\ x_2 = 0, \\ x_3 = 1. \end{cases}$$

Definition. (**Rank of a matrix**) Let $A \in M_{m\times n}(\mathbb{R})$. The *rank* of a matrix, written $r(A)$, is the largest integer r for which a nonsingular $r \times r$ submatrix of A exists. A zero matrix is said to have rank 0.

Remark. Observe that $r(A) \leq \min(m, n)$. Moreover, the rank of an nth order identity matrix is n.

Example. Consider the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. We see that matrix A is nonsingular. It means that $r(A) = 2$.

Remark. We have some simple operations on rows and columns of a matrix which do not change its rank. These operations are: interchange two rows of the matrix ($r_i \leftrightarrow r_j$), multiplication of a row by a nonzero constant ($c \cdot r_i,\ c \neq 0$), addition to an ith row a jth row multiplied by a constant ($r_i + c \cdot r_j$). The same types of operations can be used on columns. By applying such operations it is possible to convert a matrix into one whose rank can be read off by looking at the matrix.

Example. We determine a rank of the following matrix

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 1 & 0 \\ -2 & -1 & 3 \\ -1 & 4 & -2 \end{bmatrix}.$$

We see that A is the 4×3 matrix, so $r(A) \leq 3$. We have

$$r(A) = r\left(\begin{bmatrix} 1 & 2 & -3 \\ 2 & 1 & 0 \\ -2 & -1 & 3 \\ -1 & 4 & -2 \end{bmatrix}\right) \underset{\substack{r_3+r_2 \\ r_4+r_1}}{\overset{r_2-2r_1}{=}} r\left(\begin{bmatrix} 1 & 2 & -3 \\ 0 & -3 & 6 \\ 0 & 0 & 3 \\ 0 & 6 & -5 \end{bmatrix}\right) = 3,$$

since

$$\begin{vmatrix} 1 & 2 & -3 \\ 0 & -3 & 6 \\ 0 & 0 & 3 \end{vmatrix} = -9 \neq 0.$$

Theorem. Let $A \in M_{m\times n}(\mathbb{R})$. Then

$$r(A) = r(A^T).$$

Proof. Easy. $\square$

Theorem. Let A be a square matrix of order n. The following are equivalent:

1) A is invertible,
2) A is nonsingular,
3) $r(A) = n$,
4) a system $AX = B$ has a unique solution for any $n \times 1$ matrix B.

Proof. Easy. $\square$

Theorem. (**Kronecker-Capelli**) Let $A \in M_{m\times n}(\mathbb{R})$. A system $AX = B$ has a solution if and only if

$$r(A) = r(A|B).$$

Moreover, let $r(A) = r(A|B) = r$. Then

1) if $r = n$, then there is precisely one solution,
2) if $r < n$, then there are infinitely many solutions which depend on $n - r$ parameters.

(without proof)

Conclusion. A homogeneous system $AX = B$ has a solution because $r(A) = r(A|B)$.

Now we give several examples illustrating Kronecker-Capelli's Theorem.

Example. Take the following system

$$\begin{cases} x_1 + 2x_2 - 3x_3 = 2, \\ 5x_1 - x_2 + x_3 = 1. \end{cases}$$

We have

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 5 & -1 & 1 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

First, using Kronecker-Capelli's Theorem, let us find number of solutions of that system. To do it we determine ranks of A and $[A|B]$. We have

$$[A|B] = \left[\begin{array}{ccc|c} 1 & 2 & -3 & 2 \\ 5 & -1 & 1 & 1 \end{array}\right] \xrightarrow{r_2-5r_1} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 2 \\ 0 & -11 & 16 & -9 \end{array}\right].$$

So, $r(A) = r(A|B) = 2 < n = 3$. Thus the system has infinitely many solutions depended on $n - r = 3 - 2 = 1$ parameter. To solve that system let us transform the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 2 \\ 0 & -11 & 16 & -9 \end{array}\right] \xrightarrow{-\frac{1}{11}r_2} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 2 \\ 0 & 1 & -\frac{16}{11} & \frac{9}{11} \end{array}\right] \xrightarrow{r_1-2r_2} \left[\begin{array}{ccc|c} 1 & 0 & -\frac{1}{11} & \frac{4}{11} \\ 0 & 1 & -\frac{16}{11} & \frac{9}{11} \end{array}\right].$$

Hence the solution is

$$\begin{cases} x_1 = \frac{4}{11} + \frac{1}{11}x_3, \\ x_2 = \frac{9}{11} + \frac{16}{11}x_3, \end{cases} \quad \text{where} \quad x_3 \in \mathbb{R}.$$

Example. For the system

$$\begin{cases} 10x_1 - 2x_2 + 8x_3 = 1, \\ 5x_1 + x_2 - 8x_3 = 0, \\ 15x_1 - x_2 = 0 \end{cases}$$

we have

$$[A|B] = \left[\begin{array}{ccc|c} 10 & -2 & 8 & 1 \\ 5 & 1 & -8 & 0 \\ 15 & -1 & 0 & 0 \end{array}\right] \xrightarrow[\substack{r_3-3r_2 \\ r_1\leftrightarrow r_2}]{r_1-2r_2} \left[\begin{array}{ccc|c} 5 & 1 & -8 & 0 \\ 0 & -4 & 24 & 1 \\ 0 & -4 & 24 & 0 \end{array}\right] \xrightarrow{r_3-r_2} \left[\begin{array}{ccc|c} 5 & 1 & -8 & 0 \\ 0 & -4 & 24 & 1 \\ 0 & 0 & 0 & -1 \end{array}\right].$$

So, $r(A) = 2$, because

$$\begin{vmatrix} 5 & 1 & -8 \\ 0 & -4 & 24 \\ 0 & 0 & 0 \end{vmatrix} = 0 \text{ and } \begin{vmatrix} 5 & 1 \\ 0 & -4 \end{vmatrix} = -20 \neq 0$$

and $r(A|B) = 3$, because

$$\begin{vmatrix} 5 & 1 & 0 \\ 0 & -4 & 1 \\ 0 & 0 & -1 \end{vmatrix} = 20 \neq 0.$$

Hence, $r(A) \neq r(A|B)$. Thus the system has no solution.

Example. Now, we find values of k for which the system

$$\begin{cases} kx_1 + x_2 + x_3 = 1, \\ x_1 + kx_2 + x_3 = k, \\ x_1 + x_2 + kx_3 = k^2 \end{cases}$$

has a solution. We have

$$\left[\begin{array}{ccc|c} k & 1 & 1 & 1 \\ 1 & k & 1 & k \\ 1 & 1 & k & k^2 \end{array}\right] \xrightarrow{r_1 \leftrightarrow r_3} \left[\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 1 & k & 1 & k \\ k & 1 & 1 & 1 \end{array}\right] \xrightarrow[r_3 - kr_1]{r_2 - r_1} \left[\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 0 & k-1 & 1-k & k-k^2 \\ 0 & 1-k & 1-k^2 & 1-k^3 \end{array}\right]$$

$$\xrightarrow{r_3 + r_2} \left[\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 0 & k-1 & 1-k & k(1-k) \\ 0 & 0 & (1-k)(2+k) & (1-k)(1+k)^2 \end{array}\right].$$

Notice that for $k = 1$ we get

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

so in this case $r(A) = r(A|B) = 1 < n = 3$ and the system has infinitely many solutions with $n - r = 3 - 1 = 2$ parameters. Assume that $k \neq 1$. Then

$$\left[\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 0 & k-1 & 1-k & k(1-k) \\ 0 & 0 & (1-k)(2+k) & (1-k)(1+k)^2 \end{array}\right] \xrightarrow[\frac{1}{1-k} r_3]{\frac{1}{1-k} r_2} \left[\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 0 & -1 & 1 & k \\ 0 & 0 & 2+k & (1+k)^2 \end{array}\right].$$

We see that for $k = -2$ the augmented matrix takes the form

$$\left[\begin{array}{ccc|c} 1 & 1 & -2 & 4 \\ 0 & -1 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

So the system has no solution because $r(A) = 2 \neq r(A|B) = 3$. Now, consider the case when $k \neq 1$ and $k \neq -2$. Looking at the matrix

$$\left[\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 0 & -1 & 1 & k \\ 0 & 0 & 2+k & (1+k)^2 \end{array}\right]$$

we see that $r(A) = r(A|B) = 3 = n$, which shows that the system has a unique solution. Finally, the system has a solution for $k \neq -2$.

Example. Now, we discuss the solvability of the system

$$\begin{cases} 3x_1 - 2x_2 + x_3 = b, \\ 5x_1 - 8x_2 + 9x_3 = 3, \\ 2x_1 + x_2 + ax_3 = -1. \end{cases}$$

We get

$$\left[\begin{array}{ccc|c} 3 & -2 & 1 & b \\ 5 & -8 & 9 & 3 \\ 2 & 1 & a & -1 \end{array}\right] \xrightarrow[r_3-r_1]{r_2-2r_1} \left[\begin{array}{ccc|c} 3 & -2 & 1 & b \\ -1 & -4 & 7 & 3-2b \\ -1 & 3 & a-1 & -1-b \end{array}\right] \xrightarrow[\substack{r_1\leftrightarrow r_2 \\ r_2\leftrightarrow r_3}]{-1\cdot r_2} \left[\begin{array}{ccc|c} 1 & 4 & -7 & 2b-3 \\ -1 & 3 & a-1 & -1-b \\ 3 & -2 & 1 & b \end{array}\right]$$

$$\xrightarrow[r_3-3r_1]{r_2+r_1} \left[\begin{array}{ccc|c} 1 & 4 & -7 & 2b-3 \\ 0 & 7 & a-8 & b-4 \\ 0 & -14 & 22 & -5b+9 \end{array}\right] \xrightarrow{r_3+2r_2} \left[\begin{array}{ccc|c} 1 & 4 & -7 & 2b-3 \\ 0 & 7 & a-8 & b-4 \\ 0 & 0 & 2a+6 & -3b+1 \end{array}\right].$$

Discussion:

1) if $a \neq -3$, then $r(A) = r(A|B) = 3 = n$ and the system has a unique solution,

2) if $a = -3$ and $b = \frac{1}{3}$, then $r(A) = r(A|B) = 2 < 3 = n$ and the system has infinitely many solutions with 1 parameters,

3) if $a = -3$ and $b \neq \frac{1}{3}$, then $r(A) = 2 \neq r(A|B) = 3$ and the system has no solution.

4. Polynomials

Let $\mathbb{P}$ means the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$. We denote by $\mathbb{P}[x]$ the set of all polynomials of the form

$$p(x) = p_n x^n + p_{n-1}x^{n-1} + \ldots + p_1 x + p_0,$$

where $n \in \mathbb{N} \cup \{0\}$ and $p_0, p_1, \ldots, p_n \in \mathbb{P}$. So $\mathbb{P}[x]$ denotes the set of all polynomials in variable $x \in \mathbb{R}$ (or $\mathbb{C}$) and with coefficients in $\mathbb{P}$. If $\mathbb{P} = \mathbb{R}$, then p is called a *real polynomial*, and if $\mathbb{P} = \mathbb{C}$, then we call p a *complex polynomial*. If $p_n \neq 0$, then n is called the *degree* of a polynomial $p(x)$, written $n = \deg(p(x))$.

Remark. We define $\deg(0) = -\infty$, where 0 represents the *constant zero polynomial* $0(x) = 0$ for all x.

Definition. Let $p(x), q(x) \in \mathbb{P}[x]$. Operations of *polynomial addition* and *polynomial multiplication* are defined as follows

$$(p+q)(x) \underset{df}{=} p(x) + q(x),$$
$$(p \cdot q)(x) \underset{df}{=} p(x) \cdot q(x).$$

Polynomial equations:

1. An equation of the type

$$ax + b = 0,$$

where $a, b \in \mathbb{P}$ and $a \neq 0$, is called a *linear equation* and has unique solution

$$x = -\frac{b}{a}.$$

2. An equation of the type

$$ax^2 + bx + c = 0,$$

where $a, b, c \in \mathbb{P}$ and $a \neq 0$, is called a *quadratic equation*. To solve such equation we observe first of all that

$$ax^2 + bx + c = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right) = a\left(x^2 + 2\frac{b}{2a}x + \left(\frac{b}{2a}\right)^2 + \frac{c}{a} - \frac{b^2}{4a^2}\right)$$
$$= a\left(\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right) = 0$$

precisely when

$$(*) \qquad \left(x+\frac{b}{2a}\right)^2=\frac{b^2-4ac}{4a^2}.$$

Let $\mathbb{P}=\mathbb{R}$, so we consider quadratic equations with real coefficients. There are three cases.

(1) If $b^2-4ac>0$, then $(*)$ becomes

$$x+\frac{b}{2a}=\pm\frac{\sqrt{b^2-4ac}}{2a},$$

so that

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}.$$

We therefore have two distinct real solutions in this case.

(2) If $b^2-4ac=0$, then $(*)$ becomes

$$x+\frac{b}{2a}=0,$$

so that

$$x=-\frac{b}{2a}.$$

Now we have one solution which occurs twice.

(3) If $b^2-4ac<0$, then the right hand side of $(*)$ is negative. It follows that $(*)$ is never satisfied for a real number x, so that the quadratic equation has no real solution.

Let $\mathbb{P}=\mathbb{C}$. So we consider quadratic equations with complex coefficients:

$$az^2+bz+c=0,$$

where $a\neq 0$ and $\Delta=b^2-4ac=u+vi$ for some $u,v\in\mathbb{R}$. Again we have

$$\left(z+\frac{b}{2a}\right)^2=\frac{\Delta}{4a^2}.$$

Now there are two cases.

(1) If $\Delta=0$, then $\left(z+\frac{b}{2a}\right)^2=0$, so

$$z=-\frac{b}{2a}.$$

(2) If $\Delta\neq 0$, then we have

$$\left(z+\frac{b}{2a}\right)^2=\frac{\Delta}{4a^2},$$

so $(2az+b)^2 = \Delta$. Setting $t = 2az+b$, the equation is reduced to

$$t^2 = \Delta = u + vi.$$

Now, we solve this equation. Put $t = x + yi$, where $x, y \in \mathbb{R}$. Then

$$(x+yi)^2 = u + vi,$$

$$(x^2 - y^2) + 2xyi = u + vi,$$

$$\begin{cases} x^2 - y^2 = u, \\ 2xy = v. \end{cases}$$

Solving that system we get

$$\begin{cases} x = \sqrt{\frac{|\Delta|+u}{2}}, \\ y = \operatorname{sgn}(v) \cdot \sqrt{\frac{|\Delta|-u}{2}} \end{cases} \qquad \text{or} \qquad \begin{cases} x = -\sqrt{\frac{|\Delta|+u}{2}}, \\ y = -\operatorname{sgn}(v) \cdot \sqrt{\frac{|\Delta|-u}{2}}, \end{cases}$$

where

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if} \quad x > 0, \\ 0 & \text{if} \quad x = 0, \\ -1 & \text{if} \quad x < 0. \end{cases}$$

Thus,

$$t_1 = \sqrt{\frac{|\Delta|+u}{2}} + \operatorname{sgn}(v) \cdot \sqrt{\frac{|\Delta|-u}{2}} i \quad \text{and} \quad t_2 = -t_1.$$

Now, since $t = 2az + b$, we get

$$z = \frac{1}{2a}(-b + t),$$

that is, the quadratic equation has two solutions:

$$z_1 = \frac{1}{2a}(-b + t_1) \quad \text{and} \quad z_2 = \frac{1}{2a}(-b + t_2).$$

Example. We solve in $\mathbb{C}$ the following quadratic equation:

$$z^2 + 3z + 3 + i = 0.$$

For that equation: $a = 1$, $b = 3$, $c = 3+i$, $\Delta = b^2 - 4ac = 9 - 4(3+i) = -3 - 4i = u + vi$, whence $u = -3$, $v = -4$ and $|\Delta| = 5$. Thus,

$$\begin{aligned} t_1 &= \sqrt{\frac{|\Delta|+u}{2}} + \operatorname{sgn}(v) \cdot \sqrt{\frac{|\Delta|-u}{2}} i = \sqrt{\frac{5-3}{2}} - \sqrt{\frac{5+3}{2}} i = 1 - 2i, \\ t_2 &= -t_1 = -1 + 2i \end{aligned}$$

and

$$
\begin{aligned}
z_1 &= \frac{1}{2a}(-b+t_1) = \frac{1}{2}(-3+1-2i) = -1-i \quad \text{and} \\
z_2 &= \frac{1}{2a}(-b+t_2) = \frac{1}{2}(-3-1+2i) = -2+i.
\end{aligned}
$$

Remark. For polynomials equations of degree greater than 2 we do not have general formulas for solutions.

Theorem. Let $a(x), b(x) \in \mathbb{P}[x]$ and $a(x) \neq 0$. Then

$$b(x) = a(x)q(x) + r(x)$$

for unique $q(x), r(x) \in \mathbb{P}[x]$, where either $r(x) = 0$ or $\deg(r(x)) < \deg(a(x))$.

Proof. Consider all polynomials of the form

$$b(x) - a(x)Q(x),$$

where $Q(x) \in \mathbb{P}[x]$. If there exists $q(x) \in \mathbb{P}[x]$ such that $b(x) - a(x)q(x) = 0$, then our proof is complete. Suppose now that $b(x) - a(x)Q(x) \neq 0$ for any $Q(x) \in \mathbb{P}[x]$. Then among all polynomials of the form $b(x) - a(x)Q(x)$, where $Q(x) \in \mathbb{P}[x]$, there must be one with smallest degree. More precisely, there exists

$$m = \min\big\{\deg(b(x) - a(x)Q(x)) : Q(x) \in \mathbb{P}[x]\big\}.$$

Let $q(x) \in \mathbb{P}[x]$ satisfies $\deg(b(x) - a(x)q(x)) = m$ and let $r(x) = b(x) - a(x)q(x)$. Then $m = \deg(r(x)) < \deg(a(x)) = n$. Indeed, assume that $m \geq n$. Then writing $a(x) = a_n x^n + \ldots + a_1 x + a_0$ and $r(x) = r_m x^m + \ldots + r_1 x + r_0$ we have

$$
\begin{aligned}
r(x) - \big(r_m a_n^{-1} x^{m-n}\big)a(x) &= b(x) - a(x)q(x) - \big(r_m a_n^{-1} x^{m-n}\big)a(x) \\
&= b(x) - a(x)\big(q(x) + r_m a_n^{-1} x^{m-n}\big) \in \mathbb{P}[x]
\end{aligned}
$$

and $\deg\Big(r(x) - \big(r_m a_n^{-1} x^{m-n}\big)a(x)\Big) < \deg(r(x))$. This contradicts the minimality of m.

It suffices to prove that $q(x), r(x) \in \mathbb{P}[x]$ are unique. Suppose that $q_1(x), q_2(x) \in \mathbb{P}[x]$ satisfy $\deg(b(x) - a(x)q_1(x)) = m$ and $\deg(b(x) - a(x)q_2(x)) = m$. Let $r_1(x) = b(x) - a(x)q_1(x)$ and $r_2(x) = b(x) - a(x)q_2(x)$. Hence $r_1(x) - r_2(x) = a(x)(q_2(x) - q_1(x))$. If $q_1(x) \neq q_2(x)$, then

$$\deg(a(x)(q_2(x) - q_1(x))) \geq \deg(a(x)),$$

while

$$\deg(r_1(x) - r_2(x)) < \deg(a(x)),$$

and we get a contradiction. It follows that $q(x)$, and hence $r(x)$, are both unique. $\square$

Definition. A number $\alpha \in \mathbb{P}$ is called a *root* of a polynomial $p(x) \in \mathbb{P}[x]$ if

$$p(\alpha) = 0.$$

Theorem. Let $b(x) \in \mathbb{P}[x]$ and $\alpha \in \mathbb{P}$. Then there exists a unique polynomial $q(x) \in \mathbb{P}[x]$ such that

$$b(x) = (x - \alpha)q(x) + b(\alpha).$$

Proof. There exist unique polynomials $q(x), r(x) \in \mathbb{P}[x]$ such that

$$b(x) = (x - \alpha)q(x) + r(x),$$

where either $r(x) = 0$ or $\deg(r(x)) = 0$. Hence, $r(x)$ is a constant polynomial, so that

$$b(x) = (x - \alpha)q(x) + r,$$

where $r \in \mathbb{P}$. Substituting $x = \alpha$, we get $r = b(\alpha)$. $\square$

Definition. Let $a(x), b(x) \in \mathbb{P}[x]$. We say that $a(x)$ is a *factor* of $b(x)$, denoted by $a(x)|b(x)$, if there exists $c(x) \in \mathbb{P}[x]$ such that $b(x) = a(x)c(x)$.

Theorem. Let $b(x) \in \mathbb{P}[x]$. Then $\alpha \in \mathbb{P}$ is a root of $b(x)$ if and only if $x - \alpha$ is a factor of $b(x)$.

Proof. By previous Theorem there exists a unique polynomial $q(x) \in \mathbb{P}[x]$ such that

$$b(x) = (x - \alpha)q(x) + b(\alpha).$$

If $\alpha \in \mathbb{P}$ is a root of $b(x)$, then $b(\alpha) = 0$, so that $b(x) = (x - \alpha)q(x)$, whence $x - \alpha$ is a factor of $b(x)$.

Conversely, assume that $x - \alpha$ is a factor of $b(x)$. Then, by the above equality there must be $b(\alpha) = 0$. Hence α is a root of $b(x)$. $\square$

Theorem. (**Integer roots of a polynomial**) Let $p(x) = p_n x^n + p_{n-1}x^{n-1} + \ldots + p_1 x + p_0 \in \mathbb{Z}[x]$ and let $\alpha \in \mathbb{Z}\backslash\{0\}$ be a root of p. Then α is a divisor of p_0.

Proof. We have

$$p_n\alpha^n + p_{n-1}\alpha^{n-1} + \ldots + p_1\alpha + p_0 = 0,$$

that is,

$$p_0 = -\alpha\big(p_n\alpha^{n-1} + p_{n-1}\alpha^{n-2} + \ldots + p_1\big).$$

This means that α is a divisor of p_0. $\square$

Theorem. (**Rational roots of a polynomial**) Let $p(x) = p_n x^n + p_{n-1}x^{n-1} + \ldots + p_1 x + p_0 \in \mathbb{Z}[x]$, where $p_n \neq 0$, let $\frac{\alpha}{\beta}$ be a root of p, where $\alpha \in \mathbb{Z}$ and $\beta \in \mathbb{Z}\backslash\{0\}$, and let the greatest common divisor of α and β be 1. Then α is a divisor of p_0 and β is a divisor of p_n.

Proof. We have

$$p_n\left(\frac{\alpha}{\beta}\right)^n + p_{n-1}\left(\frac{\alpha}{\beta}\right)^{n-1} + \ldots + p_1\frac{\alpha}{\beta} + p_0 = 0.$$

Hence,

$$p_n\alpha^n + p_{n-1}\alpha^{n-1}\beta + \ldots + p_1\alpha\beta^{n-1} + p_0\beta^n = 0,$$

Let

$$m = p_n\alpha^{n-1} + p_{n-1}\alpha^{n-2}\beta + \ldots + p_1\beta^{n-1} \quad \text{and} \quad m' = p_{n-1}\alpha^{n-1} + p_{n-2}\alpha^{n-2}\beta + \ldots + p_0\beta^{n-1}.$$

Then $m, m' \in \mathbb{Z}$ and $\alpha m + p_0\beta^n = 0$ and $p_n\alpha^n + \beta m' = 0$. So,

$$\alpha m = -p_0\beta^n \quad \text{and} \quad \beta m' = -p_n\alpha^n.$$

Hence, α is a divisor of $p_0\beta^n$ and is not a divisor of β^n, that is, α is a divisor of p_0. Similarly, β is a divisor of $p_n\alpha^n$ and is not a divisor of α^n, that is, β is a divisor of p_n. □

Theorem. (**Fundamental theorem of algebra**) Every polynomial $p(z) \in \mathbb{C}[z]$ with $\deg(p(z)) = n$ has precisely n roots. If $p(z) = p_nz^n + \ldots + p_1z + p_0$, then there exist $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$, not necessarily distinct, such that

$$p(z) = p_n(z - \alpha_1)\cdots(z - \alpha_n).$$

(without proof)

Theorem. (**Viete's formulas**) Let $p(z) = p_nz^n + p_{n-1}z^{n-1} + \ldots + p_1z + p_0 \in \mathbb{C}[z]$ and $\deg(p(z)) = n$. Then numbers $z_1, z_2, \ldots, z_n \in \mathbb{C}$ are roots of p if and only if

$$\begin{cases} z_1 + z_2 + \ldots + z_n &= -\frac{p_{n-1}}{p_n}, \\ z_1z_2 + z_1z_3 + \ldots + z_{n-1}z_n &= \frac{p_{n-2}}{p_n}, \\ z_1z_2z_3 + z_1z_2z_4 + \ldots + z_{n-2}z_{n-1}z_n &= -\frac{p_{n-3}}{p_n}, \\ &\vdots \\ z_1z_2\cdots z_{n-1}z_n &= (-1)^n \cdot \frac{p_0}{p_n}. \end{cases}$$

(without proof)

Remark. If $p(z) = az^2 + bz + c \in \mathbb{C}[z]$, then numbers $z_1, z_2 \in \mathbb{C}$ are roots of p if and only if

$$\begin{cases} z_1 + z_2 &= -\frac{b}{a}, \\ z_1z_2 &= \frac{c}{a}. \end{cases}$$

Theorem. Let $p(x) \in \mathbb{R}[x]$ and let $\alpha \in \mathbb{C}$ be a root of $p(x)$. Then $\overline{\alpha}$ is also a root of $p(x)$.

Proof. We have $p(x) = p_n x^n + \ldots + p_1 x + p_0$, where $p_0, \ldots, p_n \in \mathbb{R}$. Since $\alpha \in \mathbb{C}$ is a root of $p(x)$, we must have $p(\alpha) = 0$. Hence,

$$0 = \overline{p(\alpha)} = \overline{p_n \alpha^n + \ldots + p_1 \alpha + p_0} = \overline{p_n}\ \overline{\alpha}^n + \ldots + \overline{p_1}\ \overline{\alpha} + \overline{p_0} = p_n \overline{\alpha}^n + \ldots + p_1 \overline{\alpha} + p_0 = p(\overline{\alpha}).$$

Thus $\overline{\alpha}$ is a root of $p(x)$. $\square$

Theorem. Every polynomial $p(x) \in \mathbb{R}[x]$ of odd degree has a root in $\mathbb{R}$.

Proof. Consider $p(x)$ as a polynomial in $\mathbb{C}[x]$. Then all roots are given by Fundamental theorem of algebra. Suppose on the contrary that none of these is real. Then, by previous Theorem, roots occur as conjugate pairs. It follows that there must be an even number of roots. This contradicts Fundamental theorem of algebra. $\square$

References

[1] O. Bretscher, *Linear algebra with applications*, Prentice Hall, New Jersey, 1997.

[2] S.I. Grossman, *Elementary linear algebra*, Saunders College Publishing, Philadelphia, 1991.

[3] W. Ledermann, *Complex Numbers*, Library of Mathematics, Routledge and Kegan Paul, London, 1962.