UNIVERSITÀ
DI SIENA
1240

Prof. Lonzi Marco

Lecture notes for the course

QUANTITATIVE METHODS FOR ECONOMIC APPLICATIONS

MATHEMATICS FOR ECONOMIC APPLICATIONS

Volume 1

Complex Numbers and Linear Algebra

AA. 2022/23

COMPLEX NUMBERS

Complex numbers arise historically to provide solutions to problems without a solution in a real environment. We begin with the following:

Definition 1 : The imaginary unit, denoted by the letter i, is the number (not real) such that $i^2 = -1$.

We can obtain the same definition supposing that there are numbers for which opposite and reciprocal coincide: $-x = \frac{1}{x}$, from which we obtain $x^2 = -1$, and so, algebraically solving, we obtain $x = \pm\sqrt{-1}$. Since, by definition, $i^2 = -1$, i and $-i$ are the solutions of such equation, therefore we obtain $\frac{1}{i} = -i$.

With regard to the powers of the imaginary unit i we have:

$i^0 = 1; i^1 = i; i^2 = -1; i^3 = i \cdot i^2 = -i; i^4 = i^2 \cdot i^2 = 1 = i^0$,

that is these are repeated with periodicity equal to 4. This allows a very quick calculation for such powers.

Example 1 : $i^{725} = i^{181 \cdot 4 + 1} = \left(i^4\right)^{181} \cdot i^1 = 1^{181} \cdot i = i$.

$i^{-321} = i^{-(80 \cdot 4) - 1} = \left(i^4\right)^{-80} \cdot i^{-1} = 1^{-80} \cdot \frac{1}{i} = -i$.

Definition 2 : Numbers of the form $k\,i, k \in \mathbb{R}$, are called imaginary (pure) numbers.

From imaginary numbers we define complex numbers:

Definition 3 : A complex number is a number of the form $a + bi$, with $a, b \in \mathbb{R}$, i.e. the sum of a real number with an imaginary number.

Number a is called the real part of the complex number $a + bi$, while bi is called the imaginary part, and b is called the coefficient of the imaginary.

A number $a + bi$ is called a complex number in algebraic form.

If $\mathbb{C} = \{a + bi, a, b \in \mathbb{R}\}$ is the set of complex numbers, we can easily see that $\mathbb{R} \subset \mathbb{C}$; in fact real numbers are a subset of complex numbers, as $a = a + 0\,i, \forall\, a \in \mathbb{R}$.

Let us now consider the pair $(a, b) \in \mathbb{R}^2$. It is easy to see that there is a bijection (one-to-one correspondence) between $\mathbb{C}$ and $\mathbb{R}^2$; to each pair (a, b) one and only one complex number in algebraic form $a + bi$ corresponds, and vice versa. There is thus a correspondence between complex numbers and points of the plane $\mathbb{R}^2$; the real part a is the abscissa, the coefficient of the imaginary b is the ordinate.

A Cartesian plane, at each point (a, b) of which the complex number $a + bi$ corresponds, is called complex (or Gauss) plane.

The horizontal axis is called the real axis, as on it numbers $a + 0i$ are placed, i.e. numbers which are real, while the vertical axis is called the imaginary axis, as on it numbers $0 + bi$ are placed, i.e. numbers which are imaginary.

The real number 0 corresponds to the couple $(0, 0)$, the real number 1 to the couple $(1, 0)$, the imaginary unit i to the couple $(0, 1)$.

OPERATIONS WITH COMPLEX NUMBERS

Definition 4 : Given two complex numbers in algebraic form $z_1 = a_1 + b_1 i$ and $z_2 = a_2 + b_2 i$, we define their sum and their difference as:

$z_1 + z_2 = (a_1 + b_1 i) + (a_2 + b_2 i) = (a_1 + a_2) + (b_1 + b_2)i$;

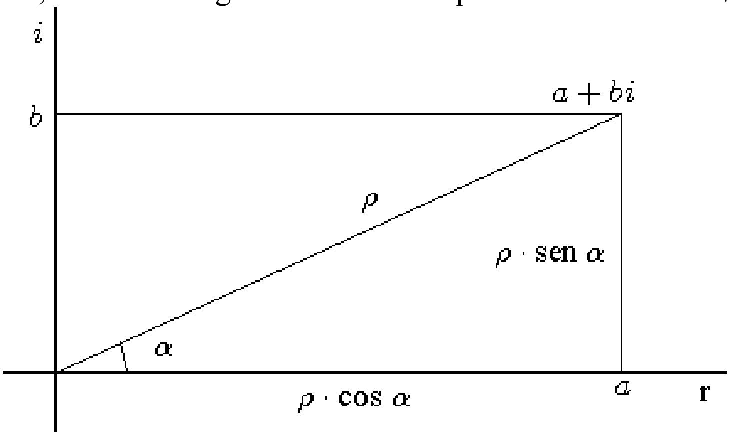
$z_1 - z_2 = (a_1 + b_1 i) - (a_2 + b_2 i) = (a_1 - a_2) + (b_1 - b_2)i$.

That is the sum (difference) of two complex numbers is a complex number having as real part the sum (difference) of the real parts and as imaginary part the sum (difference) of the imaginary parts.
Instead, using pair notation, if $z_1 = (a_1, b_1)$ and $z_2 = (a_2, b_2)$, we define:
$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2; b_1 + b_2)$ as the sum, and
$(a_1, b_1) - (a_2, b_2) = (a_1 - a_2; b_1 - b_2)$ as the difference of the two complex numbers.
Note the analogy with the sum and the difference of vectors in $\mathbb{R}^2$.
As for the product of two complex numbers in algebraic form, developing their product using the customary rules of literal calculus, and recalling that $i^2 = -1$, we obtain:
$z_1 \cdot z_2 = (a_1 + b_1 i) \cdot (a_2 + b_2 i) = a_1\, a_2 + a_1\, b_2 i + a_2\, b_1 i + b_1\, b_2\, i^2 =$
$= a_1\, a_2 + a_1\, b_2 i + a_2\, b_1 i - b_1\, b_2 = (a_1\, a_2 - b_1\, b_2) + (a_1\, b_2 + a_2\, b_1) i\,.$
Instead with pair notation we write :
$(a_1, b_1) \cdot (a_2, b_2) = (a_1\, a_2 - b_1\, b_2; a_1\, b_2 + a_2\, b_1)\,.$
It is easy to see that the neutral elements with respect to the sum and the product are still 0 and 1, i.e. the pairs $(0, 0)$ and $(1, 0)$.
Let us now calculate the reciprocal of a complex number $z = a + bi$. To do this let us introduce the concept of complex conjugate:
Definition 5 : Given the complex number $a + bi$, its complex conjugate is the complex number $a - bi$, i.e. the complex number having the same real part and the opposite for the imaginary coefficient.
The complex conjugate of z is denoted by $\overline{z}$, so we have $\overline{z} = a - bi$.
To calculate the reciprocal of z we multiply and divide by its conjugate $\overline{z}$, so we have:
$$\frac{1}{z} = \frac{1}{a+bi} = \frac{1}{a+bi} \cdot \frac{a-bi}{a-bi} = \frac{a-bi}{a^2-(bi)^2} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}\, i\,.$$
Using pair notation we have: $\dfrac{1}{(a,b)} = (a,b)^{-1} = \left(\dfrac{a}{a^2+b^2};\ -\dfrac{b}{a^2+b^2}\right)$.
Every complex number $z \neq 0$ has a unique reciprocal. Remember that $\dfrac{1}{i} = -i$.
Finally let us calculate the quotient $\dfrac{z_1}{z_2}$, treating it as the product $z_1 \cdot \dfrac{1}{z_2}$, and so we obtain:
$$\frac{z_1}{z_2} = \frac{a_1 + b_1 i}{a_2 + b_2 i} = \frac{a_1 + b_1 i}{a_2 + b_2 i} \cdot \frac{a_2 - b_2 i}{a_2 - b_2 i} = \frac{a_1 a_2 + a_2 b_1 i - a_1 b_2 i - b_1 b_2 i^2}{a_2^2 + b_2^2} =$$
$= \dfrac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2} + \dfrac{a_2 b_1 - a_1 b_2}{a_2^2 + b_2^2}\, i$. Using pair notation we have:
$$\frac{(a_1, b_1)}{(a_2, b_2)} = \left(\frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2}; \frac{a_2 b_1 - a_1 b_2}{a_2^2 + b_2^2}\right).$$

Example 2 : $(3 + 2i) - (5 - i) = -2 + 3i\,.$
$(3 + 2i) \cdot (5 - i) = 15 - 3i + 10i - 2i^2 = 17 + 7i\,.$
$$\frac{3+2i}{5-i} = \frac{3+2i}{5-i} \cdot \frac{5+i}{5+i} = \frac{15+3i+10i-2}{25+1} = \frac{13+13i}{26} = \frac{1}{2} + \frac{1}{2} i\,.$$

TRIGONOMETRIC (or POLAR) FORM FOR COMPLEX NUMBERS
Given a complex number $z = a + bi \neq 0$, as shown in the figure below, the following equalities hold:

$\begin{cases} a = \rho\cos\alpha \\ b = \rho\sin\alpha \end{cases}$, where $\rho = \sqrt{a^2+b^2} = |z|$ is the so called modulus of the complex number $z = a+bi$ while α, the angle formed by the segment joining points $(0,0)$ and (a,b) with the positive real semiaxis, is said the argument of the complex number $z = a+bi$.



Then, substituting, we have:
$z = a + bi = \rho\cos\alpha + i\,\rho\sin\alpha = \rho\,(\cos\alpha + i\sin\alpha)$,
which is called the trigonometric form of the complex number $a+bi$.
We note that $|\cos\alpha + i\sin\alpha| = \sqrt{\cos^2\alpha + \sin^2\alpha} = 1$.
Then we have $\dfrac{b}{a} = \dfrac{\rho\sin\alpha}{\rho\cos\alpha} = \operatorname{tg}\alpha$, from which we obtain $\alpha = \operatorname{arctg}\dfrac{b}{a},\ -\dfrac{\pi}{2} \le \alpha \le \dfrac{\pi}{2}$.
The trigonometric form of a complex number is not unique because, if $k \in \mathbb{Z}$, we have:
$\rho\,(\cos\alpha + i\sin\alpha) = \rho\,(\cos(\alpha+2k\pi) + i\sin(\alpha+2k\pi))$,
as trigonometric functions sinus and cosinus repeat over intervals of length 2π.

Example 3 : From $|i| = \sqrt{0+1} = 1$, we obtain : $i = 1\cdot\left(\cos\dfrac{\pi}{2} + i\sin\dfrac{\pi}{2}\right)$.
From $|-1| = 1$, we obtain $-1 = 1\cdot(\cos\pi + i\sin\pi)$.
From $|-1+i| = \sqrt{1+1} = \sqrt{2}$, we obtain : $-1+i = \sqrt{2}\left(-\dfrac{1}{\sqrt{2}} + \dfrac{1}{\sqrt{2}}i\right) =$
$= \sqrt{2}\left(\cos\dfrac{3}{4}\pi + i\sin\dfrac{3}{4}\pi\right)$; so for $z = -1+i$ we obtain $\rho = \sqrt{2}$ and $\alpha = \dfrac{3}{4}\pi$.
From $\left|2\sqrt{3}+2i\right| = \sqrt{12+4} = 4$, we get : $2\sqrt{3}+2i = 4\left(\dfrac{\sqrt{3}}{2} + \dfrac{1}{2}i\right) =$
$= 4\left(\cos\dfrac{\pi}{6} + i\sin\dfrac{\pi}{6}\right)$; so for $z = 2\sqrt{3}+2i$ we get $\rho = 4$ and $\alpha = \dfrac{\pi}{6}$.

OPERATIONS ON COMPLEX NUMBERS USING THE TRIGONOMETRIC FORM

Trigonometric form of complex numbers is not particularly useful for calculating sums and differences of complex numbers; for these operations it is more useful to work in algebraic form.
This is not true with regard to product, reciprocal, quotient, exponentiation and root extraction.
Given two complex numbers in trigonometric form:
$z_1 = \rho_1(\cos\alpha + i\sin\alpha)$ and $z_2 = \rho_2(\cos\beta + i\sin\beta)$, computing the product we have:
$z_1\cdot z_2 = \rho_1\,\rho_2\,(\cos\alpha + i\sin\alpha)\,(\cos\beta + i\sin\beta) =$
$= \rho_1\,\rho_2\left(\cos\alpha\cos\beta + i\sin\alpha\cos\beta + i\cos\alpha\sin\beta + i^2\sin\alpha\sin\beta\right) =$

$= \rho_1\,\rho_2\,(\cos\alpha\cos\beta - \sin\alpha\sin\beta + i\,(\sin\alpha\cos\beta + \cos\alpha\sin\beta)) =$
$= \rho_1\,\rho_2\,(\cos(\alpha+\beta) + i\sin(\alpha+\beta))$. And so:
Theorem 1 : To multiply two complex numbers in trigonometric form we multiply their moduli and we add their arguments.
It is easy to extend this rule to the product of as many complex numbers as one prefers.
Let us calculate, now, the reciprocal of a complex number $z \neq 0$; we have:

$$\frac{1}{z} = \frac{1}{\rho\,(\cos\alpha + i\sin\alpha)} = \frac{1}{\rho}\,\frac{(\cos\alpha - i\sin\alpha)}{(\cos\alpha + i\sin\alpha)\,(\cos\alpha - i\sin\alpha)} =$$
$$= \frac{1}{\rho}\,\frac{(\cos\alpha - i\sin\alpha)}{(\cos^2\alpha + \sin^2\alpha)} = \frac{1}{\rho}\,(\cos\alpha - i\sin\alpha) = \frac{1}{\rho}\,(\cos(-\alpha) + i\sin(-\alpha))$$. And so:

Theorem 2 : The reciprocal of a complex number in trigonometric form is a complex number having its modulus equal to the reciprocal of the modulus ρ and its argument equal to the opposite of the argument α.
Finally, let us calculate the quotient of two complex numbers in trigonometric form as the product of the first number by the reciprocal of the second. We have:

$$\frac{z_1}{z_2} = z_1 \cdot \frac{1}{z_2} = \rho_1\,(\cos\alpha + i\sin\alpha) \cdot \frac{1}{\rho_2}\,(\cos(-\beta) + i\sin(-\beta)) =$$
$$= \frac{\rho_1}{\rho_2}\,(\cos(\alpha-\beta) + i\sin(\alpha-\beta))$$. And so:

Theorem 3 : To divide two complex numbers in trigonometric form we divide their moduli and we subtract their arguments.
These formulas are also called De Moivre' formulas.

Example 4 : Since $i = 1\left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right)$ and $2\sqrt{3} + 2i = 4\left(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right)$, we obtain:

$$\frac{i}{2\sqrt{3}+2i} = \frac{1}{4}\left(\cos\left(\frac{\pi}{2} - \frac{\pi}{6}\right) + i\sin\left(\frac{\pi}{2} - \frac{\pi}{6}\right)\right) = \frac{1}{4}\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right) = \frac{1}{8} + \frac{\sqrt{3}}{8}\,i\,.$$

Since $-1 + i = \sqrt{2}\left(\cos\frac{3}{4}\pi + i\sin\frac{3}{4}\pi\right)$ we obtain:

$$\frac{1}{-1+i} = \frac{1}{\sqrt{2}}\left(\cos\left(-\frac{3\pi}{4}\right) + i\sin\left(-\frac{3\pi}{4}\right)\right) = \frac{1}{\sqrt{2}}\left(-\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\,i\right) =$$
$$= -\frac{1}{2} - \frac{1}{2}\,i\,.$$

POWERS OF COMPLEX NUMBERS USING TRIGONOMETRIC FORM
Using the product's formula, let us calculate the power of a complex number to a natural exponent.
If $z = \rho\,(\cos\alpha + i\sin\alpha)$ and if $n \in \mathbb{N}$, we have:
$z^n = [\rho\,(\cos\alpha + i\sin\alpha)]^n = \rho^n \cdot (\cos n\alpha + i\sin n\alpha)$, since the modulus is given by the product of n moduli all equal to ρ, while the argument is given by the sum of n arguments all equal to α. So:
Theorem 4 : To calculate the n-th power ($n \in \mathbb{N}$) of a complex number in trigonometric form we take the n-th power of its modulus ρ and we multiply its argument α by n.

Example 5 : Since $-1 + i = \sqrt{2}\left(\cos\frac{3}{4}\pi + i\sin\frac{3}{4}\pi\right)$, we obtain

$$(-1+i)^8 = \left(\sqrt{2}\right)^8 \left(\cos\left(8\cdot\frac{3}{4}\pi\right) + i\sin\left(8\cdot\frac{3}{4}\pi\right)\right) = 16\,(\cos 6\pi + i\sin 6\pi) = 16\,.$$

Let us consider now the powers of a complex number to an integer exponent $m \in \mathbb{Z}$. Since $\mathbb{Z}_+ = \mathbb{N}$, we need simply to define powers to negative exponent $m \in \mathbb{Z}_-$.
To this effect, let us suppose $m = -n, n \in \mathbb{N}$.
Since $z^{-n} = \left(z^{-1}\right)^n$, we have simply to apply the rule found in the case of natural exponents to the number z^{-1}, the reciprocal of z.
So we obtain, if $z = \rho\,(\cos\alpha + i\sin\alpha)$:

$$z^m = z^{-n} = \left(z^{-1}\right)^n = \left[\left(\rho\,(\cos\alpha + i\sin\alpha)\right)^{-1}\right]^n = \left[\frac{1}{\rho}\,(\cos(-\alpha) + i\sin(-\alpha))\right]^n =$$
$$= \frac{1}{\rho^n}\,(\cos(-n\,\alpha) + i\sin(-n\,\alpha)) = \rho^m\,(\cos m\,\alpha + i\sin m\,\alpha)\,.$$

So the rule is exactly the same as for powers to a natural exponent: its modulus is the m-th power of the modulus ρ and its argument is the multiple of the argument α by m.
Note that the result found for the reciprocal corresponds, of course, with the one found applying the exponent -1.

Example 6 : Since $2\sqrt{3} + 2i = 4\left(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right)$, we have also:

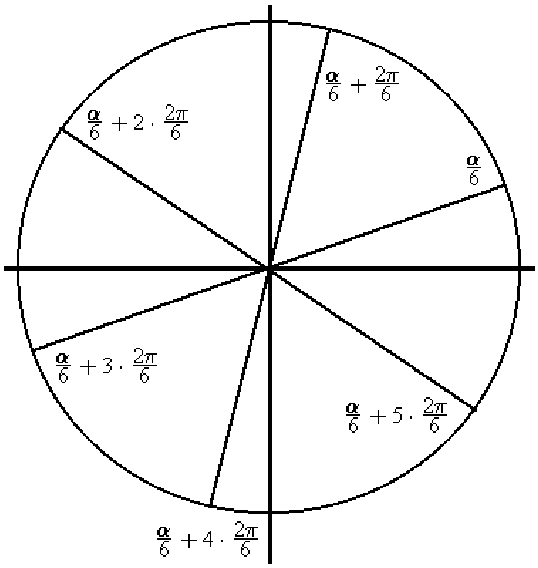
$$\left(2\sqrt{3} + 2i\right)^{-12} = 4^{-12}\left(\cos\left(-12\cdot\frac{\pi}{6}\right) + i\sin\left(-12\cdot\frac{\pi}{6}\right)\right) =$$
$$= \frac{1}{4^{12}}\,(\cos(-2\pi) + i\sin(-2\pi)) = \frac{1}{4^{12}}\,(1 + i\cdot 0) = \frac{1}{4^{12}}\,.$$

Let us study the powers to rational exponent, starting with the exponents like $\frac{1}{n}$, $n \in \mathbb{N}^*$; that is, let us study the problem of extracting the n-th root of a complex number.
We want to define $z^{\frac{1}{n}} = \sqrt[n]{z}$, with $z = \rho\,(\cos\alpha + i\sin\alpha)$.
By setting $z^{\frac{1}{n}} = \sqrt[n]{z} = w$, w unknown, let $w = x\,(\cos y + i\sin y)$, x and y unknown.
Since $z = w^n$, substituting we obtain: $\rho\,(\cos\alpha + i\sin\alpha) = x^n\,(\cos ny + i\sin ny)$.
The latter equality is satisfied if:

$$\begin{cases} \rho = x^n \\ \alpha + 2k\pi = ny \end{cases},\ \text{or if} \begin{cases} x = \sqrt[n]{\rho} \\ y = \dfrac{\alpha}{n} + k\,\dfrac{2\pi}{n}\,,\ k \in \mathbb{Z} \end{cases}.$$

The first equality has only one solution, that is the n-th positive root of ρ, while the second equality expresses the possibility that the arguments of the two complex numbers z and w^n give rise to the same point in the complex plane, although differing by integer multiples of a round.
The value $\frac{\alpha}{n}$ represents the n-th part of the argument α of the radicand z, while $\frac{2\pi}{n}$ represents the n-th part of a whole round.
If $k = 0$ we obtain $y = \frac{\alpha}{n}$, if $k = 1$ we obtain $y = \frac{\alpha}{n} + \frac{2\pi}{n}$ and so on; if $k = n - 1$ we obtain $y = \frac{\alpha}{n} + (n-1)\,\frac{2\pi}{n}$, and finally, if $k = n$ we obtain $y = \frac{\alpha}{n} + n\cdot\frac{2\pi}{n} = \frac{\alpha}{n} + 2\pi$, which in the complex plane represents the same point given by $y = \frac{\alpha}{n}$.

Having divided the round angle into n equal parts, starting from the position given by $y=\frac{\alpha}{n}$, after adding n of these parts we find ourselves again at the starting position. If $k=n+1$, $k=n+2$ and so on, we will meet the same points found earlier, and therefore the same n-th roots.
So the following Theorem is valid:
Theorem 5 : The n-th roots of a complex number z are in number of n and are given by the general formula:
$\sqrt[n]{z}=\sqrt[n]{\rho}\left(\cos\left(\frac{\alpha}{n}+k\,\frac{2\pi}{n}\right)+i\sin\left(\frac{\alpha}{n}+k\,\frac{2\pi}{n}\right)\right)$, $0\le k\le n-1, k\in\mathbb{N}$.
Every complex number $z\neq 0$ has exactly n n-th roots; these have the same modulus $\sqrt[n]{\rho}$, so they belong to a circumference having its center at $(0,0)$ and radius equal to $\sqrt[n]{\rho}$.
Since their arguments differ by an angle equal to $\frac{2\pi}{n}$, the n-th roots of a complex number z form the vertices of a regular polygon of n sides, inscribed in the circle with center $(0,0)$ and radius equal to $\sqrt[n]{\rho}$; the first of these vertices has $\frac{\alpha}{n}$ as its argument .
The following figure shows the six 6-th roots of $z=\rho\left(\cos\alpha+i\sin\alpha\right)$.



Example 7 : Let us compute $\sqrt[4]{i}$. Since $i=1\cdot\left(\cos\frac{\pi}{2}+i\sin\frac{\pi}{2}\right)$ it is:
$\sqrt[4]{i}=\sqrt[4]{1}\left(\cos\left(\frac{\pi}{8}+k\,\frac{2\pi}{4}\right)+i\sin\left(\frac{\pi}{8}+k\,\frac{2\pi}{4}\right)\right)$, $0\le k\le 3$, and so we obtain:
if $k=0$: $1\cdot\left(\cos\frac{\pi}{8}+i\sin\frac{\pi}{8}\right)$;
if $k=1$: $1\cdot\left(\cos\left(\frac{\pi}{8}+\frac{\pi}{2}\right)+i\sin\left(\frac{\pi}{8}+\frac{\pi}{2}\right)\right)=\left(\cos\frac{5\pi}{8}+i\sin\frac{5\pi}{8}\right)$;
if $k=2$: $1\cdot\left(\cos\left(\frac{\pi}{8}+\pi\right)+i\sin\left(\frac{\pi}{8}+\pi\right)\right)=\left(\cos\frac{9\pi}{8}+i\sin\frac{9\pi}{8}\right)$;
if $k=3$: $1\cdot\left(\cos\left(\frac{\pi}{8}+\frac{3\pi}{2}\right)+i\sin\left(\frac{\pi}{8}+\frac{3\pi}{2}\right)\right)=\left(\cos\frac{13\pi}{8}+i\sin\frac{13\pi}{8}\right)$.

Using the bisection (or half-angle) formulae: $\begin{cases} \sin\alpha = \sqrt{\dfrac{1-\cos 2\alpha}{2}} \\ \cos\alpha = \sqrt{\dfrac{1+\cos 2\alpha}{2}} \end{cases}$,

since: $\sin\frac{\pi}{4} = \cos\frac{\pi}{4} = \frac{\sqrt{2}}{2}$, we obtain:

$\sin\frac{\pi}{8} = \frac{\sqrt{2-\sqrt{2}}}{2}$ and $\cos\frac{\pi}{8} = \frac{\sqrt{2+\sqrt{2}}}{2}$, from which finally:

if $k=0$ it is: $\frac{\sqrt{2+\sqrt{2}}}{2} + i\,\frac{\sqrt{2-\sqrt{2}}}{2}$;

if $k=1$ it is: $-\frac{\sqrt{2-\sqrt{2}}}{2} + i\,\frac{\sqrt{2+\sqrt{2}}}{2}$;

if $k=2$ it is: $-\frac{\sqrt{2+\sqrt{2}}}{2} - i\,\frac{\sqrt{2-\sqrt{2}}}{2}$;

if $k=3$ it is: $\frac{\sqrt{2-\sqrt{2}}}{2} - i\,\frac{\sqrt{2+\sqrt{2}}}{2}$.

Example 8 : Let us compute $\sqrt[n]{1}$. Since $1 = 1\cdot(\cos 0 + i\sin 0)$ we obtain:

$\sqrt[n]{1} = \sqrt[n]{1}\cdot\left(\cos\left(\frac{0}{n} + k\,\frac{2\pi}{n}\right) + i\sin\left(\frac{0}{n} + k\,\frac{2\pi}{n}\right)\right)$, $0 \le k \le n-1, k\in\mathbb{N}$, i.e.:

$\sqrt[n]{1} = \cos\left(k\,\frac{2\pi}{n}\right) + i\sin\left(k\,\frac{2\pi}{n}\right)$, $0 \le k \le n-1$.

From $\sqrt[n]{z} = \sqrt[n]{\rho}\left(\cos\left(\frac{\alpha}{n} + k\,\frac{2\pi}{n}\right) + i\sin\left(\frac{\alpha}{n} + k\,\frac{2\pi}{n}\right)\right)$, $0 \le k \le n-1, k\in\mathbb{N}$,
using the product of complex numbers in trigonometric form, we can write:

$\sqrt[n]{z} = \sqrt[n]{\rho}\cdot\sqrt[n]{1}\cdot\left(\cos\frac{\alpha}{n} + i\sin\frac{\alpha}{n}\right)\left(\cos\left(k\,\frac{2\pi}{n}\right) + i\sin\left(k\,\frac{2\pi}{n}\right)\right)$,

where $\sqrt[n]{\rho}\left(\cos\frac{\alpha}{n} + i\sin\frac{\alpha}{n}\right)$ is the first n-th root of the number z, the one corresponding to $k=0$, while $\left(\cos\left(k\,\frac{2\pi}{n}\right) + i\sin\left(k\,\frac{2\pi}{n}\right)\right)$, as we have seen in Example 8, represents, when $0 \le k \le n-1$, the n-th roots of the unity 1. And so the following Theorem is valid:

Theorem 6 : The n-th roots of every complex number $z \ne 0$ can be obtained calculating only one root, that corresponding to $k=0$, and then multiplying this by the n n-roots of the unity 1.

Example 9 : Let us compute $\sqrt[4]{1}$ and then $\sqrt[4]{i}$. We obtain:

$\sqrt[4]{1} = \sqrt[4]{1}\cdot\left(\cos\left(k\,\frac{2\pi}{4}\right) + i\sin\left(k\,\frac{2\pi}{4}\right)\right) = \left(\cos\left(k\,\frac{\pi}{2}\right) + i\sin\left(k\,\frac{\pi}{2}\right)\right)$, $0 \le k \le 3$,

and so: if $k=0$ it is $\cos 0 + i\sin 0 = 1$; if $k=1$ we have $\cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i$;

if $k=2$ we have $\cos\pi + i\sin\pi = -1$; if $k=3$ we have $\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2} = -i$.

Taking the first root found for $\sqrt[4]{i}$, i.e., if $k=0$: $\dfrac{\sqrt{2+\sqrt{2}}}{2}+i\,\dfrac{\sqrt{2-\sqrt{2}}}{2}$, multiplying it by i, by -1 and by $-i$, we find the other 4-th roots of i we have just found. In fact:

$\left(\dfrac{\sqrt{2+\sqrt{2}}}{2}+i\,\dfrac{\sqrt{2-\sqrt{2}}}{2}\right)\cdot i = -\dfrac{\sqrt{2-\sqrt{2}}}{2}+i\,\dfrac{\sqrt{2+\sqrt{2}}}{2}$, the one for $k=1$;

$\left(\dfrac{\sqrt{2+\sqrt{2}}}{2}+i\,\dfrac{\sqrt{2-\sqrt{2}}}{2}\right)\cdot(-1) = -\dfrac{\sqrt{2+\sqrt{2}}}{2}-i\,\dfrac{\sqrt{2-\sqrt{2}}}{2}$, the one for $k=2$;

$\left(\dfrac{\sqrt{2+\sqrt{2}}}{2}+i\,\dfrac{\sqrt{2-\sqrt{2}}}{2}\right)\cdot(-i) = \dfrac{\sqrt{2-\sqrt{2}}}{2}-i\,\dfrac{\sqrt{2+\sqrt{2}}}{2}$, the one for $k=3$.

Example 10 : Let us compute $\sqrt{\left(\dfrac{\sqrt{3}}{2}-\dfrac{1}{2}\,i\right)^2}$. Since $\left|\dfrac{\sqrt{3}}{2}-\dfrac{1}{2}\,i\right|=1$, it is:

$\dfrac{\sqrt{3}}{2}-\dfrac{1}{2}\,i=\cos\left(\dfrac{11}{6}\,\pi\right)+i\sin\left(\dfrac{11}{6}\,\pi\right)$, from which we obtain:

$\left(\dfrac{\sqrt{3}}{2}-\dfrac{1}{2}\,i\right)^2=\cos\left(\dfrac{11}{3}\,\pi\right)+i\sin\left(\dfrac{11}{3}\,\pi\right)=$

$=\cos\left(\dfrac{5}{3}\,\pi\right)+i\sin\left(\dfrac{5}{3}\,\pi\right)=\dfrac{1}{2}-\dfrac{\sqrt{3}}{2}\,i$.

So $\sqrt{\dfrac{1}{2}-\dfrac{\sqrt{3}}{2}\,i}=\cos\left(\dfrac{5}{6}\,\pi+k\,\dfrac{2\pi}{2}\right)+i\sin\left(\dfrac{5}{6}\,\pi+k\,\dfrac{2\pi}{2}\right)$, $0\le k\le 1$:

if $k=0$ we obtain: $\cos\left(\dfrac{5}{6}\,\pi\right)+i\sin\left(\dfrac{5}{6}\,\pi\right)=-\dfrac{\sqrt{3}}{2}+\dfrac{1}{2}\,i$;

if $k=1$ we obtain: $\cos\left(\dfrac{11}{6}\,\pi\right)+i\sin\left(\dfrac{11}{6}\,\pi\right)=\dfrac{\sqrt{3}}{2}-\dfrac{1}{2}\,i$.

As it can be seen, therefore, it is incorrect to write $\sqrt{\left(\dfrac{\sqrt{3}}{2}-\dfrac{1}{2}\,i\right)^2}=\dfrac{\sqrt{3}}{2}-\dfrac{1}{2}\,i$.

Example 11 : Let us solve the equation $x^2+x+1=0$.
We obtain $x=\dfrac{-1\pm\sqrt{1-4}}{2}=\dfrac{-1\pm\sqrt{-3}}{2}$.
Since $\sqrt{-3}=\sqrt{3}\cdot\sqrt{-1}$, we obtain $x=\dfrac{-1\pm\sqrt{3}\,i}{2}$, and then we have two complex conjugate solutions: $x_1=-\dfrac{1}{2}+\dfrac{\sqrt{3}}{2}\,i$ and $x_2=-\dfrac{1}{2}-\dfrac{\sqrt{3}}{2}\,i$.

Example 12 : Let us solve the equation $x^3+1=0$, having only one real solution: $x=-1$.
From $x^3=-1$, we obtain $x=\sqrt[3]{-1}$, and so we have to calculate the three 3-th roots of $z=1\cdot(\cos\pi+i\sin\pi)$:
$\sqrt[3]{1}\cdot\left(\cos\left(\dfrac{\pi}{3}+k\,\dfrac{2\pi}{3}\right)+i\sin\left(\dfrac{\pi}{3}+k\,\dfrac{2\pi}{3}\right)\right)$, $0\le k\le 2$, and so:

if $k=0$ we obtain: $\cos\frac{\pi}{3}+i\sin\frac{\pi}{3}=\frac{1}{2}+\frac{\sqrt{3}}{2}\,i$;
if $k=1$ we obtain: $\cos\pi+i\sin\pi=\,-1$;
if $k=2$ we obtain: $\cos\left(\frac{5}{3}\,\pi\right)+i\sin\left(\frac{5}{3}\,\pi\right)=\frac{1}{2}-\frac{\sqrt{3}}{2}\,i$.
Therefore we have found three solutions, whose number is equal to the degree of the polynomial x^3+1.

Let us finally deal with the powers to rational exponent, $z^{\frac{m}{n}}$, $\frac{m}{n}\in\mathbb{Q}$; let us suppose that m and n are relatively prime numbers (two numbers are "relatively prime" if they have no common factors other than 1), and $m\neq 1$.
We put $z^{\frac{m}{n}}=(z^m)^{\frac{1}{n}}=\sqrt[n]{z^m}$, and we operate accordingly to the above definitions.
The power z^m gives only one result, and then we calculate the n n-th roots of this number.

THE COMPLEX EXPONENTIAL e^z, $z\in\mathbb{C}$
Given a pure imaginary number $z=x\,i$, $x\in\mathbb{R}$, let us state the following
Definition 6 : We define the complex exponential $e^{x\,i}$ as:
$e^{x\,i}=\cos x+i\sin x$
which is also called Euler's formula.
Let us see a justification (certainly not a demonstration !) of this definition using the MacLaurin' polynomials of the real functions e^x, $\sin x$ and $\cos x$, although, more correctly, we should use their power series developments. We know that it is:

$$e^x=1+x+\frac{x^2}{2\,!}+\frac{x^3}{3\,!}+....+\frac{x^n}{n\,!}+o(x^n)=\sum_{k=0}^{n}\frac{1}{k\,!}\,x^k+o(x^n).$$

$$\sin x=x-\frac{x^3}{3\,!}+\frac{x^5}{5\,!}-\frac{x^7}{7\,!}+\frac{x^9}{9\,!}-\,...+\frac{(-1)^n}{(2n+1)\,!}\,x^{2n+1}+o\big(x^{2n+1}\big)$$

$$\cos x=1-\frac{x^2}{2\,!}+\frac{x^4}{4\,!}-\frac{x^6}{6\,!}+\frac{x^8}{8\,!}-\,...+\frac{(-1)^n}{(2n)\,!}\,x^{2n}+o\big(x^{2n}\big).$$

Substituting, in a formal way, in these expressions the variable xi to the variable x, we obtain:

$$e^{x\,i}=1+(x\,i)+\frac{(x\,i)^2}{2\,!}+\frac{(x\,i)^3}{3\,!}+\frac{(x\,i)^4}{4\,!}+\frac{(x\,i)^5}{5\,!}+\frac{(x\,i)^6}{6\,!}+\frac{(x\,i)^7}{7\,!}+....\text{ and so:}$$

$$e^{x\,i}=1+x\,i-\frac{x^2}{2\,!}-\frac{x^3}{3\,!}\,i+\frac{x^4}{4\,!}+\frac{x^5}{5\,!}\,i-\frac{x^6}{6\,!}-\frac{x^7}{7\,!}\,i+....\text{ or:}$$

$$e^{x\,i}=\left(1-\frac{x^2}{2\,!}+\frac{x^4}{4\,!}-\frac{x^6}{6\,!}+....\right)+i\left(x-\frac{x^3}{3\,!}+\frac{x^5}{5\,!}-\frac{x^7}{7\,!}+....\right)\text{ and finally:}$$

$$e^{x\,i}=\cos x+i\sin x.$$

Given now a complex number $z\in\mathbb{C}, z=x+y\,i$, using the same properties as real powers, we obtain:
$e^z=e^{x+y\,i}=e^x\cdot e^{y\,i}=e^x\,(\cos y+i\sin y)$
i.e. we obtain a complex number having modulus equal to e^x and argument equal to the coefficient of the imaginary y. In fact: $|\cos y+i\sin y|=1$.
From this definition it follows immediately, $\forall\, k\in\mathbb{Z}$, that:
$e^{\,z+2k\pi\,i}=e^{x+y\,i+2k\pi\,i}=e^x\cdot e^{(y+2k\pi)\,i}=e^x\,(\cos\,(y+2k\pi)+i\sin\,(y+2k\pi))=$
$=e^x\,(\cos y+i\sin y)=e^z$, that is, the complex exponential function $z\to e^z$ is a periodic function, having an immaginay period equal to $2\pi\,i$.

Example 13 : Let us compute e^i . Being $e^i = e^{0+1\cdot i}$ we obtain:
$e^i = e^0\left(\cos 1 + i\sin 1\right) = \cos 1 + i \sin 1$.
If we compute $e^{2\pi i}$ we obtain instead $e^{2\pi i} = e^{0+2\pi i} = e^0\left(\cos 2\pi + i\sin 2\pi\right) = 1$.

LOGARITHMS OF COMPLEX NUMBERS $\log z,\, z \in \mathbb{C}$
Now let us see how to define $\log z,\, z \in \mathbb{C},\, z \neq 0$. If $\log z = w$, we obtain $z = e^w$.
If we set $w = x + y\,i$, being x and y unknown, and $z = \rho\left(\cos\alpha + i\sin\alpha\right)$, being ρ and α known values, we obtain $e^w = e^{x+y\,i} = e^x\left(\cos y + i\sin y\right) = \rho\left(\cos\alpha + i\sin\alpha\right)$, which is satisfied when : $\begin{cases} e^x = \rho \\ y = \alpha + 2k\pi,\, k \in \mathbb{Z} \end{cases}$ or if $\begin{cases} x = \log\rho \\ y = \alpha + 2k\pi,\, k \in \mathbb{Z} \end{cases}$.
We note that $\log\rho$ is always defined, since ρ, a modulus, is always a real positive number; the second equality depends on the fact that a point in the complex plane can be represented in an infinite number of ways, given the identity of representations at less than full rounds.
Substituting the equalities found we get: $\log z = w = x + y\,i = \log\rho + \left(\alpha + 2k\pi\right)i$, $k \in \mathbb{Z}$.
With this equality we define the infinite logarithms of a complex number $z \neq 0$.
They have all the same real part, $\log\rho$, while the coefficient of their imaginary part varies, adding multiples of 2π.
The values of $\log z$ form a sequence of equally spaced points along a vertical line passing through the point $(\log\rho, \alpha)$.
The value corresponding to $\alpha = 0$ is called the principal value.

Example 14 : Let us compute $\log\left(-1\right)$. Since $-1 = 1\cdot\left(\cos\pi + i\sin\pi\right)$, we obtain:
$\log\left(-1\right) = \log 1 + \left(\pi + 2k\pi\right)i = \left(2k+1\right)\pi\, i,\ k \in \mathbb{Z}$. From this we obtain also:
$e^{(2k+1)\pi\, i} = \cos\left(\left(2k+1\right)\pi\right) + i\sin\left(\left(2k+1\right)\pi\right) = -1$.
Example 15 : Let us compute $\log i$. Since $i = \left(\cos\dfrac{\pi}{2} + i\sin\dfrac{\pi}{2}\right)$, we obtain:
$\log i = \log 1 + \left(\dfrac{\pi}{2} + 2k\pi\right)i = \left(\dfrac{\pi}{2} + 2k\pi\right)i,\ k \in \mathbb{Z}$.
Example 16 : Let us compute $\log\left(1+i\right)$.
Since $1 + i = \sqrt{2}\left(\dfrac{1}{\sqrt{2}} + i\,\dfrac{1}{\sqrt{2}}\right) = \sqrt{2}\left(\cos\dfrac{\pi}{4} + i\sin\dfrac{\pi}{4}\right)$, we obtain:
$\log\left(1+i\right) = \log\sqrt{2} + \left(\dfrac{\pi}{4} + 2k\pi\right)i,\ k \in \mathbb{Z}$.

POWERS TO A COMPLEX EXPONENT
If we want to raise a complex number to a complex number, i.e. if we want to define a power such as $w^z,\, w \in \mathbb{C},\, z \in \mathbb{C},\, w \neq 0$, we use the equality, valid for every real positive number a: $a^x = e^{x\log a}$.
Definition 7 : We set $w^z = e^{z\log w}$, where both the exponential and the logarithm are the complex ones.

Example 17 : Let us compute i^i . Since $i = \cos\dfrac{\pi}{2} + i\sin\dfrac{\pi}{2}$ and $\log i = \left(\dfrac{\pi}{2} + 2k\pi\right)i$, we obtain: $i^i = e^{i\log i} = e^{i\left(\frac{\pi}{2}+2k\pi\right)i} = e^{-\left(\frac{\pi}{2}+2k\pi\right)}$, $k \in \mathbb{Z}$.
Then the power i^i takes infinite values, which are all still real.

Let us compute now $(1+i)^{1-i} = e^{(1-i)\log(1+i)}$. Since $1+i = \sqrt{2}\left(\frac{1}{\sqrt{2}} + i\,\frac{1}{\sqrt{2}}\right)$, we have just seen (Example 16) that $\log(1+i) = \log\sqrt{2} + \left(\frac{\pi}{4} + 2k\pi\right) i$, and so, substituting, we obtain:

$$e^{(1-i)\log(1+i)} = e^{(1-i)\left(\log\sqrt{2}+\left(\frac{\pi}{4}+2k\pi\right) i\right)} = e^{\log\sqrt{2}+\left(\frac{\pi}{4}+2k\pi\right) i - i\log\sqrt{2}+\left(\frac{\pi}{4}+2k\pi\right)} =$$
$$= e^{\log\sqrt{2}+\frac{\pi}{4}+2k\pi} \cdot e^{\left(\frac{\pi}{4}+2k\pi-\log\sqrt{2}\right) i} = e^{\log\sqrt{2}} \cdot e^{\frac{\pi}{4}+2k\pi} \cdot e^{\left(\frac{\pi}{4}+2k\pi-\log\sqrt{2}\right) i} =$$
$$= \sqrt{2}\, e^{\frac{\pi}{4}+2k\pi}\left(\cos\left(\frac{\pi}{4}+2k\pi-\log\sqrt{2}\right) + i\sin\left(\frac{\pi}{4}+2k\pi-\log\sqrt{2}\right)\right).$$

COMPLEX TRIGONOMETRIC FUNCTIONS

From the definition $e^{xi} = \cos x + i\sin x$, substituting xi with $(-xi)$, we obtain the following: $e^{-xi} = \cos(-x) + i\sin(-x) = \cos x - i\sin x$.

Adding and subtracting them from the two equalities $\begin{cases} e^{xi} = \cos x + i\sin x \\ e^{-xi} = \cos x - i\sin x \end{cases}$ we obtain:

$\begin{cases} e^{xi} + e^{-xi} = 2\cos x \\ e^{xi} - e^{-xi} = 2\,i\sin x \end{cases}$ and then: $\begin{cases} \cos x = \dfrac{e^{xi} + e^{-xi}}{2} \\ \sin x = \dfrac{e^{xi} - e^{-xi}}{2i} \end{cases}$.

Extending these equalities to $z \in \mathbb{C}$, we obtain the definition of the sinus and the cosinus of a complex number: $\begin{cases} \cos z = \dfrac{e^{zi} + e^{-zi}}{2} \\ \sin z = \dfrac{e^{zi} - e^{-zi}}{2i} \end{cases}$.

From these we then also obtain $\operatorname{tg} z = \dfrac{\sin z}{\cos z} = \dfrac{e^{zi} - e^{-zi}}{2i}\,\dfrac{2}{e^{zi} + e^{-zi}} = \dfrac{1}{i} \cdot \dfrac{e^{zi} - e^{-zi}}{e^{zi} + e^{-zi}}$.

LINEAR ALGEBRA

VECTORS

Let $\mathbb{R}^n$ be the n-dimension vector space, whose elements are n-tuples $(x_1, x_2, ..., x_n)$ of real numbers, resulting from the Cartesian product of $\mathbb{R}$ by itself n times. Each n-tuple is also known as a vector.

Each vector will be denoted by a capital letter or with the n-tuple of its components: $\mathbb{X} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$.

If $\mathbb{X} \in \mathbb{R}^n$ we also say that $\mathbb{X}$ has n components or that it is an n-dimension vector.

From a geometric point of view, every vector $\mathbb{X} \in \mathbb{R}^n$ identifies the straight line passing through the two points $\mathbb{O} = (0, 0, ..., 0)$ and $\mathbb{X} = (x_1, x_2, ..., x_n)$.

The vector $\mathbb{O} = (0, 0, ..., 0)$ is called the null vector.

OPERATIONS WITH VECTORS

Let us consider two vectors: $\mathbb{X}, \mathbb{Y} \in \mathbb{R}^n$, $\mathbb{X} = (x_1, x_2, ..., x_n)$ and $\mathbb{Y} = (y_1, y_2, ..., y_n)$, having the same number of components.

Definition 8 : We define vectors addition as:

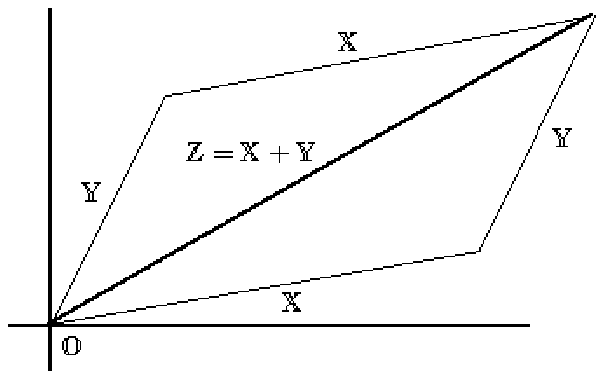
$\mathbb{X} + \mathbb{Y} = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) \in \mathbb{R}^n$.

We easily extend this definition to the sum of any number of vectors.

Example 18 : If $\mathbb{X} = (3, -1, 0)$ and $\mathbb{Y} = (5, 2, -3)$, then $\mathbb{X} + \mathbb{Y} = (8, 1, -3)$.

Vector sum obeys to the following rules:
S1) Commutativity: $\mathbb{X} + \mathbb{Y} = \mathbb{Y} + \mathbb{X}$
S2) Associativity: $\mathbb{X} + (\mathbb{Y} + \mathbb{Z}) = (\mathbb{X} + \mathbb{Y}) + \mathbb{Z}$.
S3) Identity element: there exists only one element $\mathbb{O} \in \mathbb{R}^n$, the null vector, such that $\mathbb{X} + \mathbb{O} = \mathbb{O} + \mathbb{X} = \mathbb{X}, \forall \mathbb{X} \in \mathbb{R}^n$.

Given any two vectors, $\mathbb{X} \in \mathbb{R}^n$ and $\mathbb{Y} \in \mathbb{R}^n$, let us consider the parallelogram spanned by these two vectors and its diagonal starting at the origin. This diagonal is the vector $\mathbb{Z}$ that represents the sum of the two vectors: $\mathbb{Z} = \mathbb{X} + \mathbb{Y}$. This is the so-called parallelogram rule.



Let us consider one vector: $\mathbb{X} \in \mathbb{R}^n$, $\mathbb{X} = (x_1, x_2, ..., x_n)$ and a real number (scalar) $k \in \mathbb{R}$.
Definition 9 : We define scalar multiplication as:
$k \cdot \mathbb{X} = (k\, x_1, k\, x_2, ..., k\, x_n) \in \mathbb{R}^n$.

Example 19 : If $\mathbb{X} = (3, -1, 0)$ and $k = 5,$ we obtain $5 \cdot \mathbb{X} = (15, -5, 0)$.

Scalar multiplication obeys to the following rules:
If $k_1, k_2 \in \mathbb{R}$ and $\mathbb{X}, \mathbb{Y} \in \mathbb{R}^n$:
P1) Associativity: $(k_1 \cdot k_2) \cdot \mathbb{X} = k_1 \cdot (k_2 \cdot \mathbb{X})$;
P2) Commutativity: $k_1 \cdot (k_2 \cdot \mathbb{X}) = k_2 \cdot (k_1 \cdot \mathbb{X})$;
P3) Scalar distributivity: $(k_1 + k_2) \cdot \mathbb{X} = k_1 \cdot \mathbb{X} + k_2 \cdot \mathbb{X}$;
P4) Vector distributivity: $k \cdot (\mathbb{X} + \mathbb{Y}) = k \cdot \mathbb{X} + k \cdot \mathbb{Y}$;
P5) Identity element: Multiplying by the scalar 1 does not change a vector: $1 \cdot \mathbb{X} = \mathbb{X}$.

The scalar multiplication of a vector $\mathbb{X}$ by a real positive number $k > 1$ multiplies the magnitude of the vector without changing its direction.
The scalar multiplication of a vector $\mathbb{X}$ by a real positive number $0 < k < 1$ decreases the magnitude of the vector without changing its direction.
The scalar multiplication of a vector $\mathbb{X}$ by a real negative number $k < -1$ multiplies the magnitude of the vector reversing its direction.
The scalar multiplication of a vector $\mathbb{X}$ by a real negative number $-1 < k < 0$ decreases the magnitude of the vector reversing its direction.

Definition 10 : Given p vectors $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_p \in \mathbb{R}^n$ and p scalars $k_1, k_2, ..., k_p \in \mathbb{R}$, we define the linear combination of these vectors with the coefficients k_i as the vector:

$$k_1\,\mathbb{X}_1\;+\;k_2\,\mathbb{X}_2\;+\;...\;+\;k_p\,\mathbb{X}_p=\sum_{i=1}^{p}k_i\,\mathbb{X}_i\,,$$

i.e. the sum of the p vectors $\mathbb{X}_i$, each multiplied by its scalar coefficient k_i.

Example 20 : Given $\mathbb{X}=(3,-1,0)$, $\mathbb{Y}=(5,2,-3)$ and $\mathbb{Z}=(-4,2,2)$, if $k_1=3$, $k_2=-2$ and $k_3=2$, we obtain: $3\cdot\mathbb{X}\;+\;(-2)\cdot\mathbb{Y}\;+\;2\cdot\mathbb{Z}=(-9,-3,10)$.

Definition 11 : Given any two vectors $\mathbb{X},\mathbb{Y}\in\mathbb{R}^n$, the difference of $\mathbb{X}$ and $\mathbb{Y}$ is given by their linear combination with coefficients 1 and -1:
$\mathbb{X}-\mathbb{Y}=1\cdot\mathbb{X}+(-1)\cdot\mathbb{Y}=(x_1-y_1,x_2-y_2,...,x_n-y_n)$.

Multiplying a vector $\mathbb{X}$ by -1 we obtain its additive inverse element:
$\mathbb{X}+(-1)\,\mathbb{X}=\mathbb{X}-\mathbb{X}=\mathbb{O}$.

Also the difference between two vectors can be represented graphically.
If $\mathbb{Z}=\mathbb{X}+\mathbb{Y}$, then $\mathbb{Y}=\mathbb{Z}-\mathbb{X}$, and referring to the figure relative to the parallelogram rule, we see that the vector $\mathbb{Y}$, which represents the difference, is the vector leading from the point $\mathbb{X}$, the vector that is subtracted, to the point $\mathbb{Z}$, the one from which we subtract.

VECTOR SPACES AND SUBSPACES
The operations we have defined, when executed on vectors of $\mathbb{R}^n$, give always a vector of $\mathbb{R}^n$ as their result.
The scalar multiplication of a vector $\mathbb{X}$ by a scalar k results in a vector which belongs to the same straight line of $\mathbb{X}$, without changing its direction if $k>0$, reversing its direction if $k<0$.
Two vectors $\mathbb{X},\mathbb{Y}\in\mathbb{R}^n$ belong to the same straight line if and only if $\mathbb{X}=k\cdot\mathbb{Y},k\in\mathbb{R}$; in this case the two vectors are said to be parallel vectors.
Two vectors $\mathbb{X},\mathbb{Y}\in\mathbb{R}^n$, that do not belong to the same straight line, identify a plan passing through $\mathbb{O}$, $\mathbb{X}$ and $\mathbb{Y}$, and any linear combination of these two vectors will result in a vector that belongs to the same plane.

A set in which these two operations are defined, i.e. the sum and the scalar multiplication, with the properties listed above, is said a vector space.

More precisely, we set the following
Definition 12 : A set $\mathbb{V}$ is said to be a vector space if:
$\forall\,k_1,k_2\in\mathbb{R}$ and $\forall\,\mathbb{X},\mathbb{Y}\in\mathbb{V}\;\Rightarrow\;k_1\cdot\mathbb{X}+k_2\cdot\mathbb{Y}\in\mathbb{V}$,
i.e. if every linear combination of its elements still belongs to the set.

The sets $\mathbb{R}$ (the real line), $\mathbb{R}^2$ (the real or Cartesian plane), $\mathbb{R}^3$ (the space of Euclidean geometry), ..., $\mathbb{R}^n$ are the main examples or real vector spaces, having dimensions respectively $1,2,3,...\,n$. The null vector (or the point) $\mathbb{O}$ can be considered like the vector space having its dimension equal to zero.

Example 21 : The set of polynomials $\mathrm{P}(x)=a_n\,x^n+a_{n-1}\,x^{n-1}+...+a_1\,x+a_0$, regardless of their degree, constitutes a vector space, since each linear combination of polynomials is still a polynomial.

Example 22 : The set of continuous functions at a point x_0 is a vector space since the linear combination of continuous functions at x_0 is still a continuous function at x_0. Similarly for differentiable functions at x_0.

With a similar definition, a subset $\mathbb{A} \subset \mathbb{V}$ is said to be a vector subspace of $\mathbb{V}$ if:
$\forall\, k_1, k_2 \in \mathbb{R}$ and $\forall\, \mathbb{X}, \mathbb{Y} \in \mathbb{A} \;\Rightarrow\; k_1 \cdot \mathbb{X} + k_2 \cdot \mathbb{Y} \in \mathbb{A}$.

For example, every line ($\mathbb{R}$) is a vector subspace in the plane $\mathbb{R}^2$, in $\mathbb{R}^3$, and in every $\mathbb{R}^n$; every plane ($\mathbb{R}^2$) is a vector subspace in $\mathbb{R}^3$, in $\mathbb{R}^4$, and in every $\mathbb{R}^n$, and so on.

LINEARLY DEPENDENT OR INDEPENDENT VECTORS
Given p vectors $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_p \in \mathbb{R}^n, p \leq n$;
Definition 13 : The vectors $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_p$ are said to be linearly dependent if there are p scalars $k_1, k_2, ..., k_p$, not all equal to zero (that is, with at least one of them different from zero), such that:
$k_1 \cdot \mathbb{X}_1 + k_2 \cdot \mathbb{X}_2 + ... + k_p \cdot \mathbb{X}_p = \mathbb{O}$.
If the only way to obtain the null vector $\mathbb{O}$ as the result of the linear combination is to take all the scalars k_i equal to zero, then the p vectors are said to be linearly independent.
If the p vectors are linearly dependent, each of the vectors having, in the linear combination that gives the null vector as its result, a coefficient different from zero, can be expressed as a linear combination of the other vectors. In fact, if $k_i \neq 0$, from:
$k_1 \cdot \mathbb{X}_1 + k_2 \cdot \mathbb{X}_2 + ... + k_i \cdot \mathbb{X}_i + ... + k_p \cdot \mathbb{X}_p = \mathbb{O}$, we obtain:
$k_i \cdot \mathbb{X}_i = -k_1 \cdot \mathbb{X}_1 - k_2 \cdot \mathbb{X}_2 - ... - k_{i-1} \cdot \mathbb{X}_{i-1} - k_{i+1} \cdot \mathbb{X}_{i+1} - ... - k_p \cdot \mathbb{X}_p$,
and then
$$\mathbb{X}_i = -\frac{k_1}{k_i} \cdot \mathbb{X}_1 - \frac{k_2}{k_i} \cdot \mathbb{X}_2 - ... - \frac{k_{i-1}}{k_i} \cdot \mathbb{X}_{i-1} - \frac{k_{i+1}}{k_i} \cdot \mathbb{X}_{i+1} - ... - \frac{k_p}{k_i} \cdot \mathbb{X}_p,$$
or, if $h_j = -\dfrac{k_j}{k_i}$:
$\mathbb{X}_i = h_1 \cdot \mathbb{X}_1 + h_2 \cdot \mathbb{X}_2 + ... + h_{i-1} \cdot \mathbb{X}_{i-1} + h_{i+1} \cdot \mathbb{X}_{i+1} + ... + h_p \cdot \mathbb{X}_p$.
This would not be possible if the vectors were linearly independent.
The vectors $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_p$ are linearly independent if none of them can be expressed as a linear combination of the remaining ones.

From a geometrical point of view, if we say that p vectors belonging to $\mathbb{R}^n$, with $p \leq n$, are linearly dependent, it means that they belong to a same vector subspace $\mathbb{R}^k$, whose dimension k is less than the number p of the vectors.

Example 23 : 4 vectors in $\mathbb{R}^5$ or 4 vectors in $\mathbb{R}^8$ can be linearly dependent if they are all on the same straight line ($\mathbb{R}$), or if they are all in the same plane ($\mathbb{R}^2$), or if they are all in the same $\mathbb{R}^3$ ($3 < 4$!).

The simplest example of linearly dependent vectors is provided by two vectors $\mathbb{X}, \mathbb{Y} \in \mathbb{R}^n$, if they are on the same straight line, i.e. if they are parallel vectors, i.e. if $\mathbb{Y} = k \cdot \mathbb{X}, k \in \mathbb{R}$.

Given p vectors $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_p \in \mathbb{R}^n$, if one of them is the null vector $\mathbb{O}$, then the vectors will be anyhow linearly dependent; we give simply to each nonzero vector a coefficient equal to 0 while to $\mathbb{O}$ we give any nonzero coefficient; such a linear combination will obviously have the null vector $\mathbb{O}$ as its result.

SPANNING SETS AND BASES OF A VECTOR SPACE
It can be shown, but we only state it here, that the maximum number of linearly independent vectors in $\mathbb{R}^n$ is equal to the number n of their components, or that n vectors of $\mathbb{R}^n$ are at most linearly independent among them.
As a consequence, if we choose n linearly independent vectors $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_n \in \mathbb{R}^n$, any other vector $\mathbb{Y} \in \mathbb{R}^n$ can always be expressed by a suitable linear combination of some, even all, vectors $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_n$.

Let $\mathbb{V}$ be a vector subspace, $\mathbb{V} \subseteq \mathbb{R}^n$, and $\mathbb{W} = \{\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_m\}$ a set of vectors belonging to $\mathbb{V}$.
It's interesting to consider the set of all linear combinations of these vectors. This set is called the linear span of the vectors $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_m$.
We have the following
Definition 14 : $\mathbb{W}$ is said a spanning set for $\mathbb{V}$ if every $\mathbb{Y} \in \mathbb{V}$ can be written as a linear combination of the vectors $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_m$.
And also we have the following
Definition 15 : A set of vectors $\mathbb{W} = \{\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_m\}$ is said a basis for $\mathbb{V}$ if these vectors are linearly independent and if they span $\mathbb{V}$.
So, a basis is a linearly independent spanning set for a vector space.
A basis of $\mathbb{V}$ is then a set which consists of the minimum number of vectors that span $\mathbb{V}$.
Each vector $\mathbb{Y} \in \mathbb{V}$ can always be expressed as a linear combination of the vectors of a basis of $\mathbb{V}$.

The dimension of a vector space (or subspace) is the number of the elements of any of its bases, and it is equal to the maximum number of linearly independent vectors that can be determined in it.

Example 24 : Since the dimension of $\mathbb{R}$ is equal to 1, we simply need a vector (except the null one) to span the whole $\mathbb{R}$.
For example the vector 1 generates any number $k \in \mathbb{R}$: we simply multiply 1 by k.
We need two vectors, however, not on the same straight line, to generate any other vector of $\mathbb{R}^2$, the vector space whose dimension is equal to 2.
For example: $(x, y) = x\,(1,0) + y\,(0,1), \forall\,(x, y) \in \mathbb{R}^2$.
The set of polynomials, that of continuous functions and that of differentiable functions at x_0 are instead examples of infinite dimensional vector spaces.

The simplest example of a basis of $\mathbb{R}^n$ is the so-called standard (or canonical) basis, the one made up by vectors having one component equal to 1 and all the others equal to 0.

Example 25 : The standard basis of $\mathbb{R}^2$ is E_2: $\{\mathbf{e}_1 = (1,0); \mathbf{e}_2 = (0,1)\}$, the standard basis of $\mathbb{R}^3$ is E_3: $\{\mathbf{e}_1 = (1,0,0); \mathbf{e}_2 = (0,1,0); \mathbf{e}_3 = (0,0,1)\}$.

The following theorem is valid
Theorem 7 : Given a basis of $\mathbb{R}^n$, the representation of any vector $\mathbb{Y} \in \mathbb{R}^n$ by the vectors of such basis is a unique one.
Proof: Let us proceed by contradiction, supposing that a vector $\mathbb{Y} \in \mathbb{R}^n$ may have two representations using the basis $\{\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_n\}$ of $\mathbb{R}^n$, and so:
$\mathbb{Y} = \alpha_1 \cdot \mathbb{X}_1 + \alpha_2 \cdot \mathbb{X}_2 + ... + \alpha_n \cdot \mathbb{X}_n$ and also:
$\mathbb{Y} = \beta_1 \cdot \mathbb{X}_1 + \beta_2 \cdot \mathbb{X}_2 + ... + \beta_n \cdot \mathbb{X}_n$, with some $\alpha_i \neq \beta_i$.

Subtracting member to member we obtain:
$(\alpha_1-\beta_1)\cdot\mathbb{X}_1+(\alpha_2-\beta_2)\cdot\mathbb{X}_2\ +\ ...\ +\ (\alpha_n-\beta_n)\cdot\mathbb{X}_n=\mathbb{O}$.
If at least one of the differences $(\alpha_i-\beta_i)$ was different from zero, this would mean that the vectors $\mathbb{X}_1,\mathbb{X}_2,...,\mathbb{X}_n$ are linearly dependent, contrary to the hypothesis, and so $\alpha_i=\beta_i, \forall\, i$ and the representation of $\mathbb{Y}$ is therefore unique. •

SCALAR (or DOT or INNER) PRODUCT, MODULUS, EUCLIDEAN DISTANCE

Definition 16 : Given two vectors $\mathbb{X},\mathbb{Y}\in\mathbb{R}^n$, we define their scalar (or dot or inner) product, denoted with $\mathbb{X}\cdot\mathbb{Y}$, (or $<\mathbb{X},\mathbb{Y}>$) as the sum of the products of their components with the same index: $\mathbb{X}\cdot\mathbb{Y}=x_1\,y_1\ +\ x_2\,y_2\ +\ ..\ +\ x_n\,y_n\ =\sum_{i=1}^{n}x_i\,y_i$.

We note that $\mathbb{X}\cdot\mathbb{Y}$ gives as a result a real number (a scalar) and not a vector and this explains the name of "scalar" product.

Example 26 : If $\mathbb{X}=(3,\,-1,0)$ and $\mathbb{Y}=(5,2,\,-3)$, we obtain:
$\mathbb{X}\cdot\mathbb{Y}=3\cdot5+(-1)\cdot2+0\cdot(-3)=13$.

Definition 17 : We define the modulus (or length, sometimes norm) $\|\mathbb{X}\|$ of a vector $\mathbb{X}$ as the square root of the scalar product of the vector $\mathbb{X}$ by itself: $\|\mathbb{X}\|=\sqrt{\mathbb{X}\cdot\mathbb{X}}\ =\sqrt{\sum_{i=1}^{n}x_i^2}$.

A vector whose modulus (or length) is equal to 1 is called a unit vector (or versor or normalized vector).

Given a vector $\mathbb{X}\neq\mathbb{O}$, to obtain its unit vector $\mathbb{X}_{\text{v}}$ codirectional with $\mathbb{X}$ we have simply to calculate $\mathbb{X}_{\text{v}}=\dfrac{\mathbb{X}}{\|\mathbb{X}\|}=\dfrac{1}{\|\mathbb{X}\|}\cdot\mathbb{X}$.

Example 27 : If $\mathbb{X}=(3,\,-1,0)$, since $\|\mathbb{X}\|=\sqrt{3^2+(-1)^2+0^2}=\sqrt{10}$, to obtain its unit vector we calculate $\dfrac{1}{\sqrt{10}}\cdot\mathbb{X}=\left(\dfrac{3}{\sqrt{10}},\dfrac{-1}{\sqrt{10}},0\right)$.

All the unit vectors of $\mathbb{R}^2$ can be expressed in the form $v=(\cos\alpha,\sin\alpha)$.

Definition 18 : Given two vectors $\mathbb{X}=(x_1,x_2,...,x_n)$ and $\mathbb{Y}=(y_1,y_2,...,y_n)$, we define their Euclidean distance as the real nonnegative number:
$\|\mathbb{X}-\mathbb{Y}\|=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}$.

The Euclidean distance is the length of the line segment connecting $\mathbb{X}$ and $\mathbb{Y}$.
The modulus of a vector $\mathbb{X}$ is equal to the Euclidean distance between the vector $\mathbb{X}$ and the null vector $\mathbb{O}$.

If $n=2$, $\|\mathbb{X}-\mathbb{Y}\|=\sqrt{\sum_{i=1}^{2}(x_i-y_i)^2}\ =\sqrt{(x_1-y_1)^2+(x_2-y_2)^2}$ we find again the usual formula of the distance between the two points (x_1,x_2) and (y_1,y_2) in the real (cartesian) plane.

PROPERTIES OF THE SCALAR PRODUCT

The Cauchy–Schwarz inequality states that for all vectors $\mathbb{X}, \mathbb{Y} \in \mathbb{R}^n$ the following inequality is valid: $|\mathbb{X} \cdot \mathbb{Y}| \leq \|\mathbb{X}\| \cdot \|\mathbb{Y}\|$.
Such inequality is closely linked to the following equality:
$\mathbb{X} \cdot \mathbb{Y} = \|\mathbb{X}\| \cdot \|\mathbb{Y}\| \cdot \cos \alpha$, where α is the angle between the two vectors.
This formula can be rearranged to determine the size of the angle between two nonzero vectors, in fact we obtain: $\alpha = \arccos \left(\dfrac{\mathbb{X} \cdot \mathbb{Y}}{\|\mathbb{X}\| \cdot \|\mathbb{Y}\|} \right)$.
If $\alpha = \dfrac{\pi}{2}$ the scalar product is equal to 0: $\mathbb{X} \cdot \mathbb{Y} = 0$, and we say that the two vectors are orthogonal (or perpendicular).
If $0 < \alpha < \dfrac{\pi}{2}$ we obtain $\mathbb{X} \cdot \mathbb{Y} > 0$, while if $\dfrac{\pi}{2} < \alpha < \pi$ we obtain $\mathbb{X} \cdot \mathbb{Y} < 0$; and so if the angle between the two vectors is acute their scalar product is positive, otherwise it is negative.
If $\alpha = 0$, i.e. if the vectors are parallel and with the same direction, their scalar product gives the maximum possible result, while, if $\alpha = \pi$, i.e. if the vectors are parallel but with opposite directions, their scalar product gives the minimum possible result.

For vectors $\mathbb{X}, \mathbb{Y} \in \mathbb{R}^n$ another inequality is valid, the so-called:
- triangle inequality: $\|\mathbb{X} + \mathbb{Y}\| \leq \|\mathbb{X}\| + \|\mathbb{Y}\|$,
i.e. the modulus of a sum is less than or equal to the sum of the moduli.

MATRICES
The easiest way to introduce the concept of matrix is to define matrices as a rectangular array of real numbers, arranged in rows and columns.
The individual items in a matrix are called its elements or entries.
Also the matrices will be denoted by capital letters, and, for example, we write:

$$\mathbb{A}_{m,n} = \begin{Vmatrix} a_{11} & a_{12} & a_{13} & .... & .... & a_{1n} \\ a_{21} & a_{22} & a_{23} & .... & .... & a_{2n} \\ .... & .... & .... & .... & .... & .... \\ a_{i1} & a_{i2} & a_{i3} & .... & .... & a_{in} \\ .... & .... & .... & .... & .... & .... \\ a_{m1} & a_{m2} & a_{m3} & .... & .... & a_{mn} \end{Vmatrix}.$$

The entry in the i-th row and the j-th column of a matrix is typically denoted as a_{ij}.
The first of the two indices at the base of each entry is said the row index, the second is the column index, and we will say that the matrix $\mathbb{A}$ is a matrix $(m \cdot n)$ if it is formed by m rows and n columns.
In fact a matrix can also be defined as an ordered set of vectors, horizontally the rows and vertically the columns.
Writing $\mathbb{A}_{m,n} = [a_{ij}]$, we denote the matrix consisting of m rows and n columns whose generic entry in the place (i, j) is a_{ij}.
We will denote with $\mathrm{R}_1, \mathrm{R}_2, ..., \mathrm{R}_m$ its rows, each of which is a vector of $\mathbb{R}^n$, having a number of components equal to the number n of the columns of the matrix, and similarly we will denote with $\mathrm{C}_1, \mathrm{C}_2, ..., \mathrm{C}_n$ its columns, which are vectors of $\mathbb{R}^m$, so having a number of components equal to the number m of the rows of the matrix.
We will write $\mathbb{A} = \left[\mathrm{C}_1 \middle| \mathrm{C}_2 \middle| ... \middle| \mathrm{C}_n\right]$ to indicate the matrix A by means of its columns, while we will write, preferably vertically, $\mathbb{A} = \left[\mathrm{R}_1 \middle| \mathrm{R}_2 \middle| ... \middle| \mathrm{R}_m\right]$ to indicate the matrix A by means of its rows.

A matrix is said a square matrix if the number of its rows is equal to that of its columns (this number is called the order of the matrix and the matrix will be denoted by $\mathbb{A}_n$), otherwise it is called a rectangular matrix.

SPECIAL MATRICES

The vectors may be considered as particular matrices, m by 1, or $(m \cdot 1)$, if the matrix is a column vector; 1 by n, or $(1 \cdot n)$, if the matrix is a row vector.
A submatrix is a matrix formed by selecting certain rows and columns from a bigger matrix. Given a matrix $\mathbb{A}_{m,n}$, we say submatrix $(h \cdot k)$ of $\mathbb{A}$ the matrix obtained by taking the entries of $\mathbb{A}$ that are common to h rows and k columns and discarding all the others.
It is said null matrix, denoted by $\mathbb{O}$, a matrix whose entries are all equal to zero.

SPECIAL SQUARE MATRICES

The main diagonal (or leading diagonal) of a square matrix $\mathbb{A}$ is the set of the entries a_{ii}, i.e the entries having the same row and column indexes.

Definition 19 : A square matrix is called diagonal if the only non-zero entries are the entries that belong to the main diagonal, that is the entries which have the two indexes equal.
So a diagonal matrix is a matrix in which the entries outside the main diagonal are all equal to zero.
The diagonal entries themselves may or may not be equal to zero.

Definition 20 : A diagonal matrix is called a scalar matrix if the entries of the main diagonal are all equal: $a_{ii} = k, \forall i$.

Definition 21 : A square matrix is called upper triangular if all the entries below the main diagonal are equal to zero. Conversely, a square matrix is called lower triangular if all the entries above the main diagonal are equal to zero.
A triangular matrix is one that is either lower triangular or upper triangular.
A matrix that is both upper and lower triangular is a diagonal matrix.

Example 28 : The matrices $\mathbb{A} = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 7 \end{Vmatrix}$, $\mathbb{B} = \begin{Vmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{Vmatrix}$, $\mathbb{C} = \begin{Vmatrix} 2 & 0 & 3 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{Vmatrix}$ are, respectively, a diagonal matrix, a scalar matrix and an upper triangular matrix.

Definition 22 : A square matrix is called a symmetric matrix if $a_{ij} = a_{ji}$.
The entries of a symmetric matrix are symmetric with respect to the main diagonal.

Example 29 : The matrix $\mathbb{A} = \begin{Vmatrix} 1 & 2 & -3 \\ 2 & 5 & 6 \\ -3 & 6 & 4 \end{Vmatrix}$ is a symmetric matrix.

Definition 23 : The identity matrix (or unit matrix) of order n, denoted with $\mathbb{I}_n$, is a scalar matrix having all the entries of the main diagonal equal to 1.

Example 30 : $\mathbb{I}_2 = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}$, $\mathbb{I}_3 = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}$.

Definition 24 : A permutation matrix is a matrix obtained interchanging some rows (or some columns) of the identity matrix.
So a permutation matrix is a square matrix that has exactly one entry equal to 1 in each row and each column and 0 elsewhere.

Example 31 : $\mathbb{P}_1 = \begin{Vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{Vmatrix}$ and $\mathbb{P}_2 = \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{Vmatrix}$ are permutation matrices.

BASIC MATRIX OPERATIONS
The main matrix operations are nothing but an extension of similar operations defined for vectors.

Definition 25 : Given $\mathbb{A}_{m,n} = [a_{ij}]$ and $\mathbb{B}_{m,n} = [b_{ij}]$, both with m rows and n columns, we define their sum matrix as the matrix, which is also $(m \cdot n)$, having as the entry of indices (i, j) the sum of the entries of indices (i, j) of the given matrices:
$\mathbb{C}_{m,n} = [c_{ij}] = [a_{ij} + b_{ij}]$.
The sum of two matrices can be calculated only if the two matrices have the same number of rows and columns.

Definition 26 : The scalar multiplication of a matrix $\mathbb{A}_{m,n}$ by a number k (also called a scalar) is given by multiplying every entry of $\mathbb{A}$ by k:
$k \cdot \mathbb{A}_{m,n} = (k \cdot \mathbb{A})_{m,n} = [k \cdot a_{ij}]$.
We can multiply any matrix by any scalar.

Definition 27 : A linear combination of matrices, all however having m rows and n columns, with given scalar coefficients, is defined as the matrix having as entry with indexes (i, j) the linear combination, with the same coefficients, of the entries with indexes (i, j) of the given matrices.
In the case of only two matrices, $\mathbb{A}_{m,n} = [a_{ij}]$ and $\mathbb{B}_{m,n} = [b_{ij}]$, we get:
$\alpha \cdot \mathbb{A}_{m,n} + \beta \cdot \mathbb{B}_{m,n} = [\alpha \cdot a_{ij} + \beta \cdot b_{ij}]$.

Example 32 : If $\mathbb{A} = \begin{Vmatrix} 1 & 3 & -4 \\ 0 & 1 & 2 \\ 1 & -2 & 5 \end{Vmatrix}$ and $\mathbb{B} = \begin{Vmatrix} 1 & 0 & 0 \\ 3 & -1 & -2 \\ 0 & 3 & 2 \end{Vmatrix}$, we obtain:

$$\mathbb{C} = 3 \cdot \mathbb{A} + 2 \cdot \mathbb{B} = \begin{Vmatrix} 3+2 & 9+0 & -12+0 \\ 0+6 & 3-2 & 6-4 \\ 3+0 & -6+6 & 15+4 \end{Vmatrix} = \begin{Vmatrix} 5 & 9 & -12 \\ 6 & 1 & 2 \\ 3 & 0 & 19 \end{Vmatrix}.$$

Definition 28 : The transpose of the $(m \cdot n)$ matrix $\mathbb{A}_{m,n}$ is the $(n \cdot m)$ matrix $\mathbb{A}^{\mathrm{T}}_{n,m}$, having as entry with indexes (i, j) the entry with indexes (j, i) of the given matrix, i.e. the matrix formed turning rows into columns and vice versa.

Example 33 : If $\mathbb{A}_{3,4} = \begin{Vmatrix} 1 & 3 & -4 & 2 \\ 0 & 1 & 2 & 7 \\ 1 & -2 & 5 & 6 \end{Vmatrix}$, then $\mathbb{A}^{\mathrm{T}}_{4,3} = \begin{Vmatrix} 1 & 0 & 1 \\ 3 & 1 & -2 \\ -4 & 2 & 5 \\ 2 & 7 & 6 \end{Vmatrix}$.

Properties of the Transpose
T1) $\mathbb{A} = \left(\mathbb{A}^{\mathrm{T}}\right)^{\mathrm{T}}$, i.e. the Transpose of the Transpose of a matrix $\mathbb{A}$ is the given matrix $\mathbb{A}$;
T2) $(\mathbb{A}+\mathbb{B})^{\mathrm{T}} = \mathbb{A}^{\mathrm{T}} + \mathbb{B}^{\mathrm{T}}$;
T3) $\mathbb{A} = \mathbb{A}^{\mathrm{T}}$ if and only if the matrix $\mathbb{A}$ is a symmetric matrix.

MATRIX PRODUCTS
There are many products that can be defined between two matrices. We will deal only with the so-called "rows by columns" product which, among other properties, satisfies the associative property. It is also called "matrix-multiplicative product".
Then we will deal with the "Kronecker product", an operation on two matrices of arbitrary size resulting in a block matrix.

"ROWS BY COLUMNS" PRODUCT BETWEEN MATRICES
The "rows by columns" product between two matrices is based on the scalar product of two vectors, in this case the rows of the first matrix by the columns of the second matrix.
Then two matrices will be multipliable "rows by columns" only if the row vectors of the first matrix and the column vectors of the second have the same number of entries, and this happens when the number of the columns of the first matrix is equal to the number of the rows of the second.

Definition 29 : Given two matrices $\mathbb{A}_{m,n} = [a_{ij}]$, $1 \le i \le m$, $1 \le j \le n$, and $\mathbb{B}_{n,p} = [b_{ij}]$, $1 \le i \le n$, $1 \le j \le p$, we define their "rows by columns" product $\mathbb{C}_{m,p} = \mathbb{A}_{m,n} \cdot \mathbb{B}_{n,p}$ as the matrix having as many rows, m, as those of the first matrix and as many columns, p, as those of the second one, whose entry with indexes (i, j), c_{ij}, is given by the scalar product of two vectors: $\mathrm{R}_i^{\mathbb{A}}$, the i-th row of the matrix $\mathbb{A}$ by $\mathrm{C}_j^{\mathbb{B}}$, j-th column of the matrix $\mathbb{B}$, i.e.:

$$c_{ij} = \mathrm{R}_i^{\mathbb{A}} \cdot \mathrm{C}_j^{\mathbb{B}} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj} .$$

This matrix product satisfies the associative property: $\mathbb{A} \cdot (\mathbb{B} \cdot \mathbb{C}) = (\mathbb{A} \cdot \mathbb{B}) \cdot \mathbb{C}$, while it does not satisfy the commutative property: in general $\mathbb{A} \cdot \mathbb{B}$ and $\mathbb{B} \cdot \mathbb{A}$ are different matrices; equality might be fulfilled if the two matrices were both square and of the same order, but there are examples of how, also in this case, it is not worth, in general, the commutative property.

Example 34 : Given the two matrices $\mathbb{A}_{2,2} = \begin{Vmatrix} 3 & -2 \\ 0 & 5 \end{Vmatrix}$ and $\mathbb{B}_{2,2} = \begin{Vmatrix} 1 & -4 \\ -3 & 7 \end{Vmatrix}$, the result is:

$$\mathbb{A}_{2,2} \cdot \mathbb{B}_{2,2} = \begin{Vmatrix} 3 \cdot 1 + (-2) \cdot (-3) & 3 \cdot (-4) + (-2) \cdot 7 \\ 0 \cdot 1 + 5 \cdot (-3) & 0 \cdot (-4) + 5 \cdot 7 \end{Vmatrix} = \begin{Vmatrix} 9 & -26 \\ -15 & 35 \end{Vmatrix} \text{ while}$$

$$\mathbb{B}_{2,2} \cdot \mathbb{A}_{2,2} = \begin{Vmatrix} 1 \cdot 3 + (-4) \cdot 0 & 1 \cdot (-2) + (-4) \cdot 5 \\ (-3) \cdot 3 + 7 \cdot 0 & (-3) \cdot (-2) + 7 \cdot 5 \end{Vmatrix} = \begin{Vmatrix} 3 & -22 \\ -9 & 41 \end{Vmatrix} .$$

After calculating the two products it is thus evident that $\mathbb{A} \cdot \mathbb{B} \neq \mathbb{B} \cdot \mathbb{A}$.

If we consider a matrix $\mathbb{A}_{m,n}$ and the unit matrices $\mathbb{I}_m$ and $\mathbb{I}_n$, the following are valid: $\mathbb{A}_{m,n} \cdot \mathbb{I}_n = \mathbb{A}_{m,n} = \mathbb{I}_m \cdot \mathbb{A}_{m,n}$.

We should also note that, unlike multiplication of real numbers, the product of two matrices can give as a result the null matrix $\mathbb{O}$ even if none of the two matrix is a null matrix. So this matrix product does not obey the product cancellation law.

Example 35 : If $\mathbb{A} = \begin{Vmatrix} 1 & -1 & 1 \\ -3 & 2 & -1 \\ -2 & 1 & 0 \end{Vmatrix}$ and $\mathbb{B} = \begin{Vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{Vmatrix}$, performing their "rows by columns" product, we obtain $\mathbb{A} \cdot \mathbb{B} = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$ while $\mathbb{B} \cdot \mathbb{A} = \begin{Vmatrix} -11 & 6 & -1 \\ -22 & 12 & -2 \\ -11 & 6 & -1 \end{Vmatrix}$.

For the "rows by columns" product the following properties hold:
P1) Associative: $\mathbb{A} \cdot \mathbb{B} \cdot \mathbb{C} = (\mathbb{A} \cdot \mathbb{B}) \cdot \mathbb{C} = \mathbb{A} \cdot (\mathbb{B} \cdot \mathbb{C})$;
P2) Distributive over matrix addition:
$\mathbb{A} \cdot (\mathbb{B} + \mathbb{C}) = \mathbb{A} \cdot \mathbb{B} + \mathbb{A} \cdot \mathbb{C}$ and $(\mathbb{B} + \mathbb{C}) \cdot \mathbb{A} = \mathbb{B} \cdot \mathbb{A} + \mathbb{C} \cdot \mathbb{A}$;
P3) Scalar multiplication: $(k \cdot \mathbb{A}) \cdot \mathbb{B} = k \cdot (\mathbb{A} \cdot \mathbb{B})$;
P4) Transpose: $(\mathbb{A} \cdot \mathbb{B})^{\mathrm{T}} = \mathbb{B}^{\mathrm{T}} \cdot \mathbb{A}^{\mathrm{T}}$, i.e. the transpose of a product is equal to the product of the transposes, but in the opposite order.

Square matrices can be multiplied by themselves repeatedly in the same way as real numbers, because such product always gives as result matrices having the same number of rows and columns. This repeated multiplication can be described as a power of the matrix. This is not possible for rectangular matrices.
If $\mathbb{A}$ is a square matrix, we write: $\mathbb{A}^2 = \mathbb{A} \cdot \mathbb{A}$. Similarly we denote by $\mathbb{A}^k$ the product, k times, of the matrix $\mathbb{A}$ by itself , to obtain the k-th power of the matrix $\mathbb{A}$.

The transpose of a power of a matrix is equal to the power of the transpose: $\left(\mathbb{A}^k\right)^{\mathrm{T}} = \left(\mathbb{A}^{\mathrm{T}}\right)^k$.
In fact: $\left(\mathbb{A}^k\right)^{\mathrm{T}} = (\mathbb{A} \cdot \mathbb{A} \cdot ... \cdot \mathbb{A} \cdot \mathbb{A})^{\mathrm{T}} = \mathbb{A}^{\mathrm{T}} \cdot \mathbb{A}^{\mathrm{T}} \cdot ... \cdot \mathbb{A}^{\mathrm{T}} \cdot \mathbb{A}^{\mathrm{T}} = \left(\mathbb{A}^{\mathrm{T}}\right)^k$.

If $\mathbb{B} = \mathbb{A}^2$, we usually also write: $\mathbb{A} = \mathbb{B}^{\frac{1}{2}}$.

Example 36 : If $\mathbb{A} = \begin{Vmatrix} 1 & -1 \\ 2 & 1 \end{Vmatrix}$, we obtain $\mathbb{A}^2 = \mathbb{A} \cdot \mathbb{A} = \begin{Vmatrix} -1 & -2 \\ 4 & -1 \end{Vmatrix} = \mathbb{B}$ and so:
$\begin{Vmatrix} -1 & -2 \\ 4 & -1 \end{Vmatrix} = \begin{Vmatrix} 1 & -1 \\ 2 & 1 \end{Vmatrix}^2$ and $\begin{Vmatrix} 1 & -1 \\ 2 & 1 \end{Vmatrix} = \begin{Vmatrix} -1 & -2 \\ 4 & -1 \end{Vmatrix}^{\frac{1}{2}}$.

An idempotent matrix is a square matrix such that, when multiplied by itself, it yields itself: $\mathbb{A}^2 = \mathbb{A}$.
A nilpotent matrix is a square matrix $\mathbb{A}$ such that: $\mathbb{A}^k = \mathbb{O}$. k is called the index (or degree) of $\mathbb{A}$.

KRONECKER PRODUCT ON MATRICES

There are several other types of matrices products. One of these, much used in statistics, is the so-called Kronecker product, which is indicated with the symbol $\otimes$.

Definition 30 : Let us consider two matrices: $\mathbb{A}_{m,n} = [a_{ij}], 1 \le i \le m, 1 \le j \le n$, and $\mathbb{B}_{p,q} = [b_{ij}], 1 \le i \le p, 1 \le j \le q$.

We define the matrix obtained by the Kronecker product $\mathbb{K}_{m\cdot p, n\cdot q} = \mathbb{A}_{m,n} \otimes \mathbb{B}_{p,q}$, as a matrix, having $m \cdot p$ rows (the product between the number of rows of $\mathbb{A}$ and those of $\mathbb{B}$) and $n \cdot q$ columns (the product between the number of columns of $\mathbb{A}$ and those of $\mathbb{B}$), obtained in the following way: in the place of its entry k_{ij} we put the matrix given by the scalar product $a_{ij} \cdot \mathbb{B}$.

Example 37 : Given the two matrices $\mathbb{A}_{2,2} = \begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}$ and $\mathbb{B}_{2,3} = \begin{Vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{Vmatrix}$, we obtain: $\mathbb{A}_{2,2} \otimes \mathbb{B}_{2,3} = \mathbb{K}_{2\cdot 2, 2\cdot 3} = \mathbb{K}_{4,6} = \begin{Vmatrix} a_{11} \cdot \mathbb{B} & a_{12} \cdot \mathbb{B} \\ a_{21} \cdot \mathbb{B} & a_{22} \cdot \mathbb{B} \end{Vmatrix} =$

$$= \begin{Vmatrix} a_{11} \cdot \begin{Vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{Vmatrix} & a_{12} \cdot \begin{Vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{Vmatrix} \\ a_{21} \cdot \begin{Vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{Vmatrix} & a_{22} \cdot \begin{Vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{Vmatrix} \end{Vmatrix} =$$

$$= \begin{Vmatrix} a_{11} \cdot b_{11} & a_{11} \cdot b_{12} & a_{11} \cdot b_{13} & a_{12} \cdot b_{11} & a_{12} \cdot b_{12} & a_{12} \cdot b_{13} \\ a_{11} \cdot b_{21} & a_{11} \cdot b_{22} & a_{11} \cdot b_{23} & a_{12} \cdot b_{21} & a_{12} \cdot b_{22} & a_{12} \cdot b_{23} \\ a_{21} \cdot b_{11} & a_{21} \cdot b_{12} & a_{21} \cdot b_{13} & a_{22} \cdot b_{11} & a_{22} \cdot b_{12} & a_{22} \cdot b_{13} \\ a_{21} \cdot b_{21} & a_{21} \cdot b_{22} & a_{21} \cdot b_{23} & a_{22} \cdot b_{21} & a_{22} \cdot b_{22} & a_{22} \cdot b_{23} \end{Vmatrix}.$$

Example 38 : Given the matrices $\mathbb{A}_{2,2} = \begin{Vmatrix} 1 & 2 \\ 3 & 4 \end{Vmatrix}$ and $\mathbb{B}_{3,4} = \begin{Vmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{Vmatrix}$, we obtain:

$$\mathbb{A}_{2,2} \otimes \mathbb{B}_{3,4} = \mathbb{K}_{2\cdot 3, 2\cdot 4} = \mathbb{K}_{6,8} = \begin{Vmatrix} 1 \cdot \mathbb{B} & 2 \cdot \mathbb{B} \\ 3 \cdot \mathbb{B} & 4 \cdot \mathbb{B} \end{Vmatrix} =$$

$$= \begin{Vmatrix} 1 \cdot \begin{Vmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{Vmatrix} & 2 \cdot \begin{Vmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{Vmatrix} \\ 3 \cdot \begin{Vmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{Vmatrix} & 4 \cdot \begin{Vmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{Vmatrix} \end{Vmatrix} =$$

$$= \begin{Vmatrix} 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 \\ 3 & 3 & 3 & 3 & 6 & 6 & 6 & 6 \\ 3 & 3 & 3 & 3 & 4 & 4 & 4 & 4 \\ 6 & 6 & 6 & 6 & 8 & 8 & 8 & 8 \\ 9 & 9 & 9 & 9 & 12 & 12 & 12 & 12 \end{Vmatrix}.$$

There are no restrictions on the implementability of the Kronecker product; it is possible to multiply any two matrices together regardless of their size.

We can easily see that $\mathbb{A} \otimes \mathbb{O} = \mathbb{O} \otimes \mathbb{A} = \mathbb{O}$.

For the Kronecker product the cancellation law is valid, i.e.:

$\mathbb{A} \otimes \mathbb{B} = \mathbb{O} \Rightarrow \mathbb{A} = \mathbb{O}$ and/or $\mathbb{B} = \mathbb{O}$.

The Kronecker product does not satisfy the commutative property: in general $\mathbb{A} \otimes \mathbb{B}$ and $\mathbb{B} \otimes \mathbb{A}$ are different matrices.

The Kronecker product satisfies the associative property:
$\mathbb{A} \otimes \mathbb{B} \otimes \mathbb{C} = \mathbb{A} \otimes (\mathbb{B} \otimes \mathbb{C}) = (\mathbb{A} \otimes \mathbb{B}) \otimes \mathbb{C}, \forall\, \mathbb{A}, \mathbb{B}$ and $\mathbb{C}$.
The Kronecker product satisfies the distributive property: $(\mathbb{A} + \mathbb{B}) \otimes \mathbb{C} = \mathbb{A} \otimes \mathbb{C} + \mathbb{B} \otimes \mathbb{C}$.
For the transpose matrix, unlike the "rows by columns" product, the following property holds: $(\mathbb{A} \otimes \mathbb{B})^{\mathrm{T}} = \mathbb{A}^{\mathrm{T}} \otimes \mathbb{B}^{\mathrm{T}}$.

ELEMENTARY OPERATIONS ON THE LINES OF A MATRIX

Elementary operations on the lines of a matrix are ways to change the matrices. There are three types of line operations: line switching, that is interchanging two lines of a matrix; line multiplication, that is multiplying all the entries of a line by a non-zero constant; and finally line addition, which means replacing a line with the addition of the line itself to a linear combination of other lines.
These line operations can be very useful to compute the Determinant and the Rank of a matrix, and also to solve linear equations and to find inverses.
We will use such elementary operations on the lines (rows and/or columns) of a matrix:
1) interchanging two rows or columns: $\mathrm{L}_i \leftrightarrows \mathrm{L}_j$
2) multiplying the entries of a line by a constant $k \neq 0$: $\mathrm{L}_i \leftarrow k \cdot \mathrm{L}_i$;
3) replacing a line with the addition of the line itself to a linear combination of other lines: $\mathrm{L}_i \leftarrow \mathrm{L}_i + \sum \alpha_j \cdot \mathrm{L}_j \ (j \neq i)$.

The elementary operations can also be obtained multiplying the given matrix $\mathbb{A}$ by a suitable matrix $\mathbb{E}$, called "elementary matrix".
The $\mathbb{A} \cdot \mathbb{E}$ product coincides with elementary operations on columns, the product $\mathbb{E} \cdot \mathbb{A}$ with elementary operations on rows.

Example 39 : If $\mathbb{A} = \begin{Vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{Vmatrix}$, using the permutation matrix $\mathbb{E} = \begin{Vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{Vmatrix}$, we obtain: $\begin{Vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{Vmatrix} \cdot \begin{Vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{Vmatrix} = \begin{Vmatrix} 2 & 1 & 3 \\ 5 & 4 & 6 \\ 8 & 7 & 9 \end{Vmatrix}$, so we have switched the first column with the second column;
we obtain $\begin{Vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{Vmatrix} \cdot \begin{Vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{Vmatrix} = \begin{Vmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{Vmatrix}$, so we have switched the first row with the second row.

Example 40 : If $\mathbb{A} = \begin{Vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{Vmatrix}$, using the matrix $\mathbb{E} = \begin{Vmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{Vmatrix}$, we obtain:
$\begin{Vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{Vmatrix} \cdot \begin{Vmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{Vmatrix} = \begin{Vmatrix} 4 & 1 & 1 \\ 10 & 1 & 1 \\ 16 & 1 & 1 \end{Vmatrix}$, so we have replaced:
- the first column with the sum of the first column with the third column;
- the second column with the difference between the second column and the first column;
- the third column with the difference between the third column and the second column.

We obtain $\begin{Vmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{Vmatrix} \cdot \begin{Vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{Vmatrix} = \begin{Vmatrix} -3 & -3 & -3 \\ -3 & -3 & -3 \\ 8 & 10 & 12 \end{Vmatrix}$, so we have replaced:
- the first row with the difference between the first row and the second row;
- the second row with the difference between the second row and the third row;
- the third row with the sum of the third row with the first row.

Example 41 : If $\mathbb{A} = \begin{Vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{Vmatrix}$, using the matrix $\mathbb{E} = \begin{Vmatrix} 1 & 2 & 0 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \end{Vmatrix}$, we obtain:
$\begin{Vmatrix} 1 & 2 & 0 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \end{Vmatrix} \cdot \begin{Vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{Vmatrix} = \begin{Vmatrix} 9 & 12 & 15 \\ -10 & -11 & -12 \\ 8 & 10 & 12 \end{Vmatrix}$, so we have replaced:
- the first row with the sum of the first row with the double of the second row;
- the second row with the difference between the second row and the double of the third row;
- the third row with the sum of the third row with the first row.

THE DETERMINANT
From now on we will consider only square matrices $\mathbb{A}_n$. Following first the traditional form, we give the
Definition 31 : We define the Determinant of a square matrix $\mathbb{A}_n$, denoted with $\det(\mathbb{A})$ or $|\mathbb{A}|$, as the sum of the $n\,!$ products of entries of the matrix where:
- each product has n terms, containing one and only one entry for each row and each column;
- each of these products is taken with its own sign or changing its sign depending on whether the permutation of the first indices of the entries of the product is or not of the same class of the permutation of the second indices.
Permutations can be of even or odd class; the sequence of the first indexes and the sequence of the second indexes of the entries of each product have to be brought to the original sequence $1, 2, 3, ..., n$; the permutation is of even (or odd) class if the original sequence can be obtained by an even (or odd) number of switches of the indexes.

To illustrate this, let us consider the matrix $\mathbb{A}_3 = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}$.
The 3! possible products, each of 3 entries being not on the same row or column, are the following:
1) $a_{11}\,a_{22}\,a_{33}$; first indexes $1, 2, 3$; second indexes $1, 2, 3$;
2) $a_{11}\,a_{23}\,a_{32}$; first indexes $1, 2, 3$; second indexes $1, 3, 2$;
3) $a_{12}\,a_{23}\,a_{31}$; first indexes $1, 2, 3$; second indexes $2, 3, 1$;
4) $a_{12}\,a_{21}\,a_{33}$; first indexes $1, 2, 3$; second indexes $2, 1, 3$;
5) $a_{13}\,a_{21}\,a_{32}$; first indexes $1, 2, 3$; second indexes $3, 1, 2$;
6) $a_{13}\,a_{22}\,a_{31}$; first indexes $1, 2, 3$; second indexes $3, 2, 1$.
In the six products that we have formed the first group of indices is always the same: $1, 2, 3$, so we need 0 (even number) switches to bring it to the original sequence $1, 2, 3$.
Let us consider now the second indexes.
For products $1)$, $3)$ and $5)$ we need respectively 0, 2 and 2 switches, i.e. an even number like 0; so these products are to be taken with their own sign.

For products $2)$, $4)$ and $6)$ we need respectively 1, 1 and 3 switches, i.e. an odd number; so these products are to be taken changing their own sign.
The Determinant of $\mathbb{A}$ is then equal to:
$|\mathbb{A}| = a_{11}\, a_{22}\, a_{33} - a_{11}\, a_{23}\, a_{32} + a_{12}\, a_{23}\, a_{31} - a_{12}\, a_{21}\, a_{33} + a_{13}\, a_{21}\, a_{32} - a_{13}\, a_{22}\, a_{31}$.

At this point we should list all the properties which follow from the definition of Determinant. We can list more rapidly these results presenting a more modern definition for Determinant, which is the following:
Definition 32 (axiomatic definition for the Determinant): The Determinant $|\mathbb{A}|$ of a square matrix $\mathbb{A}_n$ is a multilinear and alternating function of the rows and/or columns (i.e. of the lines) of the matrix, that associates a real number to $\mathbb{A}_n$, and such that $\det(\mathbb{I}_n) = 1$.
So we will write: $|\mathbb{A}| = f(\mathrm{R}_1, \mathrm{R}_2, ..., \mathrm{R}_n) = f(\mathrm{C}_1, \mathrm{C}_2, ..., \mathrm{C}_n)$.

Let us see the latter definition more precisely.
When we say that $f(\mathbb{L}_1, ..., \mathbb{L}_n)$ is a multilinear function we mean that:
$f(\mathbb{L}_1, ..., \mathbb{L}_{k-1}, \alpha\,\mathbb{X} + \beta\,\mathbb{Y}, \mathbb{L}_{k+1}, ..., \mathbb{L}_n) =$
$= \alpha\, f(\mathbb{L}_1, ..., \mathbb{L}_{k-1}, \mathbb{X}, \mathbb{L}_{k+1}, ..., \mathbb{L}_n) + \beta\, f(\mathbb{L}_1, ..., \mathbb{L}_{k-1}, \mathbb{Y}, \mathbb{L}_{k+1}, ..., \mathbb{L}_n)$,
that is, whatever line we have used, the image of a linear combination is equal to the linear combination of the images.
When we say that $f(\mathbb{L}_1, ..., \mathbb{L}_n)$ is an alternating function we mean that:
$f(\mathbb{L}_1, ..., \mathbb{X}, ..., \mathbb{Y}, ..., \mathbb{L}_n) = -f(\mathbb{L}_1, ..., \mathbb{Y}, ..., \mathbb{X}, ..., \mathbb{L}_n)$,
that is, swapping the position of two variables (lines), the value of the function changes its sign.

This second definition allows us to list quickly the main properties of the Determinant:
P1) - The Determinant changes its sign if we interchange two lines of the matrix (alternating property).

Example 42 : $|\mathbb{A}_3| = \begin{vmatrix} 1 & 2 & 6 \\ 0 & 3 & 5 \\ 4 & 8 & 7 \end{vmatrix} = - \begin{vmatrix} 0 & 3 & 5 \\ 1 & 2 & 6 \\ 4 & 8 & 7 \end{vmatrix} = \begin{vmatrix} 5 & 3 & 0 \\ 6 & 2 & 1 \\ 7 & 8 & 4 \end{vmatrix}$.

P2)- The Determinant of a matrix $\mathbb{A}_n$ is equal to that of its Transpose $\mathbb{A}_n^{\mathrm{T}}$.

Example 43 : $\begin{vmatrix} 1 & 2 & 0 \\ 5 & 8 & 3 \\ 4 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 5 & 4 \\ 2 & 8 & 2 \\ 0 & 3 & 0 \end{vmatrix}$.

As a consequence of the multilinear and alternating properties, the following fundamental property is valid:
P3) - The Determinant of a matrix is equal to 0 if and only if the rows (and the columns) of the matrix are linearly dependent vectors.

A matrix $\mathbb{A}$ whose Determinant is different from 0 is called a non-singular matrix, otherwise, if $|\mathbb{A}| = 0$, it is called a singular matrix.

From property P3) other properties follow:
P4) - If all the entries of a line of the matrix are equal to 0, the Determinant is equal to 0.

P5) - If two lines are proportional (i.e. the entries of a line are multiple, for the same scalar, of the entries of the other), the Determinant is equal to 0.

Example 44 : $\begin{vmatrix} 1 & 3 & 2 \\ 3 & 9 & 6 \\ 0 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 2 \\ 3\cdot 1 & 3\cdot 3 & 3\cdot 2 \\ 0 & 4 & 1 \end{vmatrix} = 0\,.$

P6) - If we multiply all the entries of a line of the matrix $\mathbb{A}$ by a scalar k, the Determinant of this new matrix is equal to the Determinant of $\mathbb{A}$ multiplied by k.

Example 45 : $\begin{vmatrix} 1 & 3 & 2 \\ 3 & 6 & 0 \\ 0 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 2 \\ 3\cdot 1 & 3\cdot 2 & 3\cdot 0 \\ 0 & 4 & 1 \end{vmatrix} = 3\cdot \begin{vmatrix} 1 & 3 & 2 \\ 1 & 2 & 0 \\ 0 & 4 & 1 \end{vmatrix}\,.$

P7) - If a line $\mathbb{L}$ of a matrix $\mathbb{A}$ can be expressed as the sum of two or more lines, the Determinant of $\mathbb{A}$ is equal to the sum of two or more Determinants, each having the same lines of $\mathbb{A}$, except the line $\mathbb{L}$, instead of which we have to put, one at a time, the various addends of $\mathbb{L}$ (multilinearity property).

Example 46 : $\begin{vmatrix} 1 & 3 & 2 \\ 2+5 & 1+3 & 0+3 \\ 0 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 2 \\ 2 & 1 & 0 \\ 0 & 4 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 3 & 2 \\ 5 & 3 & 3 \\ 0 & 4 & 1 \end{vmatrix}\,.$

P8) - The value of the Determinant of a matrix does not change if a line is replaced by any linear combination of the line itself with other lines of the matrix.
This property is very important from the point of view of the practical calculus of the Determinants, since, applying it suitably, it allows us to generate lines that contain the largest possible number of zero entries, greatly reducing the calculations required to find the value of the Determinant.

Example 47 : Adding to the first line the second multiplied by 4 and the third multiplied by (-3), $(\mathrm{R}_1 \leftarrow \mathrm{R}_1 + 4\,\mathrm{R}_2 - 3\,\mathrm{R}_3)$, we obtain:

$$\begin{vmatrix} 3 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 3 \end{vmatrix} = \begin{vmatrix} 3+4\cdot 0-3\cdot 1 & 1+4\cdot 2-3\cdot 1 & 2+4\cdot 1-3\cdot 3 \\ 0 & 2 & 1 \\ 1 & 1 & 3 \end{vmatrix} = \begin{vmatrix} 0 & 6 & -3 \\ 0 & 2 & 1 \\ 1 & 1 & 3 \end{vmatrix},$$

i.e. the value of the Determinant does not change.

Example 48 : Given the three vectors $\mathbb{X} = (1,2,4)$, $\mathbb{Y} = (-1,1,-1)$ and $\mathbb{Z} = (1,5,7)$, we want to check whether they are linearly independent or dependent. We use property P3); so we construct the matrix $\mathbb{A}_3$ having $\mathbb{X}$, $\mathbb{Y}$ and $\mathbb{Z}$ as rows.
Adding the third row to the second row, and subtracting the first row to the third row $(\mathrm{R}_2 \leftarrow \mathrm{R}_2 + \mathrm{R}_3$ and $\mathrm{R}_3 \leftarrow \mathrm{R}_3 - \mathrm{R}_1)$ we obtain:

$|\mathbb{A}_3| = \begin{vmatrix} 1 & 2 & 4 \\ -1 & 1 & -1 \\ 1 & 5 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 4 \\ 0 & 6 & 6 \\ 0 & 3 & 3 \end{vmatrix} = 0$, since the second and the third line are proportional. Then the three vectors are linearly dependent vectors.

LAPLACE' THEOREMS (or LAPLACE' EXPANSION)

To calculate practically the value of a Determinant, in a better way than using the definition, we can use the so-called Laplace' first Theorem; to state this theorem we need some definitions.

Definition 33 : Given a square matrix $\mathbb{A}_n$, for every entry a_{ij} the Minor M_{ij} of a_{ij} is defined to be the Determinant of the $((n-1)\cdot(n-1))$-submatrix that results from $\mathbb{A}_n$ removing the i-th row and the j-th column, i.e. the lines to which the entry belongs.

Definition 34 : Given a square matrix $\mathbb{A}_n$, for every entry a_{ij} the Cofactor A_{ij} of a_{ij} is defined as $\mathrm{A}_{ij} = (-1)^{i+j}\cdot \mathrm{M}_{ij}$, i.e. the Minor M_{ij} with its own sign if the sum $i+j$ of the indexes of the entry is even; otherwise, if the sum $i+j$ of the indexes of the entry is odd, we must change its sign.

Then the following applies:

Theorem (I° Laplace' Theorem): The Determinant of any square matrix $\mathbb{A}_n$ is obtained by the sum of the products of the entries of any line of the matrix by their own Cofactors.

For example, if we calculate the Determinant through the i-th row, we obtain:

$$|\mathbb{A}_n| = \det(\mathbb{A}_n) = a_{i1}\,\mathrm{A}_{i1} + a_{i2}\,\mathrm{A}_{i2} + \ldots + a_{ik}\,\mathrm{A}_{ik} + \ldots + a_{in}\,\mathrm{A}_{in} = \sum_{k=1}^{n} a_{ik}\,\mathrm{A}_{ik}\,.$$

Laplace' theorem then allows to calculate the Determinant of any square matrix of order n by calculating n Determinants of matrices of order $n-1$, which are to be calculated by means of $n-1$ Determinants of order $n-2$, and so on until obtaining Determinants of matrices of order 2.

Let us begin calculating the Determinant of the matrices of the lower orders.

For a $(1\cdot 1)$ matrix, as $\mathbb{A}_{1,1} = a_{1,1}$, we have $|\mathbb{A}_{1,1}| = a_{1,1}$.

Then we calculate the Determinant of a matrix of order 2.

Given the matrix $\mathbb{A} = \begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}$, we obtain: $\det(\mathbb{A}) = a_{11}\,a_{22} - a_{12}\,a_{21}$.

We consider now the matrix of order 3: $\mathbb{A} = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}$.

Developing the Determinant, for example, for the entries of the first row, we will have:

$|\mathbb{A}| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$, from which we obtain:

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

Example 49 : Let us calculate the Determinant of the matrix $\mathbb{A} = \begin{Vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{Vmatrix}$.

Using Laplace' theorem, developing for the entries of the first row, we obtain:

$$|\mathbb{A}| = 1\cdot\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} + (-2)\cdot\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3\cdot\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} =$$
$$= (45-48) - 2\,(36-42) + 3\,(32-35) = -3 + 12 - 9 = 0\,.$$

Now let us calculate the same Determinant using the property P8).

Using elementary operations on the rows, we replace the second row with the row itself minus the first row multiplied by 4: $R_2 \leftarrow R_2 - 4R_1$, and then we replace the third row with the row itself minus the first row multiplied by 7: $R_3 \leftarrow R_3 - 7R_1$ and so we obtain:

$$|\mathbb{A}| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 4-4 & 5-8 & 6-12 \\ 7-7 & 8-14 & 9-21 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{vmatrix}.$$

Using Laplace' theorem, developing for the entries of the first column, we obtain:

$$|\mathbb{A}| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \cdot \begin{vmatrix} -3 & -6 \\ -6 & -12 \end{vmatrix} = 36 - 36 = 0.$$

Even without making calculations it can be seen that the latter Determinant is 0: in fact the third line is the double of the second, and so for property P5) it follows that $|\mathbb{A}| = 0$.

Since the Determinant is 0, the three rows (and the three columns) are linearly dependent vectors. Using the previous calculations, we can also find the linear combination between these three rows.

From $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{vmatrix} = \begin{vmatrix} R_1 \\ R_2 - 4 \cdot R_1 \\ R_3 - 7 \cdot R_1 \end{vmatrix}$ we obtain:

$R_3 - 7 \cdot R_1 = 2 \cdot (R_2 - 4 \cdot R_1)$ and so: $R_3 = 2 \cdot R_2 - R_1$ or $R_1 - 2R_2 + R_3 = 0$.

It can be easily seen that the Determinant of a diagonal matrix is given by the product of the entries of the main diagonal.

The product of the entries of the main diagonal expresses also the value of the Determinant of a triangular matrix (lower triangular or upper triangular).

The Determinant of the unit matrix is equal to 1, whatever the order of the unit matrix is.

Example 50 : Let us compute the Determinant of the matrices:

$$\mathbb{A} = \left\| \begin{matrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & 5 & 0 \\ 1 & 2 & 3 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{matrix} \right\| \text{ and } \mathbb{B} = \left\| \begin{matrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 & 2 & 1 \\ 0 & 0 & 4 & 3 & 2 & 1 \\ 0 & 5 & 4 & 3 & 2 & 1 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{matrix} \right\|.$$

With appropriate exchanges between the rows, the matrix $\mathbb{A}$ can be brought to the shape of the diagonal matrix $\mathbb{D}$:

$$\mathbb{D} = \left\| \begin{matrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & 0 & 0 \\ 1 & 2 & 3 & 4 & 5 & 0 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \right\|, \text{ whose Determinant is then equal to } 6! = 720 \ .$$

As we need 8 exchanges between the rows to bring $\mathbb{A}$ to $\mathbb{D}$, it is $|\mathbb{A}| = |\mathbb{D}|$.

Operating as above, now interchanging the columns, also the matrix $\mathbb{B}$ can be brought to the shape of the diagonal matrix $\mathbb{D}$ of the previous example.

As we need 15 exchanges between the columns to bring $\mathbb{A}$ to $\mathbb{D}$, it is $|\mathbb{B}| = -|\mathbb{D}| = -720$.

Example 51 : Given the vectors $\mathbb{X} = \left(2, k, e^h\right)$, $\mathbb{Y} = (1,1,0)$ and $\mathbb{Z} = \left(2, e^h, k\right)$, let us determine the set of pairs (h, k) that make them linearly dependent vectors.
We construct the matrix $\mathbb{A}_3$ having the three vectors as rows, and then we impose that its Determinant must be equal to 0:

$$|\mathbb{A}_3| = \begin{vmatrix} 2 & k & e^h \\ 1 & 1 & 0 \\ 2 & e^h & k \end{vmatrix} = -1 \cdot \left(k^2 - e^{2h}\right) + 1 \cdot \left(2k - 2e^h\right) = e^{2h} - 2e^h - \left(k^2 - 2k\right) = 0 \,.$$

Considering it as a second degree polynomial equation with the unknown e^h, it is verified if:
$e^h = 1 \pm \sqrt{1 - 2k + k^2} = 1 \pm (1 - k)$, i.e. if $k = e^h$ and if $k = 2 - e^h$.

BINET' THEOREM

The calculus of the Determinant cannot be switched with sum and difference operations. In general, the Determinant of a matrix which is the sum of two matrices is not equal to the sum of the Determinants of the two matrices.
There is a formula, called Cauchy-Binet' formula, which is an identity for the Determinant of the product of two rectangular matrices of transpose shapes (so that the product is well defined and square). We will present only a special case of this formula, i.e. the statement that the Determinant of a matrix that is given by the product of two square matrices is equal to the product of their two Determinants.

Theorem (Binet): Given two matrices $\mathbb{A}_n$ and $\mathbb{B}_n$ of the same order n, if $\mathbb{C}_n = \mathbb{A}_n \cdot \mathbb{B}_n$ then $\det(\mathbb{C}_n) = \det(\mathbb{A}_n) \cdot \det(\mathbb{B}_n)$.

Example 52 : Given the two matrices $\mathbb{A} = \begin{Vmatrix} 1 & -1 & 1 \\ 2 & 0 & 4 \\ 4 & 1 & 8 \end{Vmatrix}$ and $\mathbb{B} = \begin{Vmatrix} 1 & 3 & 2 \\ 1 & 0 & 2 \\ 1 & 0 & 1 \end{Vmatrix}$, let us verify the validity of Binet' theorem.
We have, developing the Determinant of $\mathbb{A}$ for the entries of the second row, and developing the Determinant of $\mathbb{B}$ for the entries of the second column we have:

$$|\mathbb{A}| = -2 \cdot \begin{vmatrix} -1 & 1 \\ 1 & 8 \end{vmatrix} - 4 \cdot \begin{vmatrix} 1 & -1 \\ 4 & 1 \end{vmatrix} = -2 \cdot (-8 - 1) - 4 \cdot (1 + 4) = -2 \text{ and}$$

$$|\mathbb{B}| = -3 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = -3 \cdot (1 - 2) = 3 \,.$$

We now calculate the matrix product $\mathbb{A} \cdot \mathbb{B}$ and we obtain:

$$\mathbb{A} \cdot \mathbb{B} = \begin{Vmatrix} 1\cdot1 - 1\cdot1 + 1\cdot1 & 1\cdot3 - 1\cdot0 + 1\cdot0 & 1\cdot2 - 1\cdot2 + 1\cdot1 \\ 2\cdot1 + 0\cdot1 + 4\cdot1 & 2\cdot3 + 0\cdot0 + 4\cdot0 & 2\cdot2 + 0\cdot2 + 4\cdot1 \\ 4\cdot1 + 1\cdot1 + 8\cdot1 & 4\cdot3 + 1\cdot0 + 8\cdot0 & 4\cdot2 + 1\cdot2 + 8\cdot1 \end{Vmatrix} =$$

$$= \begin{Vmatrix} 1 & 3 & 1 \\ 6 & 6 & 8 \\ 13 & 12 & 18 \end{Vmatrix} \text{, and so } |\mathbb{A} \cdot \mathbb{B}| = \begin{vmatrix} 1 & 3 & 1 \\ 4 & 0 & 6 \\ 9 & 0 & 14 \end{vmatrix} = -3 \cdot (56 - 54) = -6 \,.$$

To calculate more quickly the Determinant of $\mathbb{A} \cdot \mathbb{B}$ we have replaced the second row with the second row itself minus twice the first row: $R_2 \leftarrow R_2 - 2\,R_1$, and the third row with the third row itself minus the first row multiplied by 4: $R_3 \leftarrow R_3 - 4\,R_1$.
Since $|\mathbb{A}| \cdot |\mathbb{B}| = -6$, the equality is verified.
If we compute also $\mathbb{B} \cdot \mathbb{A}$, we obtain $|\mathbb{B} \cdot \mathbb{A}| = -6$, even if $\mathbb{B} \cdot \mathbb{A} \neq \mathbb{A} \cdot \mathbb{B}$.

Then it can be shown that also the following applies:

Theorem (II° Laplace's Theorem): The sum of the products of the entries of a row (or column) of a matrix by the Cofactors of the entries of another row (or column) is always equal to 0.

THE RANK

Let us consider now matrices of any type, both rectangular and square; we give the following very important:

Definition 35 : The Rank of a matrix is the maximum order of its nonzero Minors.

The Rank of the matrix $\mathbb{A}$ will be denoted as $\text{Rank}(\mathbb{A})$.

Then the Rank of a matrix $\mathbb{A}$ is given by the order of the largest non-singular square submatrix included in the matrix $\mathbb{A}$.

By property P3) of the Determinant, we can also say that $\text{Rank}(\mathbb{A})$ expresses the maximum number of linearly independent rows (and therefore of columns), contained in the matrix $\mathbb{A}$.

Only the null matrix $\mathbb{O}$, i.e. the matrix having all its entries equal to 0, has the Rank equal to 0, while a non-singular square matrix has its Rank equal to its order.

Example 53 : Given the matrix $\mathbb{A}_4 = \begin{Vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & -1 \\ 4 & 2 & 2 & 4 \\ 3 & -1 & 1 & 2 \end{Vmatrix}$, it is valid that $|\mathbb{A}| = 0$.

Then the Rank of the matrix cannot be equal to 4, since its 4 rows (and its 4 columns) are linearly dependent vectors. So let us consider the submatrix $\mathbb{B}$ formed by the entries belonging to the 1^, 2^ and 4^ row, and to the 1^, 2^ and 3^ column: $\mathbb{B}_3 = \begin{Vmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 3 & -1 & 1 \end{Vmatrix}$.

Since $|\mathbb{B}| = 18$, we have found a nonzero Minor of the third order, and so $\text{Rank}(\mathbb{A}) = 3$.

We can also compute the Rank of this matrix using elementary operations on the rows (or on the columns), which do not change the Determinants of the Minors and then the Rank of the matrix.

Let us perform elementary operations: $\text{R}_3 \leftarrow \text{R}_3 - 4\,\text{R}_1$ and $\text{R}_4 \leftarrow \text{R}_4 - 3\,\text{R}_1$ to obtain:

$$\begin{Vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & -6 & 6 & -8 \\ 0 & -7 & 4 & -7 \end{Vmatrix}.$$

In this new matrix, let us then perform: $\text{R}_3 \leftarrow \text{R}_3 + 6\,\text{R}_2$ and $\text{R}_4 \leftarrow \text{R}_4 + 7\,\text{R}_2$ to obtain:

$$\begin{Vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 18 & -14 \\ 0 & 0 & 18 & -14 \end{Vmatrix}.$$

As we can see, the third and the fourth rows are equal, so the Rank cannot be equal to 4.

Since $\begin{vmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 18 \end{vmatrix} = 18 \neq 0$, the Rank of $\mathbb{A}_4$ is equal to 3.

Example 54 : Let us compute the Rank of the following matrices:

$$\mathbb{A} = \begin{Vmatrix} 1 & 0 & 1 & 1 & 3 \\ 2 & 1 & 3 & 4 & 7 \\ -1 & 3 & 2 & 5 & 1 \end{Vmatrix} ; \ \mathbb{B} = \begin{Vmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 3 & -2 \\ 5 & -1 & 7 & 4 \\ 3 & -2 & 0 & 4 \end{Vmatrix} ;$$

$$\mathbb{C} = \begin{Vmatrix} 1 & 0 & -1 & 2 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ -2 & 0 & 2 & -4 & -2 \end{Vmatrix} .$$

$\mathrm{Rank}(\mathbb{A}) = 3$, in fact : $\begin{vmatrix} 1 & 1 & 3 \\ 3 & 4 & 7 \\ 2 & 5 & 1 \end{vmatrix} = 1 \neq 0$;

$\mathrm{Rank}(\mathbb{B}) = 3$, since $|\mathbb{B}| = 0$ while $\begin{vmatrix} 1 & 3 & -2 \\ -1 & 7 & 4 \\ -2 & 0 & 4 \end{vmatrix} = -12 \neq 0$;

$\mathrm{Rank}(\mathbb{C}) = 2$, since all its Minors of the third order are 0 while $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \neq 0$.

Example 55 : Let us check, depending on the variation of the parameters m and k, the Rank of the matrix:

$$\mathbb{A} = \begin{Vmatrix} 1 & m & k \\ 2 & k & m \\ 3 & k+m & k+m \end{Vmatrix} .$$

The third row of the matrix is equal to the sum of the first and the second row, so $|\mathbb{A}| = 0$ and the Rank will never be equal to 3.

So let us consider the following submatrix: $\begin{Vmatrix} 1 & m & k \\ 2 & k & m \end{Vmatrix}$.

The Rank of the matrix is equal to 2 if:

$\begin{vmatrix} 1 & m \\ 2 & k \end{vmatrix} = k - 2m \neq 0$, i.e. $k \neq 2m$, or if:

$\begin{vmatrix} 1 & k \\ 2 & m \end{vmatrix} = m - 2k \neq 0$, i.e. $m \neq 2k$, or if:

$\begin{vmatrix} m & k \\ k & m \end{vmatrix} = m^2 - k^2 \neq 0$, i.e. $m \neq \pm k$.

None of these conditions will be satisfied if $k = m = 0$, and in this case the result is $\mathrm{Rank}(\mathbb{A}) = 1$, as in the first column there are nonzero entries.

Example 56 : Let us check, depending on the variation of the parameters m and k, through elementary row operations, the Rank of the matrix: $\mathbb{A} = \begin{Vmatrix} 1 & -1 & 0 & 2 & 1 \\ 3 & 1 & 2 & 1 & k \\ 1 & 3 & k & m & 0 \end{Vmatrix}$.

Let us begin with the following: $R_2 \leftarrow R_2 - 3R_1$ and $R_3 \leftarrow R_3 - R_1$ to obtain the matrix:

$\begin{Vmatrix} 1 & -1 & 0 & 2 & 1 \\ 0 & 4 & 2 & -5 & k-3 \\ 0 & 4 & k & m-2 & -1 \end{Vmatrix}$. Then we perform the $R_3 \leftarrow R_3 - R_2$ to obtain:

$$\begin{Vmatrix} 1 & -1 & 0 & 2 & 1 \\ 0 & 4 & 2 & -5 & k-3 \\ 0 & 0 & k-2 & m+3 & 2-k \end{Vmatrix} .$$

As it can be seen, we have obtained, through the first two columns, a triangular matrix, in which, to complete a third order Minor, we can put the third or the fourth or the fifth column of the given matrix. It is immediate to find the Determinant of these triangular matrices of order 3 . So we conclude that:
$\text{Rank}(\mathbb{A}) = 3$ if $k - 2 \neq 0 \Rightarrow k \neq 2$ or if $m + 3 \neq 0 \Rightarrow m \neq -3$.
If $k = 2$ and $m = -3$ we obtain:
$\begin{Vmatrix} 1 & -1 & 0 & 2 & 1 \\ 0 & 4 & 2 & -5 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{Vmatrix}$, and so $\text{Rank}(\mathbb{A}) = 2$ because the rows are reduced to two and also $\begin{vmatrix} 1 & -1 \\ 0 & 4 \end{vmatrix} = 4 \neq 0$, i.e. there is a nonzero Minor of order 2.

Without giving the proof, we state the following:
Theorem 8 : Every matrix $\mathbb{A}$ and its transpose $\mathbb{A}^{\mathrm{T}}$ have the same Rank:
$\text{Rank}(\mathbb{A}) = \text{Rank}(\mathbb{A}^{\mathrm{T}})$.
Theorem 9 : The Rank of the matrix product $\mathbb{A} \cdot \mathbb{B}$ is less than or equal to the Ranks of each of the two matrices $\mathbb{A}$ and $\mathbb{B}$: $\text{Rank}(\mathbb{A} \cdot \mathbb{B}) \leq \min\{\text{Rank}(\mathbb{A}); \text{Rank}(\mathbb{B})\}$.

Example 57 : Using the matrices $\mathbb{A} = \begin{Vmatrix} 1 & -1 & 1 \\ -3 & 2 & -1 \\ -2 & 1 & 0 \end{Vmatrix}$ and $\mathbb{B} = \begin{Vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{Vmatrix}$, performing the two products $\mathbb{A} \cdot \mathbb{B} = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$ and $\mathbb{B} \cdot \mathbb{A} = \begin{Vmatrix} -11 & 6 & -1 \\ -22 & 12 & -2 \\ -11 & 6 & -1 \end{Vmatrix}$, we see that
$\text{Rank}(\mathbb{A}) = 2$, $\text{Rank}(\mathbb{B}) = 1$, $\text{Rank}(\mathbb{A} \cdot \mathbb{B}) = 0$, $\text{Rank}(\mathbb{B} \cdot \mathbb{A}) = 1$.
The Rank of $\mathbb{A} \cdot \mathbb{B}$ shows us how the Rank of the matrix product $\mathbb{A} \cdot \mathbb{B}$ can also be less than the minimum Rank among the Ranks of the matrices $\mathbb{A}$ and $\mathbb{B}$.

Example 58 : Given the two matrices $\mathbb{A} = \begin{Vmatrix} 1 & x \\ x & 1 \end{Vmatrix}$ and $\mathbb{B} = \begin{Vmatrix} 1 & 2 & 3 \\ 1 & 2 & 2 \end{Vmatrix}$, let us determine if there exist values of x such that the Rank of $\mathbb{C} = \mathbb{A} \cdot \mathbb{B}$ is equal to 2.
Since the Rank of a product is less than or equal to those of the two matrices, firstly we need: $1 - x^2 \neq 0 \Rightarrow x \neq \pm 1$, so the matrix $\mathbb{A}$ has Rank equal to 2, the same Rank of $\mathbb{B}$, although this is not enough to ensure that $\text{Rank}(\mathbb{C})$ will be equal to 2.
Performing the product, we obtain: $\mathbb{C} = \begin{Vmatrix} 1+x & 2+2x & 3+2x \\ 1+x & 2+2x & 3x+2 \end{Vmatrix}$.
To have the Rank of $\mathbb{C}$ equal to 2, we need:
$(1+x)(2+3x) - (1+x)(3+2x) = (1+x)(x-1) \neq 0$, i.e. $x \neq \pm 1$,
since, as it canbe easily seen, the first and second columns of $\mathbb{C}$ are proportional vectors.

Using the Kronecker product, the following instead is valid:
Theorem 10 : If $\text{Rank}(\mathbb{A}) = n$ and $\text{Rank}(\mathbb{B}) = m$, then $\text{Rank}(\mathbb{A} \otimes \mathbb{B}) = n \cdot m$.

INVERSE MATRIX
Given a square matrix $\mathbb{A}_n$, we want to determine a square matrix, denoted by $\mathbb{A}_n^{-1}$, such that: $\mathbb{A}_n^{-1} \cdot \mathbb{A}_n = \mathbb{A}_n \cdot \mathbb{A}_n^{-1} = \mathbb{I}_n$, i.e. such that the product, both from the right and from the left, of this matrix by the matrix $\mathbb{A}_n$ gives, as its result, the unit matrix. If such a matrix exists, it will be called the inverse matrix of $\mathbb{A}_n$.

We immediately present an important property. In fact the following is valid:
Theorem 11 : If a square matrix $\mathbb{A}_n$ admits the inverse matrix $\mathbb{A}_n^{-1}$, then $\mathbb{A}_n^{-1}$ is unique.
Proof: If there were two different inverse matrices $\mathbb{A}^{-1}$ and $\mathbb{B}^{-1}$, applying the definition, we would have:
$\mathbb{A}^{-1} \cdot \mathbb{A} = \mathbb{A} \cdot \mathbb{A}^{-1} = \mathbb{I}$ and $\mathbb{B}^{-1} \cdot \mathbb{A} = \mathbb{A} \cdot \mathbb{B}^{-1} = \mathbb{I}$.
Applying the associative property, we obtain:
$\mathbb{A}^{-1} \cdot \mathbb{A} \cdot \mathbb{B}^{-1} = \left(\mathbb{A}^{-1} \cdot \mathbb{A}\right) \cdot \mathbb{B}^{-1} = \mathbb{I} \cdot \mathbb{B}^{-1} = \mathbb{B}^{-1}$ and also:
$\mathbb{A}^{-1} \cdot \mathbb{A} \cdot \mathbb{B}^{-1} = \mathbb{A}^{-1} \cdot \left(\mathbb{A} \cdot \mathbb{B}^{-1}\right) = \mathbb{A}^{-1} \cdot \mathbb{I} = \mathbb{A}^{-1}$, i.e. $\mathbb{B}^{-1} = \mathbb{A}^{-1}$. •

We must say that the problem of the existence, uniqueness and determination of the inverse matrix is a much more general problem, also studied for rectangular matrices by means of the so-called Moore-Penrose' inverse, with results different from those found for square matrices. Here we treat only the case of the inverse of a nonsingular square matrix.

Having verified that it exists, let's see, now, how to build the inverse of a nonsingular square matrix.

Definition 36 : Adjugate (or classical adjoint) matrix adj($\mathbb{A}_n$).
In the Anglo-Saxon mathematical literature the adjugate (or classical adjoint) matrix adj($\mathbb{A}_n$) (or $\mathbb{A}_n^*$) is the transpose of the matrix whose entries are the cofactors of the entries of $\mathbb{A}_n$.
In the Italian mathematical literature the adjugate (or adjoint) matrix adj($\mathbb{A}_n$) (or $\mathbb{A}_n^*$) often means the matrix whose entries are the cofactors of the entries of $\mathbb{A}_n$.

If we use the Italian version, then the following is valid:
Theorem 12 : The Inverse $\mathbb{A}_n^{-1}$ of a nonsingular square matrix $\mathbb{A}_n$ is given by the transpose of the adjugate matrix adj($\mathbb{A}_n$), divided by the determinant of $\mathbb{A}$, namely:
$\mathbb{A}_n^{-1} = \dfrac{1}{|\mathbb{A}|} \cdot \left(\mathbb{A}_n^*\right)^{\mathrm{T}}$.
So:
-we have to calculate the Determinant of the matrix $\mathbb{A}_n$ and verify that it is different from 0;
- we must calculate the adjugate matrix adj($\mathbb{A}_n$) replacing each entry of $\mathbb{A}_n$ with its cofactor;
- we must transpose the adjugate matrix;
- we must divide each entry of the transpose of the adjugate matrix for the Determinant $|\mathbb{A}|$.

Indicating with $\mathbf{A}_{ij}$ the Cofactor of the entry a_{ij}, we can write:

$$\mathbb{A}_n^{-1} = \left\| \begin{matrix} \dfrac{\mathrm{A}_{11}}{|\mathbb{A}|} & \dfrac{\mathrm{A}_{21}}{|\mathbb{A}|} & \dots & \dfrac{\mathrm{A}_{n1}}{|\mathbb{A}|} \\ \dfrac{\mathrm{A}_{12}}{|\mathbb{A}|} & \dfrac{\mathrm{A}_{22}}{|\mathbb{A}|} & \dots & \dfrac{\mathrm{A}_{n2}}{|\mathbb{A}|} \\ \dots & \dots & \dots & \dots \\ \dfrac{\mathrm{A}_{1\mathrm{i}}}{|\mathbb{A}|} & \dfrac{\mathrm{A}_{2\mathrm{i}}}{|\mathbb{A}|} & \dots & \dfrac{\mathrm{A}_{ni}}{|\mathbb{A}|} \\ \dots & \dots & \dots & \dots \\ \dfrac{\mathrm{A}_{1n}}{|\mathbb{A}|} & \dfrac{\mathrm{A}_{2n}}{|\mathbb{A}|} & \dots & \dfrac{\mathrm{A}_{nn}}{|\mathbb{A}|} \end{matrix} \right\|.$$

Proof: Let us consider the product $\mathbb{A}_n^{-1} \cdot \mathbb{A}_n = \mathbb{B}_n$.

The entry b_{ij} of the result matrix $\mathbb{B}_n$ is given by the product between the i-th row of $\mathbb{A}_n^{-1}$ and the j-th column of $\mathbb{A}_n$, i.e.:
$$b_{ij} = \frac{\mathrm{A}_{1i}}{|\mathbb{A}|}\, a_{1j} + \frac{\mathrm{A}_{2i}}{|\mathbb{A}|}\, a_{2j} + ... + \frac{\mathrm{A}_{ni}}{|\mathbb{A}|}\, a_{nj} = \frac{1}{|\mathbb{A}|}\,[\, \mathrm{A}_{1i}\, a_{1j} \;+\; \mathrm{A}_{2i}\, a_{2j} \;+ ... +\; \mathrm{A}_{ni}\, a_{nj}\,]\,.$$
If $i = j$, the term into brackets is the development of the Determinant of $\mathbb{A}_n$ using the j-th column, and so the result is equal to 1, while, if $i \neq j$, for the second Laplace' theorem, the result is equal to 0.
Similar considerations for the product $\mathbb{A}_n \cdot \mathbb{A}_n^{-1}$, from which it follows that $\mathbb{B}_n = \mathbb{I}_n$. •

Example 59 : Let us compute the inverse matrix of $\mathbb{A}_{33} = \begin{Vmatrix} 3 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{Vmatrix}$.

It is $\mathrm{adj}(\mathbb{A}_n) = \begin{Vmatrix} 1 & 1 & -1 \\ -1 & 5 & -1 \\ 0 & -4 & 2 \end{Vmatrix}$. Then $\left(\mathrm{adj}(\mathbb{A}_n)\right)^{\mathrm{T}} = \begin{Vmatrix} 1 & -1 & 0 \\ 1 & 5 & -4 \\ -1 & -1 & 2 \end{Vmatrix}$,

and since $\det(\mathbb{A}) = 2$, we finally obtain: $\mathbb{A}^{-1} = \begin{Vmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{5}{2} & -2 \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{Vmatrix}$.

Performing the products we can see that:
$$\mathbb{A} \cdot \mathbb{A}^{-1} = \begin{Vmatrix} 3 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{Vmatrix} \cdot \begin{Vmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{5}{2} & -2 \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{Vmatrix} = \mathbb{I}_3 \text{ and that}$$
$$\mathbb{A}^{-1} \cdot \mathbb{A} = \begin{Vmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{5}{2} & -2 \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{Vmatrix} \cdot \begin{Vmatrix} 3 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{Vmatrix} = \mathbb{I}_3\,.$$

Finally, for the inverse matrix the following properties are valid:
I1) $\left(\mathbb{A}^{-1}\right)^{-1} = \mathbb{A}$ (the inverse of the inverse matrix is the matrix itself);
I2) $(\mathbb{A} \cdot \mathbb{B})^{-1} = \mathbb{B}^{-1} \cdot \mathbb{A}^{-1}$:
in fact $(\mathbb{A} \cdot \mathbb{B})^{-1} \cdot (\mathbb{A} \cdot \mathbb{B}) = \mathbb{B}^{-1} \cdot \left(\mathbb{A}^{-1} \cdot \mathbb{A}\right) \cdot \mathbb{B} = \mathbb{B}^{-1} \cdot \mathbb{I} \cdot \mathbb{B} = \mathbb{I}$;
I3) $\left|\mathbb{A}^{-1}\right| = \dfrac{1}{|\mathbb{A}|}$:
in fact, applying Binet' theorem, we obtain $\left|\mathbb{A}^{-1} \cdot \mathbb{A}\right| = \left|\mathbb{A}^{-1}\right| \cdot |\mathbb{A}| = |\mathbb{I}| = 1$ and so the thesis;
I4) $\left(\mathbb{A}^{\mathrm{T}}\right)^{-1} = \left(\mathbb{A}^{-1}\right)^{\mathrm{T}}$, i.e. the inverse of the transpose matrix and the transpose of the inverse matrix are the same matrix:
in fact, from $\mathbb{A}^{-1} \cdot \mathbb{A} = \mathbb{I}$ we obtain $\left(\mathbb{A}^{-1} \cdot \mathbb{A}\right)^{\mathrm{T}} = \mathbb{I}^{\mathrm{T}}$, and so $\mathbb{A}^{\mathrm{T}} \cdot \left(\mathbb{A}^{-1}\right)^{\mathrm{T}} = \mathbb{I}$.
Since $\mathbb{I} = \mathbb{A}^{\mathrm{T}} \cdot \left(\mathbb{A}^{\mathrm{T}}\right)^{-1}$ we have the thesis;
I5) The inverse of a diagonal matrix is still a diagonal matrix, having as entry b_{ii} the reciprocal of the entry a_{ii}: $b_{ii} = \dfrac{1}{a_{ii}}$;
I6) The inverse of a symmetric matrix is a symmetric matrix:
in fact, from $\mathbb{A} = \mathbb{A}^{\mathrm{T}}$ we obtain $\mathbb{A}^{-1} = \left(\mathbb{A}^{\mathrm{T}}\right)^{-1}$, but, for the property I4) it is $\mathbb{A}^{-1} = \left(\mathbb{A}^{-1}\right)^{\mathrm{T}}$, and so the thesis.

Example 60 : Let us consider the matrix $\mathbb{A} = \begin{Vmatrix} 3 & 8 \\ -1 & -3 \end{Vmatrix}$. It is $\mathbb{A} \cdot \mathbb{A} = \mathbb{A}^2 = \mathbb{I}$:

$$\mathbb{A} \cdot \mathbb{A} = \mathbb{A}^2 = \begin{Vmatrix} 3 & 8 \\ -1 & -3 \end{Vmatrix} \cdot \begin{Vmatrix} 3 & 8 \\ -1 & -3 \end{Vmatrix} = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} = \mathbb{I}$$

and so, since $|\mathbb{A}| = -1 \neq 0$, we obtain $\mathbb{A}^{-1} \cdot \mathbb{A} \cdot \mathbb{A} = \mathbb{A}^{-1} \cdot \mathbb{I}$ i.e. $\mathbb{A} = \mathbb{A}^{-1}$.

For the Kronecker product, otherwise the "rows by columns" product, regarding the inverse matrix, the following property is valid:
K1) if $\mathbb{A}$ and $\mathbb{B}$ are invertible matrices, then $(\mathbb{A} \otimes \mathbb{B})^{-1} = \mathbb{A}^{-1} \otimes \mathbb{B}^{-1}$.

To calculate the inverse matrix we can also follow another procedure, based on elementary operations. We write, one beside the other, the given matrix $\mathbb{A}$ and the unit matrix, to form a single matrix. By elementary operations on the rows, we transform the original matrix $\mathbb{A}$ to become the unit matrix: the matrix on the right is then the inverse matrix $\mathbb{A}^{-1}$.

Example 61 : Let us compute the inverse matrix of the matrix $\mathbb{A} = \begin{Vmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{Vmatrix}$ using elementary operations on rows. We begin from $\left\Vert \begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array} \right\Vert = [\mathbb{A} \,|\, \mathbb{I}]$.

Let us begin with $R_2 \leftarrow R_2 - R_1$ and $R_3 \leftarrow R_3 - R_1$ to obtain: $\left\Vert \begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 2 & 2 & -1 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \end{array} \right\Vert$.

Then by $R_2 \leftarrow \frac{1}{2} R_2$ we obtain: $\left\Vert \begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \end{array} \right\Vert$.

Then by $R_3 \leftarrow R_3 - R_2$ we obtain: $\left\Vert \begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & 1 \end{array} \right\Vert$.

Finally, by $R_2 \leftarrow R_2 - R_3$ and by $R_1 \leftarrow R_1 + R_3$ we obtain:

$$\left\Vert \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & 1 \end{array} \right\Vert = \left[\mathbb{I} \,|\, \mathbb{A}^{-1}\right].$$

Since the matrix on the left is the unit matrix, the one on the right is the matrix $\mathbb{A}^{-1}$. And thus:

$$\mathbb{A}^{-1} = \begin{Vmatrix} \frac{1}{2} & -\frac{1}{2} & 1 \\ 0 & 1 & -1 \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{Vmatrix}.$$

We conclude with the following
Theorem 13 : The only idempotent and nonsingular matrix $\mathbb{A}_n$ is the unit matrix $\mathbb{I}_n$.
Proof: If $\mathbb{A} \cdot \mathbb{A} = \mathbb{A}$, and if $\mathbb{A}$ is a non singular matrix, multiplying on the left by the inverse matrix we obtain: $\mathbb{A}^{-1} \cdot \mathbb{A} \cdot \mathbb{A} = \mathbb{A}^{-1} \cdot \mathbb{A}$ i.e. $\left(\mathbb{A}^{-1} \cdot \mathbb{A}\right) \cdot \mathbb{A} = \mathbb{I} \cdot \mathbb{A} = \mathbb{A} = \mathbb{A}^{-1} \cdot \mathbb{A} = \mathbb{I}$ and so $\mathbb{A} = \mathbb{I}$.
And thus a matrix, different from the unit matrix, to be idempotent must be singular.

LINEAR MAPS AND LINEAR SYSTEMS

Two of the most important applications of matrix calculus are the study of linear maps (or linear functions) and solving systems of linear equations.

Let us consider a matrix $\mathbb{A}_{m,n}$, and let $\mathbb{X} \in \mathbb{R}^n$ be a column vector. The "rows by columns" product $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1}$ gives as a result a column vector $\mathbb{Y}_{m,1} \in \mathbb{R}^m$, i.e. $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$.

If we consider the vector $\mathbb{X}$ as an independent variable, through the "rows by columns" product of the matrix A by the column vector $\mathbb{X}$ we can construct a function (or map) $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ that associates to each vector $\mathbb{X} \in \mathbb{R}^n$ one and only one vector $\mathbb{Y} \in \mathbb{R}^m$, such that $\mathbb{Y} = \mathbb{A} \cdot \mathbb{X}$.

The functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ given in this form are called linear maps.

$\mathbb{R}^n$ is said the domain of the function, $\mathbb{R}^m$ is said the codomain of the function.

Conversely, if we consider $\mathbb{Y}$ as an assigned vector, we want to check if there are and how many vectors $\mathbb{X}$ there are that satisfy such equation. Solving such type of problem is what we call to solve a linear system.

LINEAR MAPS

Definition 37 : A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be a linear map if the following two conditions are satisfied:

1) $f(\mathbb{X}_1 + \mathbb{X}_2) = f(\mathbb{X}_1) + f(\mathbb{X}_2), \forall \mathbb{X}_1, \mathbb{X}_2 \in \mathbb{R}^n$ and

2) $f(k \cdot \mathbb{X}) = k \cdot f(\mathbb{X}), \forall k \in \mathbb{R}$

or if

Linearity property: $f(k_1 \mathbb{X}_1 + k_2 \mathbb{X}_2) = k_1 f(\mathbb{X}_1) + k_2 f(\mathbb{X}_2), \forall k \in \mathbb{R}, \forall \mathbb{X}_1, \mathbb{X}_2 \in \mathbb{R}^n$.

So a map is said to be linear if the image of any linear combination coincides with the linear combination of the images.

Expressing the vector $\mathbb{X} = (x_1,\, x_2, ..., x_n)$ using the vectors of a basis of $\mathbb{R}^n$, for example the standard basis, from $\mathbb{X} = x_1\, \mathbf{e}_1 + x_2\, \mathbf{e}_2 + ... +\; x_n\, \mathbf{e}_n$, we get, by the linearity property:

$f(\mathbb{X}) = f(x_1\, \mathbf{e}_1 + x_2\, \mathbf{e}_2 + ... +\; x_n\, \mathbf{e}_n) = x_1\, f(\mathbf{e}_1) + x_2\, f(\mathbf{e}_2) + ... +\; x_n\, f(\mathbf{e}_n)$.

Therefore it is sufficient to know the image of the elements of the chosen basis to have the image of any element $\mathbb{X}$.

Example 62 : We consider $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$, such that: $\begin{cases} f(\mathbf{e}_1) = (1,1,2) \\ f(\mathbf{e}_2) = (2,0,1) \end{cases}$. So we obtain:

$f(3,5) = f(3\mathbf{e}_1 + 5\mathbf{e}_2) = 3\, f(\mathbf{e}_1) + 5\, f(\mathbf{e}_2) = 3\,(1,1,2) + 5\,(2,0,1) = (13,3,11)$.

Given $\mathbb{X} = (x_1,\, x_2, ..., x_n)$ and $\mathbb{Y} = (y_1, y_2, ..., y_m)$, we can also write:

$f(\mathbb{X}) = f(x_1, x_2, .., x_n) = (y_1, y_2, .., y_m) = (f_1(x_1, x_2, .., x_n), .., f_m\,(x_1, x_2, .., x_n)) = \mathbb{Y}$.

So, a map is a linear map if each of the functions $f_i\,(x_1, x_2, ..., x_n)$ is a linear map, namely in the form $f_i\,(x_1, x_2, ..., x_n) = a_{i1}\, x_1 + a_{i2}\, x_2 +\; ... + a_{in}\, x_n$.

From this it follows that every linear map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ can be expressed in the form:

$\mathbb{Y} = f(\mathbb{X}) = \mathbb{A} \cdot \mathbb{X}$, where $\mathbb{A}$ is a matrix having m rows and n columns.

Conversely, to each matrix $\mathbb{A}_{m,n}$ corresponds a linear map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

In fact, for the properties of the "rows by columns" product, the definition of linear map is satisfied:

1) $\mathbb{A} \cdot (\mathbb{X}_1 + \mathbb{X}_2) = \mathbb{A} \cdot \mathbb{X}_1 + \mathbb{A} \cdot \mathbb{X}_2$ i.e. $f(\mathbb{X}_1 + \mathbb{X}_2) = f(\mathbb{X}_1) + f(\mathbb{X}_2)$,

2) $\mathbb{A} \cdot (k\, \mathbb{X}) = k\, \mathbb{A} \cdot \mathbb{X}$ i.e. $f(k\, \mathbb{X}) = k\, f(\mathbb{X})$, $\forall k \in \mathbb{R}$

and finally:

$\mathbb{A} \cdot (k_1 \mathbb{X}_1 + k_2 \mathbb{X}_2) = k_1 \mathbb{A}\mathbb{X}_1 + k_2 \mathbb{A}\mathbb{X}_2 \Rightarrow f(k_1 \mathbb{X}_1 + k_2 \mathbb{X}_2) = k_1 f(\mathbb{X}_1) + k_2 f(\mathbb{X}_2)$.

From $\mathbb{Y}_{m,1} = \mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1}$ it follows $\mathbb{A}_{m,n} \cdot \mathbb{O} = \mathbb{O}$ or $f(\mathbb{O}) = \mathbb{O}$ for any linear map f.

Example 63 : Let us consider $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, $f(x_1, x_2, x_3) = (2x_1 + x_3, x_1 + x_2 - 3\,x_3)$.
Consisting only of linear combinations of (x_1, x_2, x_3), f is a linear map, which can be written as: $\begin{Vmatrix} 2 & 0 & 1 \\ 1 & 1 & -3 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} = \begin{Vmatrix} y_1 \\ y_2 \end{Vmatrix}$.

LINEAR MAPS AS LINEAR COMBINATIONS

If we denote by $\mathrm{C}_1, \mathrm{C}_2, ..., \mathrm{C}_n$ the columns of the matrix $\mathbb{A}_{m,n}$, and if $\mathbb{X} = (x_1, x_2, ..., x_n)$, the product $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$ can also be seen as a linear combination of the columns of the matrix $\mathbb{A}_{m,n}$: $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathrm{C}_1 \cdot x_1 + \mathrm{C}_2 \cdot x_2 + ... + \mathrm{C}_n \cdot x_n = \mathbb{Y}_{m,1}$, and then it can be interpreted as the search for a suitable linear combination of the columns of the matrix $\mathbb{A}_{m,n}$ by which we express the vector $\mathbb{Y}$.

LINEAR SYSTEMS

A system of linear equations with m equations and n unknowns (or variables) $x_1, x_2, ..., x_n$ can be represented as:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = y_2 \\ \quad \text{..} \\ a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n = y_i \\ \quad \text{..} \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = y_m \end{cases} \Leftrightarrow \mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}\,.$$

If $\mathbb{A}_{m,n}$ is the coefficient matrix, i.e. the matrix whose entries a_{ij} are the coefficients of the unknowns $(x_1, x_2, ..., x_n)$, if $\mathbb{X}$ is the column vector $(n \cdot 1)$ having as components the unknowns $(x_1, x_2, ..., x_n)$, and if $\mathbb{Y}$ is the column vector $(m \cdot 1)$ having as components the constant terms $(y_1, y_2, ..., y_m)$, then the linear system can be expressed by the matrix product $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$, that is, it can be seen as a linear map f, with matrix $\mathbb{A}_{m,n}$, which associates, as image under f, to the unknown vector $\mathbb{X}_{n,1}$ the vector of the constant terms $\mathbb{Y}_{m,1}$; and so finding a solution is like determining the inverse image (or preimage) of $\mathbb{Y}$.

We can also write the system as $\mathrm{C}_1 \cdot x_1 + \mathrm{C}_2 \cdot x_2 + ... + \mathrm{C}_n \cdot x_n = \mathbb{Y}$, that is, we can see every solution $x_1, x_2, ..., x_n$ of the system as the n coefficients that allow us to express the vector of constant terms $\mathbb{Y}$ as a linear combination of the columns C_i of the matrix $\mathbb{A}_{n,m}$.

To solve a linear system means to determine all its possible solutions, i.e. all the n-tuples $(x_1, x_2, ..., x_n)$ that simultaneously satisfy the m given equations.
If the constant terms $y_1, ..., y_m$ are all zero, that is, if the vector of the constant terms is the null vector $\mathbb{O}$, the linear system is said an "homogeneous system".

SQUARE MATRIX SYSTEMS - CRAMER'S RULE

Let us now consider a system of linear equations with as many equations as unknowns (or variables), i.e. the matrix $\mathbb{A}$ is a square matrix: $\mathbb{A}_n$.
If $\mathbb{A}_n$ is a non-singular matrix, then its inverse $\mathbb{A}_n^{-1}$ exists and it is unique and so, multiplying on the left by $\mathbb{A}_n^{-1}$ the two members of the equation $\mathbb{A}_n \cdot \mathbb{X} = \mathbb{Y}$, we obtain:
$\mathbb{A}_n^{-1} \cdot (\mathbb{A}_n \cdot \mathbb{X}) = \left(\mathbb{A}_n^{-1} \cdot \mathbb{A}_n\right) \cdot \mathbb{X} = \mathbb{I}_n \cdot \mathbb{X} = \mathbb{X} = \mathbb{A}_n^{-1} \cdot \mathbb{Y}$,

i.e. we can obtain the solution of the system multiplying the inverse matrix of $\mathbb{A}$ by the vector of the constant terms: $\mathbb{X} = \mathbb{A}_n^{-1} \cdot \mathbb{Y}$.

This is explained in the following:
Theorem 14 (Cramer's Rule) : A linear equations system with as many equations as unknowns, whose coefficient matrix is a non-singular square matrix, admits one and only one solution $(x_1, x_2, ..., x_n)$. The individual values for the unknowns (or variables) are given by a quotient, the denominator of which is always the Determinant of the coefficient matrix, and the numerator is the Determinant of the matrix obtained by replacing, in the coefficient matrix, the column of the coefficients of the unknown we want to compute with the vector of the constant terms.
Proof: If $\mathbb{A}_n$ is a non-singular square matrix, then $\mathbb{X} = \mathbb{A}_n^{-1} \cdot \mathbb{Y}$; and so:

$$\left\| \begin{matrix} x_1 \\ x_2 \\ .... \\ x_i \\ .... \\ x_n \end{matrix} \right\| = \left\| \begin{matrix} \frac{\mathrm{A}_{11}}{|\mathbb{A}|} & \frac{\mathrm{A}_{21}}{|\mathbb{A}|} & ..... & \frac{\mathrm{A}_{n1}}{|\mathbb{A}|} \\ \frac{\mathrm{A}_{12}}{|\mathbb{A}|} & \frac{\mathrm{A}_{22}}{|\mathbb{A}|} & ..... & \frac{\mathrm{A}_{n2}}{|\mathbb{A}|} \\ ..... & ..... & ..... & ..... \\ \frac{\mathrm{A}_{1i}}{|\mathbb{A}|} & \frac{\mathrm{A}_{2i}}{|\mathbb{A}|} & ..... & \frac{\mathrm{A}_{ni}}{|\mathbb{A}|} \\ ..... & ..... & ..... & ..... \\ \frac{\mathrm{A}_{1n}}{|\mathbb{A}|} & \frac{\mathrm{A}_{2n}}{|\mathbb{A}|} & ..... & \frac{\mathrm{A}_{nn}}{|\mathbb{A}|} \end{matrix} \right\| \cdot \left\| \begin{matrix} y_1 \\ y_2 \\ .... \\ y_i \\ .... \\ y_n \end{matrix} \right\|$$

from which, performing the product of the i-th row of the matrix $\mathbb{A}_n^{-1}$ by the vector $\mathbb{Y}$, we obtain: $x_i = \dfrac{(\mathrm{A}_{1i}\, y_1 \ + \ \mathrm{A}_{2i}\, y_2 \ + ... + \ \mathrm{A}_{ni}\, y_n)}{|\mathbb{A}|}$.

The numerator of this fraction is nothing but the development, using the i-th column, of a Determinant having the same columns of the matrix $\mathbb{A}$, except the i-th one, instead of which there is the column of the constant terms, and so we obtain:

$$x_i = \frac{\begin{vmatrix} a_{11} & a_{12} & ..... & a_{1i-1} & y_1 & a_{1i+1} & ..... & a_{1n} \\ a_{21} & a_{22} & ..... & a_{2i-1} & y_2 & a_{2i+1} & ..... & a_{2n} \\ ..... & ..... & ..... & ..... & ..... & ..... & ..... & ..... \\ a_{n1} & a_{n2} & & a_{ni-1} & y_n & a_{ni+1} & & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & ..... & a_{1i-1} & a_{1i} & a_{1i+1} & ..... & a_{1n} \\ a_{21} & a_{22} & ..... & a_{2i-1} & a_{2i} & a_{2i+1} & ..... & a_{2n} \\ ..... & ..... & ..... & ..... & ..... & ..... & ..... & ..... \\ a_{n1} & a_{n2} & & a_{ni-1} & a_{ni} & a_{ni+1} & & a_{nn} \end{vmatrix}}$$

namely the thesis of Cramer's Rule. ●

If a system with a square non-singular matrix $\mathbb{A}$ has one and only one solution, this means that the columns of the matrix $\mathbb{A}$ are linearly independent vectors, so they are a basis for $\mathbb{R}^n$ and then there exists one and only one way to express every constant vector as a linear combination of these columns.

Example 64 : Let us determine the values of the variables x, y and z for which it is satisfied:

$$\begin{Vmatrix} x & 2 & y \\ 3 & 1 & z \\ y & 1 & 4 \end{Vmatrix} \cdot \begin{Vmatrix} 2 \\ x \\ 4 \end{Vmatrix} = \begin{Vmatrix} z \\ y \\ z \end{Vmatrix} .$$

Performing the product, we obtain $\begin{Vmatrix} 4x+4y \\ 6+x+4z \\ 2y+x+16 \end{Vmatrix} = \begin{Vmatrix} z \\ y \\ z \end{Vmatrix}$, and then, by equating the components, we obtain the system $\begin{cases} 4x+4y-z=0 \\ x-y+4z=-6 \\ x+2y-z=-16 \end{cases}$ having 3 equations and 3 unknowns.

Let us compute the Determinant of the coefficient matrix, to obtain:

$$\begin{vmatrix} 4 & 4 & -1 \\ 1 & -1 & 4 \\ 1 & 2 & -1 \end{vmatrix} = \begin{vmatrix} 4 & 4 & -1 \\ 1 & -1 & 4 \\ 0 & 3 & -5 \end{vmatrix} = 4 \cdot \begin{vmatrix} -1 & 4 \\ 3 & -5 \end{vmatrix} - 1 \cdot \begin{vmatrix} 4 & -1 \\ 3 & -5 \end{vmatrix} = -11 .$$

From Cramer's Rule, the system has one and only one solution, which is given by:

$$x = \frac{\begin{vmatrix} 0 & 4 & -1 \\ -6 & -1 & 4 \\ -16 & 2 & -1 \end{vmatrix}}{-11} = \frac{252}{11} \ ; \ y = \frac{\begin{vmatrix} 4 & 0 & -1 \\ 1 & -6 & 4 \\ 1 & -16 & -1 \end{vmatrix}}{-11} = -\frac{290}{11} \ ;$$

$$z = \frac{\begin{vmatrix} 4 & 4 & 0 \\ 1 & -1 & -6 \\ 1 & 2 & -16 \end{vmatrix}}{-11} = -\frac{152}{11} .$$

Example 65 : Let us determine the values of the parameters m and k for which the system:

$\begin{cases} x+2y+kz=1 \\ x+3y+kz=2 \\ mx+4y+kz=0 \end{cases}$ has one and only one solution.

From Cramer's Rule, the Determinant of the coefficient matrix must be different from zero, so we need:

$$\begin{vmatrix} 1 & 2 & k \\ 1 & 3 & k \\ m & 4 & k \end{vmatrix} = \begin{vmatrix} 1 & 2 & k \\ 0 & 1 & 0 \\ m & 4 & k \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & k \\ m & k \end{vmatrix} = k\,(1-m) \neq 0 .$$

This is satisfied when it is, simultaneously, $m \neq 1$ and $k \neq 0$.

LINEAR HOMOGENEOUS SYSTEMS

Since $\mathbb{A} \cdot \mathbb{O} = \mathbb{O}, \forall \mathbb{A}$, it is clear that every homogeneous system always admits at least the zero solution $\mathbb{X} = \mathbb{O}$.

If the homogeneous system has as many equations as unknowns, and if its coefficient matrix is non-singular, then, by Cramer's Rule, the system has only one solution that is, as just said, the null solution $\mathbb{X} = \mathbb{O}$.

This means that the columns of the coefficient matrix are linearly independent vectors, and so their only linear combination which gives as a result the null vector (i.e. the vector of the constant terms) must have coefficients (i.e. the solution of the system) all equal to 0.

In order that a homogeneous system with a square coefficient matrix has instead other solutions in addition to the null one, the Determinant of the coefficient matrix $\mathbb{A}$ should be equal to 0.

Example 66 : Given the three vectors $\mathbb{X} = (1,2,4)$, $\mathbb{Y} = (-1,1,-1)$ and $\mathbb{Z} = (1,5,7)$, after verifying that they are linearly dependent vectors, let us determine the coefficients of their linear combination that gives as a result the null vector.
We obtain, performing $\mathrm{R}_2 \leftarrow \mathrm{R}_2 + \mathrm{R}_1$ and $\mathrm{R}_3 \leftarrow \mathrm{R}_3 - \mathrm{R}_1$:
$\begin{vmatrix} 1 & 2 & 4 \\ -1 & 1 & -1 \\ 1 & 5 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 4 \\ 0 & 3 & 3 \\ 0 & 3 & 3 \end{vmatrix} = 0$, since the second and the third row are equal.
Then: $\mathrm{R}_2 + \mathrm{R}_1 = \mathrm{R}_3 - \mathrm{R}_1$, i.e.: $2\,\mathbb{X} + \mathbb{Y} - \mathbb{Z} = \mathbb{O}$.
We can also use another procedure, that of linear systems.
We must find α, β and γ such that: $\alpha\,\mathbb{X} + \beta\,\mathbb{Y} + \gamma\,\mathbb{Z} = \mathbb{O}$, or such that:
$\begin{cases} \alpha - \beta + \gamma = 0 \\ 2\alpha + \beta + 5\gamma = 0 \\ 4\alpha - \beta + 7\gamma = 0 \end{cases}$.
The Determinant of the coefficient matrix is clearly equal to 0, since the columns are the linearly dependent vectors $\mathbb{X}, \mathbb{Y}$ and $\mathbb{Z}$.
We subtract the second equation from the third and we have the system:
$\begin{cases} \alpha - \beta + \gamma = 0 \\ 2\alpha + \beta + 5\gamma = 0 \\ 2\alpha - 2\beta + 2\gamma = 0 \end{cases}$;
now the third equation is equal to the double of the first and so it can be discarded.
There remain only the first two equations and, leaving γ as a free variable, we obtain:
$\begin{cases} \alpha - \beta = -\gamma \\ 2\alpha + \beta = -5\gamma \end{cases}$ and from these easily we solve: $\begin{cases} \alpha = -2\gamma \\ \beta = -\gamma \end{cases}$.
Therefore there are, as a function of γ, ∞^1 linear combinations of $\mathbb{X}$, $\mathbb{Y}$ and $\mathbb{Z}$ that give as a result the null vector, all expressible in the form: $-2\gamma\,\mathbb{X} - \gamma\,\mathbb{Y} + \gamma\,\mathbb{Z} = \mathbb{O}$, which generalizes the solution found above.

ROUCHE'-CAPELLI THEOREM
$\mathbb{A}_{m,n}$ is the coefficient matrix, i.e. the matrix whose entries a_{ij} are the coefficients of the unknowns (or variables) $(x_1, x_2, ..., x_n)$.
The augmented matrix $(\mathbb{A} \,|\, \mathbb{Y})_{m,n+1}$ is obtained adding the column of constant terms to the coefficient matrix.
This theorem, valid for any linear system $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$, ensures, under suitable assumptions, the existence of solutions for the linear system and then allows to compute their number, given the rank of its coefficient matrix and the rank of its augmented matrix.
The following applies:
Theorem 15 (Rouchè-Capelli) : A linear system, whatever the number of its equations and its unknowns, has solutions if and only if the Rank of the coefficient matrix $\mathbb{A}$ is equal to the Rank of the augmented matrix $(\mathbb{A} \,|\, \mathbb{Y})$. That is:
$\exists\, \mathbb{X}$: $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1} \Leftrightarrow \mathrm{Rank}(\mathbb{A}) = \mathrm{Rank}(\mathbb{A} \,|\, \mathbb{Y})$.
Proof: This theorem states that a linear system has solutions if and only if the vector of constant terms, together with the columns of the coefficient matrix, are a set of linearly dependent vectors because, adding to the columns of $\mathbb{A}$ the vector of constant terms, the Rank, namely the number of linearly independent vectors, does not increase; then the vector of constant terms can be expressed as a linear combination of the columns of the coefficient matrix, i.e.:
$\mathbb{Y} = \mathrm{C}_1 \cdot x_1 + \mathrm{C}_2 \cdot x_2 + ... + \mathrm{C}_n \cdot x_n \Leftrightarrow \mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$.
If the Ranks are different, it results that: $\mathrm{Rank}(\mathbb{A} \,|\, \mathbb{Y}) = \mathrm{Rank}(\mathbb{A}) + 1$, and this means that the vector of constant terms $\mathbb{Y}$ is linearly independent from the columns of the coefficient

matrix, so it is not possible to express $\mathbb{Y}$ as a linear combination of them, and so the system has no solutions. ●

If the system has solutions, the common Rank of coefficient matrix and augmented matrix also expresses the number of the significant equations of the system: if m is the number of the equations and k is the common Rank (obviously with $m \geq k$), this means that $m-k$ equations are a linear combination of only k of them, and so they can be discarded.

With regard to the search of solutions, we must proceed in the following way: after determining the common Rank k of the coefficient matrix and the augmented matrix, we use only k equations in k unknowns, provided they form a non-singular submatrix; the remaining $m-k$ equations, as mentioned earlier, are discarded; k unknowns remain as such while the remaining $n-k$ unknowns are to be brought with the constant terms, and then the system, which has now k equations and k unknowns, can be solved using Cramer's rule.
The value of each of the k remaining unknowns becomes a function of the other $n-k$ unknowns, which remain undetermined, but may take any real value. We write in this case that the system admits ∞^{n-k} solutions.
We can summarize the above as follows:
from $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$, if $k = \text{Rank}(\mathbb{A}) = \text{Rank}\left(\mathbb{A} \middle| \mathbb{Y}\right)$ is the common Rank of the coefficient matrix and the augmented matrix, after having discarded the $m-k$ useless equations and their constant terms, we obtain, using block notation:
$\mathbb{A}_{k,n} = \left[\mathbb{A}'_{k,k} \,\middle|\, \mathbb{A}''_{k,n-k}\right]$ and $\mathbb{X} = \left[\mathbb{X}_{k,1} \,\middle|\, \mathbb{X}_{n-k,1}\right]$ and so:
$\left[\mathbb{A}'_{k,k} \,\middle|\, \mathbb{A}''_{k,n-k}\right] \cdot \left[\mathbb{X}_{k,1} \,\middle|\, \mathbb{X}_{n-k,1}\right] = \mathbb{A}'_{k,k} \cdot \mathbb{X}_{k,1} + \mathbb{A}''_{k,n-k} \cdot \mathbb{X}_{n-k,1} = \mathbb{Y}_{k,1}$ and then:
$\mathbb{A}'_{k,k} \cdot \mathbb{X}_{k,1} = \mathbb{Y}_{k,1} - \mathbb{A}''_{k,n-k} \cdot \mathbb{X}_{n-k,1}$, and so the solution:
$\mathbb{X}_{k,1} = \left(\mathbb{A}'_{k,k}\right)^{-1} \cdot \left(\mathbb{Y}_{k,1} - \mathbb{A}''_{k,n-k} \cdot \mathbb{X}_{n-k,1}\right)$, where $\left(\mathbb{A}'_{k,k}\right)^{-1}$ is the inverse matrix of $\mathbb{A}'_{k,k}$, having as entries the coefficient of the remaining k unknows, that is surely a non-singular matrix by construction.
The solution, depending from the $n-k$ variables $\mathbb{X}_{n-k,1}$, is the general solution of the system.

If the system is an homogeneous one, it is $\mathbb{Y}_{k,1} = \mathbb{O}$, and we have the general solution in the form: $\mathbb{X}_{k,1} = -\left(\mathbb{A}'_{k,k}\right)^{-1} \cdot \mathbb{A}''_{k,n-k} \cdot \mathbb{X}_{n-k,1}$.

Example 67 : Let us study, varying the parameter k, the existence and number of solutions of the linear system: $\begin{cases} 3x_1 - x_2 + 2x_3 + x_4 = 7 \\ x_1 + x_2 - 4x_3 + 3x_4 = k \\ 5x_1 - 3x_2 + 8x_3 - x_4 = 2 \end{cases}$.

Let us first consider the coefficient matrix $\begin{Vmatrix} 3 & -1 & 2 & 1 \\ 1 & 1 & -4 & 3 \\ 5 & -3 & 8 & -1 \end{Vmatrix}$. Its Rank is equal to 2; in fact $\mathrm{R}_3 = 2\,\mathrm{R}_1 - \mathrm{R}_2$. In order for the system to have solutions we need this relationship also between the constant terms, so that the Rank of the augmented matrix is equal to 2. Therefore it should be: $14 - k = 2$, and so $k = 12$.
If $k \neq 12$ the system has no solutions.
If $k = 12$ the system is reduced to only two equations, and we can discard the third, since it is a linear combination of the first two.
Then we solve the system:

$$\begin{cases} 3x_1 - \ x_2 + 2x_3 + \ x_4 = 7 \\ x_1 + \ x_2 - 4x_3 + \ 3x_4 = 12 \end{cases}$$

in which we bring with the constant terms the unknowns x_3 and x_4 to obtain:

$$\begin{cases} 3x_1 - \ x_2 = \ -2x_3 - x_4 + 7 \\ x_1 + \ x_2 = 4x_3 - \ 3x_4 \ + 12 \end{cases}$$

and then, using Cramer's rule, and since $\begin{vmatrix} 3 & -1 \\ 1 & 1 \end{vmatrix} = 4$, we obtain:

$$\begin{cases} x_1 = \dfrac{1}{2}x_3 - x_4 + \dfrac{19}{4} \\ x_2 = \dfrac{7}{2}x_3 - 2x_4 + \dfrac{29}{4} \end{cases},$$

and finally the solution $\mathbb{X} = \left(\dfrac{1}{2}x_3 - x_4 + \dfrac{19}{4}; \dfrac{7}{2}x_3 - 2x_4 + \dfrac{29}{4}; x_3; x_4\right)$.

Example 68 : Let us determine for which values of the parameters m and k the linear system $\begin{cases} x + 2y + 3z = m \\ x + my + kz = k \end{cases}$ has solutions.

The coefficient matrix and the augmented matrix of the system are:

$\begin{Vmatrix} 1 & 2 & 3 \\ 1 & m & k \end{Vmatrix}$ and $\begin{Vmatrix} 1 & 2 & 3 & m \\ 1 & m & k & k \end{Vmatrix}$. The Rank of the coefficient matrix is equal to 2 if:

$\begin{vmatrix} 1 & 2 \\ 1 & m \end{vmatrix} \neq 0$, or if $\begin{vmatrix} 1 & 3 \\ 1 & k \end{vmatrix} \neq 0$, or if $\begin{vmatrix} 2 & 3 \\ m & k \end{vmatrix} \neq 0$, and so if:

$m - 2 \neq 0 \Rightarrow m \neq 2$, or if $k - 3 \neq 0 \Rightarrow k \neq 3$, or if $2k - 3m \neq 0 \Rightarrow k \neq \dfrac{3}{2}m$.

If these conditions are satisfied, the Rank of the coefficient matrix is maximum, and then it will be equal to that of the augmented matrix, and so the system has solutions.

Being a system of two equations and three unknowns, it has $\infty^{3-2} = \infty^1$ solutions.

If $m = 2$ and $k = 3$ (and so also $k = \dfrac{3}{2}m$), the Rank of the coefficient matrix cannot be equal to 2 (it is equal to 1, since there are entries different from 0); the augmented matrix is $\begin{Vmatrix} 1 & 2 & 3 & 2 \\ 1 & 2 & 3 & 3 \end{Vmatrix}$ whose Rank is equal to 2, since $\begin{vmatrix} 3 & 2 \\ 3 & 3 \end{vmatrix} \neq 0$ and so the system, if $m = 2$ and $k = 3$ has no solutions.

Example 69 : Let us solve the system $\begin{cases} x + 2y + z = 2 \\ 3x + y - 2z = 1 \\ 4x - 3y - z = 3 \\ 2x + 4y + 2z = 4 \end{cases}$.

The system has four equations and three unknowns, and so, if there are solutions, at least one equation must be a linear combination of the other. Since the fourth equation is equal to twice the first, it may be discarded.

The system is thus reduced to the following: $\begin{cases} x + 2y + z = 2 \\ 3x + y - 2z = 1 \\ 4x - 3y - z = 3 \end{cases}$.

Applying Cramer's rule, we calculate the Determinant and we obtain:

$$\begin{vmatrix} 1 & 2 & 1 \\ 3 & 1 & -2 \\ 4 & -3 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 1 \\ 3 & 1 & -2 \\ 5 & -1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 1 \\ 5 & 5 & 0 \\ 5 & -1 & 0 \end{vmatrix} = 1 \cdot \begin{vmatrix} 5 & 5 \\ 5 & -1 \end{vmatrix} = -30.$$

Therefore the system has one and only one solution, which is given by:

$$x=\frac{\begin{vmatrix}2&2&1\\1&1&-2\\3&-3&-1\end{vmatrix}}{-30}=\frac{\begin{vmatrix}2&0&1\\1&0&-2\\3&-6&-1\end{vmatrix}}{-30}=-\frac{1}{30}\cdot 6\cdot(-4-1)=1\,;$$

$$y=\frac{\begin{vmatrix}1&2&1\\3&1&-2\\4&3&-1\end{vmatrix}}{-30}=\frac{\begin{vmatrix}1&2&0\\3&1&-5\\4&3&-5\end{vmatrix}}{-30}=\frac{\begin{vmatrix}1&2&0\\3&1&-5\\1&2&0\end{vmatrix}}{-30}=0\,;$$

$$z=\frac{\begin{vmatrix}1&2&2\\3&1&1\\4&-3&3\end{vmatrix}}{-30}=\frac{\begin{vmatrix}-5&0&0\\3&1&1\\4&-3&3\end{vmatrix}}{-30}=-\frac{1}{30}\cdot(-5)\,(3+3)=1\,.$$

Example 70 : Let us find the value of the parameter k so that the vector $\mathbb{Y}=(1,7,2,k)$ is a linear combination of the vectors $\mathbb{X}_1=(1,-1,0,2)$ and $\mathbb{X}_2=(2,2,1,-1)$

First of all we see that $\mathbb{X}_1$ and $\mathbb{X}_2$ are linearly independent vectors, since the Rank of the matrix $\begin{Vmatrix}1&-1&0&2\\2&2&1&-1\end{Vmatrix}$ is equal to 2. Then we calculate, using the elementary operations on the rows, the Rank of the matrix $\begin{Vmatrix}1&-1&0&2\\2&2&1&-1\\1&7&2&k\end{Vmatrix}$. We obtain:

$$\begin{Vmatrix}1&-1&0&2\\2&2&1&-1\\1&7&2&k\end{Vmatrix}\to\begin{Vmatrix}1&-1&0&2\\0&4&1&-5\\0&8&2&k-2\end{Vmatrix}\to\begin{Vmatrix}1&-1&0&2\\0&4&1&-5\\0&0&0&k+8\end{Vmatrix}.$$

If $\mathbb{Y}$ is a linear combination of $\mathbb{X}_1$ and $\mathbb{X}_2$ the Rank of the latter matrix must remain equal to 2, and this happens if $k=-8$.

To apply the Rouchè-Capelli theorem we should have put the three vectors as columns, but we can also operate by rows, since $\text{Rank}(\mathbb{A})=\text{Rank}(\mathbb{A}^{\text{T}})$.

THE RELATIONSHIP BETWEEN THE SOLUTIONS OF A LINEAR SYSTEM AND THOSE OF THE ASSOCIATED HOMOGENEOUS SYSTEM

Given the linear system $\mathbb{A}_{m,n}\cdot\mathbb{X}_{n,1}=\mathbb{Y}_{m,1}$, with $\mathbb{Y}\neq\mathbb{O}$, we say that $\mathbb{A}_{m,n}\cdot\mathbb{X}_{n,1}=\mathbb{O}$ is its associated homogeneous system. The following is valid:

Theorem 16 : The linear system $\mathbb{A}_{m,n}\cdot\mathbb{X}_{n,1}=\mathbb{Y}_{m,1}$ has solutions if and only if each of its solutions $\mathbb{X}$ can be expressed as the sum of a particular solution $\mathbb{X}^*$ of the non-homogeneous system with the general solution $\mathbb{Z}$ of the associated homogeneous system.

Proof: If $\mathbb{X}$ and $\mathbb{X}^*$ are solutions of the linear system, from $\mathbb{A}\cdot\mathbb{X}=\mathbb{Y}$ and $\mathbb{A}\cdot\mathbb{X}^*=\mathbb{Y}$, by subtraction we obtain $\mathbb{A}\cdot(\mathbb{X}-\mathbb{X}^*)=\mathbb{O}$, and so $\mathbb{X}-\mathbb{X}^*$ is a solution of the associated homogeneous system, therefore $\mathbb{X}-\mathbb{X}^*=\mathbb{Z}$ and so $\mathbb{X}=\mathbb{X}^*+\mathbb{Z}$.

On the contrary, if $\mathbb{X}=\mathbb{X}^*+\mathbb{Z}$, where $\mathbb{X}^*$ is a solution of the non-homogeneous system and $\mathbb{Z}$ is the general solution of the associated homogeneous system, we obtain:

$\mathbb{A}\cdot(\mathbb{X}^*+\mathbb{Z})=\mathbb{A}\cdot\mathbb{X}^*+\mathbb{A}\cdot\mathbb{Z}=\mathbb{Y}+\mathbb{O}=\mathbb{Y}$,

and so $\mathbb{X}$ is a solution of the non-homogeneous system. ●

Example 71 : Let us study the solvability of the system:

$$\begin{cases} x_1+x_2-2x_3+x_4+3x_5=1 \\ 2x_1-x_2+2x_3+2x_4+6x_5=2 \\ 4x_1+x_2-2x_3+4x_4+2x_5=0 \end{cases}.$$

Examining the coefficient matrix $\begin{Vmatrix} 1 & 1 & -2 & 1 & 3 \\ 2 & -1 & 2 & 2 & 6 \\ 4 & 1 & -2 & 4 & 2 \end{Vmatrix}$, if we study the four Minors of order 3 that can be constructed using the first 4 columns, we see that they all have Determinant equal to 0. Considering the Minor of order 3 constituted with the first, second and fifth column, we have:

$$\begin{vmatrix} 1 & 1 & 3 \\ 2 & -1 & 6 \\ 4 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 2 & -1 & 0 \\ 4 & 1 & -10 \end{vmatrix} = -10 \cdot \begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} = -10 \cdot (-3) = 30 \neq 0\,.$$

The coefficient matrix has Rank equal to 3, i.e. maximum, therefore also the augmented matrix has Rank equal to 3, so the system has solutions, and these solutions can be obtained bringing the unknowns x_3 and x_4 to constant terms, and so we obtain:

$$\begin{cases} x_1+x_2+3x_5= \ 1+2x_3-x_4 \\ 2x_1-x_2+6x_5= \ 2-2x_3-2x_4 \\ 4x_1+x_2+2x_5= \ 2x_3-4x_4 \end{cases}$$, from which, using Cramer's rule, we obtain:

$$x_1 = \frac{1}{30}\begin{vmatrix} 1+2x_3-x_4 & 1 & 3 \\ 2-2x_3-2x_4 & -1 & 6 \\ 2x_3-4x_4 & 1 & 2 \end{vmatrix} = -x_4-\frac{1}{5}\,,$$

$$x_2 = \frac{1}{30}\begin{vmatrix} 1 & 1+2x_3-x_4 & 3 \\ 2 & 2-2x_3-2x_4 & 6 \\ 4 & 2x_3-4x_4 & 2 \end{vmatrix} = 2x_3\,,$$

$$x_5 = \frac{1}{30}\begin{vmatrix} 1 & 1 & 1+2x_3-x_4 \\ 2 & -1 & 2-2x_3-2x_4 \\ 4 & 1 & 2x_3-4x_4 \end{vmatrix} = \frac{2}{5}\,,$$

and so the system has ∞^2 solutions: $\mathbb{X} = \left(-x_4-\frac{1}{5}\,, 2x_3\,, x_3\,, x_4\,, \frac{2}{5}\right)$.

Let us consider the associated homogeneous system $\begin{cases} x_1+x_2-2x_3+x_4+3x_5=0 \\ 2x_1-x_2+2x_3+2x_4+6x_5=0 \\ 4x_1+x_2-2x_3+4x_4+2x_5=0 \end{cases}$;

for its coefficient matrix the same considerations previously made are valid, so it has the solution:

$$x_1 = \frac{1}{30}\begin{vmatrix} 2x_3-x_4 & 1 & 3 \\ -2x_3-2x_4 & -1 & 6 \\ 2x_3-4x_4 & 1 & 2 \end{vmatrix} = -x_4\,,$$

$$x_2 = \frac{1}{30}\begin{vmatrix} 1 & 2x_3-x_4 & 3 \\ 2 & -2x_3-2x_4 & 6 \\ 4 & 2x_3-4x_4 & 2 \end{vmatrix} = 2x_3\,,$$

$$x_5 = \frac{1}{30}\begin{vmatrix} 1 & 1 & 2x_3-x_4 \\ 2 & -1 & -2x_3-2x_4 \\ 4 & 1 & 2x_3-4x_4 \end{vmatrix} = 0\,.$$

So the general solution of the associated homogeneous system is $\mathbb{Z} = (-x_4, 2x_3, x_3, x_4, 0)$.

Having verified that $\mathbb{X}^* = \left(-\dfrac{1}{5}, 0, 0, 0, \dfrac{2}{5}\right)$ is a particular solution of the system, then the equality: $\mathbb{X} = \mathbb{X}^* + \mathbb{Z}$ is satisfied.

KERNEL AND IMAGE OF A LINEAR MAP

Definition 38 : The Image (or Range): $\mathrm{Im}(\mathbb{A})$ of a linear map $\mathbb{Y} = \mathbb{A} \cdot \mathbb{X}$ is the set:
$\mathrm{Im}(\mathbb{A}) = \{\mathbb{Y} \in \mathbb{R}^m \colon \exists\, \mathbb{X} \in \mathbb{R}^n, \mathbb{Y} = \mathbb{A} \cdot \mathbb{X}\}$.
Therefore the Image is a subset of the codomain of the linear map.

Theorem 17 : The Image of a linear map is a vector space whose dimension is equal to the Rank of the matrix $\mathbb{A}_{m,n}$.
Proof: If $\mathbb{Y}_1 \in \mathrm{Im}(\mathbb{A})$ and $\mathbb{Y}_2 \in \mathrm{Im}(\mathbb{A})$, it is $\mathbb{Y}_1 = \mathbb{A} \cdot \mathbb{X}_1$ and $\mathbb{Y}_2 = \mathbb{A} \cdot \mathbb{X}_2$, from which it follows, for the properties of the matrix product:
$k_1\, \mathbb{Y}_1 + k_2\, \mathbb{Y}_2 = k_1\, \mathbb{A} \cdot \mathbb{X}_1 + k_2\, \mathbb{A} \cdot \mathbb{X}_2 = \mathbb{A} \cdot (k_1\, \mathbb{X}_1 + k_2\, \mathbb{X}_2)$,
and so, since $k_1\, \mathbb{X}_1 + k_2\, \mathbb{X}_2 \in \mathbb{R}^n$, it follows that $k_1\, \mathbb{Y}_1 + k_2\, \mathbb{Y}_2 \in \mathrm{Im}(\mathbb{A})$, i.e. the Image (or Range) of a linear map is a vector space.
Considering the linear map as a linear combination of the columns of the matrix $\mathbb{A}$, from:
$\mathbb{Y} = \mathrm{C}_1 \cdot x_1 + \mathrm{C}_2 \cdot x_2 + \ \dots\ + \mathrm{C}_n \cdot x_n$, we immediately see that the Range is the vector space spanned by the columns of $\mathbb{A}$, and so its dimension is given by the maximum number of independent columns of $\mathbb{A}$, i.e. by Rank($\mathbb{A}$). •

Definition 39 : The Kernel (or nullspace): $\mathrm{Ker}(\mathbb{A})$ of a linear map $\mathbb{Y} = \mathbb{A} \cdot \mathbb{X}$ is the set:
$\mathrm{Ker}(\mathbb{A}) = \{\mathbb{X} \in \mathbb{R}^n \colon \mathbb{A} \cdot \mathbb{X} = \mathbb{O}\}$,
i.e. the set of the vectors of the domain having as image the null vector.

Theorem 18: The Kernel of a linear map is a vector subspace of the domain $\mathbb{R}^n$.
Proof: If $\mathbb{X}_1$ and $\mathbb{X}_2$ belong to the Kernel, since:
$\mathbb{A} \cdot (k_1\, \mathbb{X}_1 + k_2\, \mathbb{X}_2) = k_1\, \mathbb{A} \cdot \mathbb{X}_1 + k_2\, \mathbb{A} \cdot \mathbb{X}_2 = \mathbb{O} + \mathbb{O} = \mathbb{O}$,
we immediately see that the Kernel is a vector space, more exactly a vector subspace of the domain $\mathbb{R}^n$. •

We must notice the similarities between the Kernel of a linear map and the solutions of a linear homogeneous system, and also the similarities between the Image of a linear map and the linear non-homogeneous systems that have solutions.

Also the following is valid:
Theorem 19 : In a linear map, linearly dependent vectors have linearly dependent images.
Proof: From $k_1 \mathbb{X}_1 + k_2\, \mathbb{X}_2 + \dots + k_p\, \mathbb{X}_p = \mathbb{O}$, with at least one $k_i \neq 0$, it follows:
$\mathbb{A} \cdot (k_1 \mathbb{X}_1 + k_2\, \mathbb{X}_2 + \dots + k_p\, \mathbb{X}_p) = \mathbb{A} \cdot k_1 \mathbb{X}_1 + \mathbb{A} \cdot k_2\, \mathbb{X}_2 \dots + \mathbb{A} \cdot k_p\, \mathbb{X}_p = \mathbb{O}$ from which:
$k_1\, \mathbb{A} \cdot \mathbb{X}_1 + k_2\, \mathbb{A} \cdot \mathbb{X}_2 \dots + k_p\, \mathbb{A} \cdot \mathbb{X}_p = \mathbb{O}$ or:
$k_1\, f(\mathbb{X}_1) + k_2\, f(\mathbb{X}_2) \dots + k_p\, f(\mathbb{X}_p) = \mathbb{O}$, with at least one $k_i \neq 0$ and so the images are linearly dependent vectors. •

Consequently, also the following is valid:
Theorem 20 : In a linear map, linearly independent images are generated by linearly independent vectors.

RELATIONSHIP BETWEEN KERNEL AND IMAGE DIMENTIONS: RANK-NULLITY THEOREM

Given the linear map $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$, if $\mathbb{Y} \in \text{Im}(\mathbb{A})$ surely $\text{Rank}(\mathbb{A}) = \text{Rank}\big(\mathbb{A} \,\big|\, \mathbb{Y}\big)$.
We have seen that the dimension of the Image of a linear map is equal to the Rank of the matrix $\mathbb{A}_{m,n}$: $\text{Dim}(\text{Im}(\mathbb{A})) = \text{Rank}(\mathbb{A}) = k$.
Considering the associated homogeneous system $(\mathbb{Y}_{k,1} = \mathbb{O})$ we have got its solution:
$\mathbb{X}_{k,1} = -\left(\mathbb{A}'_{k,k}\right)^{-1} \cdot \mathbb{A}''_{k,n-k} \cdot \mathbb{X}_{n-k,1}$,
from which we see that $\text{Dim}(\text{Ker}(\mathbb{A})) = n - k$, i.e. the dimension of the Kernel is equal to the number of unknowns brought with the constant terms, i.e. the number of variables which remain independent.
The dimension of the Kernel of $\mathbb{A}$ is also called the nullity of $\mathbb{A}$.

The following theorem summarizes what we have already said:
Theorem 21 ("rank-nullity theorem" or "Sylvester's theorem" or "Image theorem") :
The Rank and the nullity of a matrix add up to the number of columns of the matrix itself, i.e.: adding the dimensions of the Image and the dimensions of the Kernel of a linear map we obtain the dimension of the domain, i.e.:
$\text{Dim}(\text{Im}(\mathbb{A})) + \text{Dim}(\text{Ker}(\mathbb{A})) = \text{Dim}(\mathbb{R}^n)$ or:
$\text{Dim}(\text{Im}(\mathbb{A})) + \text{Dim}(\text{Ker}(\mathbb{A})) =$ number of the columns of $\mathbb{A}$.
In fact, as previously seen, $k + (n - k) = n$.

Example 72 : Given the linear map $\mathbb{A} \cdot \mathbb{X} = \mathbb{Y}$, with $\mathbb{A} = \left\|\begin{matrix} k & 2 & k \\ -1 & 0 & -3 \\ 1 & -2 & 1 \end{matrix}\right\|$, let us determine, varying the parameter k, the dimension of the Kernel and the dimension of the Image of the linear map.
Since $\text{Det}(\mathbb{A}) = -4(1+k)$, it follows that:
- if $k = -1$, $\text{Rank}(\mathbb{A}) = 2 \Rightarrow \text{Dim}(\text{Im}) = 2$ and $\text{Dim}(\text{Ker}) = 3 - 2 = 1$;
- if $k \neq -1$, $\text{Rank}(\mathbb{A}) = 3 \Rightarrow \text{Dim}(\text{Im}) = 3$ and $\text{Dim}(\text{Ker}) = 3 - 3 = 0$.

SURJECTIVE, INJECTIVE, BIJECTIVE AND INVERTIBLE LINEAR MAPS
A function is surjective (onto) if every element of the codomain is mapped to by at least one element of the domain. That is, the image and the codomain of the function are the same set.
A function is injective (one-to-one) if every element of the codomain is mapped to by at most one element of the domain.
A function is bijective (one-to-one and onto or one-to-one correspondence) if every element of the codomain is mapped to by exactly one element of the domain. That is, the function is both injective and surjective.
A function is an invertible one if it is both injective and surjective, i.e. if it is bijective.
Since linear maps are functions $\mathbb{R}^n \longrightarrow \mathbb{R}^m$, we study the problem of the existence of their inverse function, which requires first to check when a linear map is surjective and when it is injective.

The linear map: $\mathbb{R}^n \longrightarrow \mathbb{R}^m$, $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$ is a surjective map if its Image (Range) is equal to the whole $\mathbb{R}^m$, i.e. if $\text{Dim}(\text{Im}(\mathbb{A})) = m$.
Therefore it must be $\text{Dim}(\text{Im}(\mathbb{A})) = \text{Rank}(\mathbb{A}) = m$ ($\leq n$ number of the columns of $\mathbb{A}$).

The linear map: $\mathbb{R}^n \longrightarrow \mathbb{R}^m$, $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$ is an injective map if:
$f(\mathbb{X}_1) = f(\mathbb{X}_2) \Rightarrow \mathbb{X}_1 = \mathbb{X}_2$.
But $f(\mathbb{X}_1) = f(\mathbb{X}_2) \Rightarrow f(\mathbb{X}_1 - \mathbb{X}_2) = \mathbb{O}$, i.e. $\mathbb{X}_1 - \mathbb{X}_2 \in \text{Ker}(\mathbb{A})$.

And so $\mathbb{X}_1 = \mathbb{X}_2$ if and only if $\text{Ker}(\mathbb{A}) = \{\mathbb{O}\}$, i.e. $\text{Dim}(\text{Ker}(\mathbb{A})) = 0$, i.e.:
$(n-k) = 0 \Rightarrow \text{Rank}(\mathbb{A}) = k = n$ ($\leq m$ number of the rows of $\mathbb{A}$).

The linear map: $\mathbb{R}^n \longrightarrow \mathbb{R}^m$, $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$ is an invertible map if it is surjective and injective, and as previously seen it must be: $\text{Rank}(\mathbb{A}) = k = m = n$, and then the matrix $\mathbb{A}$ must be a square and non-singular matrix.
From $\mathbb{A}_n \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{n,1}$, if $\mathbb{A}$ is a square and non-singular matrix, we obtain $\mathbb{X}_{n,1} = \mathbb{A}_n^{-1} \cdot \mathbb{Y}_{n,1}$, and so $\mathbb{Y}_{n,1} = \mathbb{A}_n^{-1} \cdot \mathbb{X}_{n,1}$, which is the inverse map of the linear map.

Example 73 : For the linear map $\mathbb{A} \cdot \mathbb{X} = \mathbb{Y}$, with $\mathbb{A} = \left\| \begin{matrix} k & 2 & k \\ -1 & 0 & -3 \\ 1 & -2 & 1 \end{matrix} \right\|$, as seen in Example 72, we have:
-if $k = -1$, $\text{Dim}(\text{Im}) = 2$ and $\text{Dim}(\text{Ker}) = 1$, and so the map is neither injective nor surjective;
-if $k \neq -1$, $\text{Dim}(\text{Im}) = 3$ and $\text{Dim}(\text{Ker}) = 0$, and so the map is injective and surjective, and then invertible.

COMPOSITION OF LINEAR MAPS
Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{m,1}$ and $g : \mathbb{R}^m \longrightarrow \mathbb{R}^p$, $\mathbb{B}_{p,m} \cdot \mathbb{Y}_{m,1} = \mathbb{Z}_{p,1}$. We see that the composition of two linear maps is still a linear map.
In fact, if $\mathbb{Y} = f(\mathbb{X}) = \mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1}$ and $\mathbb{Z} = g(\mathbb{Y}) = \mathbb{B}_{p,m} \cdot \mathbb{Y}_{m,1}$, it is:
$\mathbb{Z} = g(\mathbb{Y}) = g(f(\mathbb{X})) = \mathbb{B}_{p,m} \cdot \mathbb{Y}_{m,1} = \mathbb{B}_{p,m} \cdot (\mathbb{A}_{m,n} \cdot \mathbb{X}_{n,1}) = (\mathbb{B}_{p,m} \cdot \mathbb{A}_{m,n}) \cdot \mathbb{X}_{n,1}$.
Then the composed linear map is given by $g \circ f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, $\mathbb{C}_{p,n} \cdot \mathbb{X}_{n,1} = \mathbb{Z}_{p,1}$, with $\mathbb{C}_{p,n} = \mathbb{B}_{p,m} \cdot \mathbb{A}_{m,n}$.
The composition of two or more linear maps is still a linear map having for matrix the product of the matrices of the composed linear maps.

Example 74 : Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$, $\mathbb{A}_{4,2} \cdot \mathbb{X}_{2,1} = \mathbb{Y}_{4,1}$ and $g : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$, $\mathbb{B}_{3,4} \cdot \mathbb{Y}_{4,1} = \mathbb{Z}_{3,1}$, with: $\mathbb{A}_{4,2} = \left\| \begin{matrix} 1 & 0 \\ 0 & 2 \\ 1 & 1 \\ 2 & 1 \end{matrix} \right\|$ and $\mathbb{B}_{3,4} = \left\| \begin{matrix} 0 & 1 & 2 & 1 \\ 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 2 \end{matrix} \right\|$. The composition of these linear maps is
$g \circ f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$, $\mathbb{C}_{3,2} \cdot \mathbb{X}_{2,1} = \mathbb{Z}_{3,1}$, with

$$\mathbb{C}_{3,2} = \mathbb{B}_{3,4} \cdot \mathbb{A}_{4,2} = \left\| \begin{matrix} 0 & 1 & 2 & 1 \\ 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 2 \end{matrix} \right\| \cdot \left\| \begin{matrix} 1 & 0 \\ 0 & 2 \\ 1 & 1 \\ 2 & 1 \end{matrix} \right\| = \left\| \begin{matrix} 4 & 5 \\ 1 & 2 \\ 7 & 5 \end{matrix} \right\|.$$

ORTHONORMAL BASES - ORTHOGONAL MATRICES
Two vectors are orthogonal if and only if their scalar (or dot) product is equal to 0. A vector whose modulus (or length) is equal to 1 is a unit vector (or versor or normalized vector).
A set of vectors form an orthonormal set if all the vectors in the set are unit vectors and are mutually orthogonal.
Definition 40 : An orthonormal basis for a vector space $\mathbb{R}^n$ is a basis for $\mathbb{R}^n$ whose vectors are orthonormal vectors.

If $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_n$ form an orthonormal basis, since they are mutually orthogonal unit vectors, it follows that: $\mathbb{X}_i \cdot \mathbb{X}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$.

Definition 41 : A matrix $\mathbb{A}$ is orthogonal if $\mathbb{A} \cdot \mathbb{A}^{\mathrm{T}} = \mathbb{A}^{\mathrm{T}} \cdot \mathbb{A} = \mathbb{I}$.
This definition requires that the matrix is a square and non-singular matrix, so an invertible matrix; since $\mathbb{A}^{-1}$ is the only matrix for which the property $\mathbb{A}^{-1} \cdot \mathbb{A} = \mathbb{A} \cdot \mathbb{A}^{-1} = \mathbb{I}$ applies, it follows that a matrix $\mathbb{A}$ is orthogonal if its transpose is equal to its inverse: $\mathbb{A}^{\mathrm{T}} = \mathbb{A}^{-1}$.
From this definition it also follows that the rows (and the columns) of an orthogonal matrix form an orthonormal basis, i.e. its rows and its columns are mutually orthogonal unit vectors.

It is easily seen that all unit matrices $\mathbb{I}_n$ and all permutation matrices are orthogonal matrices.

For orthogonal matrices the following properties are valid:
O1) If $\mathbb{A}_n$ and $\mathbb{B}_n$ are orthogonal matrices, also $\mathbb{A} \cdot \mathbb{B}$ and $\mathbb{B} \cdot \mathbb{A}$ are orthogonal matrices:
in fact $(\mathbb{A} \cdot \mathbb{B})^{\mathrm{T}} \cdot (\mathbb{A} \cdot \mathbb{B}) = \mathbb{B}^{\mathrm{T}} \cdot \mathbb{A}^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{B} = \mathbb{B}^{\mathrm{T}} \cdot \mathbb{I} \cdot \mathbb{B} = \mathbb{I}$, so the thesis;
O2) If $\mathbb{A}$ is an orthogonal matrix, then $|\mathbb{A}| = \pm 1$:
in fact, from $\left|\mathbb{A} \cdot \mathbb{A}^{\mathrm{T}}\right| = |\mathbb{I}|$, we obtain $|\mathbb{A}| \cdot \left|\mathbb{A}^{\mathrm{T}}\right| = |\mathbb{A}|^2 = |\mathbb{I}| = 1$ and so $|\mathbb{A}| = \pm 1$;
O3) If $\mathbb{A}$ is an orthogonal matrix, also $\mathbb{A}^{\mathrm{T}}$ and $\mathbb{A}^{-1}$ are orthogonal matrices:
in fact $\mathbb{A}^{\mathrm{T}} \cdot \left(\mathbb{A}^{\mathrm{T}}\right)^{\mathrm{T}} = \mathbb{A}^{\mathrm{T}} \cdot \mathbb{A} = \mathbb{I}$ while
$\mathbb{A}^{-1} \cdot \left(\mathbb{A}^{-1}\right)^{\mathrm{T}} = \mathbb{A}^{-1} \cdot \left(\mathbb{A}^{\mathrm{T}}\right)^{-1} = \mathbb{A}^{\mathrm{T}} \cdot \left(\mathbb{A}^{\mathrm{T}}\right)^{-1} = \mathbb{I}$;
O4) If $\mathbb{A}$ is an orthogonal matrix, also $\mathbb{A}^k$ is an orthogonal matrix:
in fact: $\left(\mathbb{A}^k\right)^{\mathrm{T}} \cdot \mathbb{A}^k = \left(\mathbb{A}^{\mathrm{T}}\right)^k \cdot \mathbb{A}^k = \mathbb{A}^{\mathrm{T}} \cdot \mathbb{A}^{\mathrm{T}} \cdot ... \cdot \left(\mathbb{A}^{\mathrm{T}} \cdot \mathbb{A}\right) \cdot ... \cdot \mathbb{A} \cdot \mathbb{A} = \mathbb{I}$.

ORTHONORMAL BASES: GRAM-SCHMIDT PROCESS

If $\{\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_n\}$ is a basis of $\mathbb{R}^n$, but conditions $\mathbb{X}_i \cdot \mathbb{X}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ are not satisfied, the Gram-Schmidt process allows us to construct, starting from $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_n$, an orthonormal basis.

We start with the first vector of the basis, $\mathbb{X}_1$, and we set $\mathbb{Z}_1 = \mathbb{X}_1$.
The unit vector $\mathbb{W}_1 = \dfrac{\mathbb{Z}_1}{\|\mathbb{Z}_1\|}$ is the first vector of the orthonormal basis.
To find the second vector, $\mathbb{W}_2$, first we determine a vector $\mathbb{Z}_2$, expressed in the form:
$\mathbb{Z}_2 = \mathbb{X}_2 - \alpha\, \mathbb{W}_1$,
i.e. the second vector of the original basis minus a multiple of the first vector found for the orthonormal basis.
$\mathbb{Z}_2$ has to be orthogonal to $\mathbb{W}_1$, and so we need:
$\mathbb{Z}_2 \cdot \mathbb{W}_1 = (\mathbb{X}_2 - \alpha\, \mathbb{W}_1) \cdot \mathbb{W}_1 = \mathbb{X}_2 \cdot \mathbb{W}_1 - \alpha\, \mathbb{W}_1 \cdot \mathbb{W}_1 = 0$,
from which, since $\mathbb{W}_1 \cdot \mathbb{W}_1 = 1$, we obtain:
$\alpha = \mathbb{X}_2 \cdot \mathbb{W}_1 = \ < \mathbb{X}_2, \mathbb{W}_1 >$ in order to obtain:
$\mathbb{Z}_2 = \mathbb{X}_2 - \ < \mathbb{X}_2, \mathbb{W}_1 > \ \cdot \mathbb{W}_1$.
Then the second vector of the orthonormal basis is $\mathbb{W}_2 = \dfrac{\mathbb{Z}_2}{\|\mathbb{Z}_2\|}$.
To determine the third vector, we set:
$\mathbb{Z}_3 = \mathbb{X}_3 - \alpha\, \mathbb{W}_1 - \beta\, \mathbb{W}_2$.
$\mathbb{Z}_3$ has to be orthogonal to $\mathbb{W}_1$ and $\mathbb{W}_2$ and so:

$$\begin{cases} \mathbb{Z}_3 \cdot \mathbb{W}_1 = (\mathbb{X}_3 - \alpha\,\mathbb{W}_1 - \beta\,\mathbb{W}_2) \cdot \mathbb{W}_1 = \mathbb{X}_3 \cdot \mathbb{W}_1 - \alpha\,\mathbb{W}_1 \cdot \mathbb{W}_1 - \beta\,\mathbb{W}_2 \cdot \mathbb{W}_1 \\ \mathbb{Z}_3 \cdot \mathbb{W}_2 = (\mathbb{X}_3 - \alpha\,\mathbb{W}_1 - \beta\,\mathbb{W}_2) \cdot \mathbb{W}_2 = \mathbb{X}_3 \cdot \mathbb{W}_2 - \alpha\,\mathbb{W}_1 \cdot \mathbb{W}_2 - \beta\,\mathbb{W}_2 \cdot \mathbb{W}_2 \end{cases}.$$

Since $\mathbb{W}_1 \cdot \mathbb{W}_1 = \mathbb{W}_2 \cdot \mathbb{W}_2 = 1$ and $\mathbb{W}_1 \cdot \mathbb{W}_2 = \mathbb{W}_2 \cdot \mathbb{W}_1 = 0$, we obtain:

$$\begin{cases} \alpha = \mathbb{X}_3 \cdot \mathbb{W}_1 = \ < \mathbb{X}_3, \mathbb{W}_1 > \\ \beta = \mathbb{X}_3 \cdot \mathbb{W}_2 = \ < \mathbb{X}_3, \mathbb{W}_2 > \end{cases} \text{ and so:}$$

$\mathbb{Z}_3 = \mathbb{X}_3 - \ < \mathbb{X}_3, \mathbb{W}_1 > \ \cdot \mathbb{W}_1 - \ < \mathbb{X}_3, \mathbb{W}_2 > \ \cdot \mathbb{W}_2$.

Then the third vector of the orthonormal basis is $\mathbb{W}_3 = \dfrac{\mathbb{Z}_3}{\|\mathbb{Z}_3\|}$.

Generalizing the process, the i-th vector of the orthonormal basis is $\mathbb{W}_i = \dfrac{\mathbb{Z}_i}{\|\mathbb{Z}_i\|}$, with

$\mathbb{Z}_i = \mathbb{X}_i - \ < \mathbb{X}_i, \mathbb{W}_1 > \ \cdot \mathbb{W}_1 - \ < \mathbb{X}_i, \mathbb{W}_2 > \ \cdot \mathbb{W}_2 - ... - \ < \mathbb{X}_i, \mathbb{W}_{i-1} > \ \cdot \mathbb{W}_{i-1}$.

Example 75 : Let us consider the basis of $\mathbb{R}^3$ consisting of these three vectors:
$\mathbb{X}_1 = (0,1,1)$, $\mathbb{X}_2 = (1,0,1)$, $\mathbb{X}_3 = (1,1,0)$ and we want to determine, starting from this, an orthonormal basis. Obviously, we have checked that the three vectors are linearly independent and that they are not perpendicular two by two.

$\mathbb{W}_1 = \dfrac{\mathbb{Z}_1}{\|\mathbb{Z}_1\|} = \dfrac{\mathbb{X}_1}{\|\mathbb{X}_1\|} = \left(0, \dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right)$ is the first vector of the orthonormal basis.

From $\mathbb{Z}_2 = \mathbb{X}_2 - \ < \mathbb{X}_2, \mathbb{W}_1 > \ \cdot \mathbb{W}_1$ we obtain, since $< \mathbb{X}_2, \mathbb{W}_1 > \ = \dfrac{1}{\sqrt{2}}$:

$\mathbb{Z}_2 = (1,0,1) - \dfrac{1}{\sqrt{2}}\left(0, \dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right) = \left(1, -\dfrac{1}{2}, \dfrac{1}{2}\right)$, and since $\|\mathbb{Z}_2\| = \dfrac{\sqrt{3}}{\sqrt{2}}$ we obtain:

$\mathbb{W}_2 = \dfrac{\mathbb{Z}_2}{\|\mathbb{Z}_2\|} = \left(\dfrac{2}{\sqrt{6}}, -\dfrac{1}{\sqrt{6}}, \dfrac{1}{\sqrt{6}}\right)$.

Finally, to find the third vector, we calculate first:
$\mathbb{Z}_3 = \mathbb{X}_3 - \ < \mathbb{X}_3, \mathbb{W}_1 > \ \cdot \mathbb{W}_1 - \ < \mathbb{X}_3, \mathbb{W}_2 > \ \cdot \mathbb{W}_2$ and since:
$< \mathbb{X}_3, \mathbb{W}_1 > \ = \dfrac{1}{\sqrt{2}}$ and $< \mathbb{X}_3, \mathbb{W}_2 > \ = \dfrac{1}{\sqrt{6}}$ we obtain:

$$\mathbb{Z}_3 = (1,1,0) - \frac{1}{\sqrt{2}}\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) - \frac{1}{\sqrt{6}}\left(\frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) = \left(\frac{2}{3}, \frac{2}{3}, -\frac{2}{3}\right)$$

from which, since $\|\mathbb{Z}_3\| = \dfrac{2}{\sqrt{3}}$, we obtain $\mathbb{W}_3 = \left(\dfrac{1}{\sqrt{3}}, \dfrac{1}{\sqrt{3}}, -\dfrac{1}{\sqrt{3}}\right)$.

The vectors:

$$\{\mathbb{W}_1, \mathbb{W}_2, \mathbb{W}_3\} = \left\{\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right), \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)\right\}$$

constitute an orthonormal basis.

CHANGE OF BASIS

To express every element of a vector space we need a basis for this space. We have seen that every vector of $\mathbb{R}^n$ is expressed in one and only one way as a linear combination of the n elements of the chosen basis.

If not otherwise specified, writing $\mathbb{X} = (x_1, x_2, ..., x_n)$ we mean the vector $\mathbb{X}$ expressed under the standard basis, i.e. $\mathbb{X} = \mathbb{X}_e = x_1\,\mathbf{e}_1 + x_2\,\mathbf{e}_2 + ... + x_n\,\mathbf{e}_n$.

If we choose any other basis, always formed by n linearly independent vectors, we want to find the coordinates of the vector $\mathbb{X}$ under the new basis.

The problem can be reduced to linear maps and their appropriate compositions.
Being $\mathbb{X} = \mathbb{I}_n \cdot \mathbb{X}$, since the columns of $\mathbb{I}_n$ are the elements of the standard basis, we see that every vector $\mathbb{X} \in \mathbb{R}^n$ can be written as a product of a suitable matrix, which, in the case of the standard basis, is the unit matrix $\mathbb{I}_n$, multiplied by the column vector $\mathbb{X}$ formed by its coordinates under the chosen basis.
If we choose a basis other than the standard, for example $\{\mathbb{W}_1, \mathbb{W}_2, ..., \mathbb{W}_n\}$, and if the vector $\mathbb{X}$ has, under this basis, coordinates $\mathbb{X}_{\mathrm{w}} = (\alpha_1, \alpha_2, ..., \alpha_n)$, it is:
$\mathbb{X}_e = \alpha_1 \mathbb{W}_1 + \alpha_2 \mathbb{W}_2 + ... + \alpha_n \mathbb{W}_n$,
that can be written as:

$$\mathbb{X}_e = \mathbb{W} \cdot \mathbb{X}_{\mathrm{w}} = \left[\mathbb{W}_1 \middle| \mathbb{W}_2 \middle| ... \middle| \mathbb{W}_n\right] \cdot \begin{Vmatrix} \alpha_1 \\ \alpha_2 \\ ... \\ \alpha_n \end{Vmatrix},$$

i.e. the product of a matrix $\mathbb{W}$, whose columns are the vectors $\mathbb{W}_1, \mathbb{W}_2, ..., \mathbb{W}_n$ of the basis, by the column vector $\mathbb{X}_{\mathrm{w}}$ formed by the coordinates of $\mathbb{X}_e$ under the new basis.
Vectors $\mathbb{W}_1, \mathbb{W}_2, ..., \mathbb{W}_n$ are a basis, so they are linearly independent vectors, and so the matrix $\mathbb{W}$ is non-singular and then invertible.
Being $\mathbb{W}^{-1} \cdot \mathbb{X}_e = \mathbb{X}_{\mathrm{w}}$, by means of the product $\mathbb{W}^{-1} \cdot \mathbb{X}_e$ we obtain the coordinates $\mathbb{X}_{\mathrm{w}}$ of the vector $\mathbb{X}$ under the new basis $\{\mathbb{W}_1, \mathbb{W}_2, ..., \mathbb{W}_n\}$.
If we want to change any vector of the space to this new basis $\{\mathbb{W}_1, \mathbb{W}_2, ..., \mathbb{W}_n\}$, we only need to left-multiply its components (or coordinates) by the inverse of the matrix $\mathbb{W}$.
The matrix $\mathbb{W}^{-1}$ is said the transition matrix (or change-of-basis matrix) from the standard basis to the new basis.
Therefore a change of basis is a linear invertible map $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ from a vector space into itself.

If we choose a further new basis for $\mathbb{R}^n$: $\{\mathbb{V}_1, \mathbb{V}_2, ..., \mathbb{V}_n\}$ and if we want to express the vector $\mathbb{X}_e$ under this basis, operating as in the previous case we have:
$\mathbb{X}_e = \mathbb{V} \cdot \mathbb{X}_{\mathrm{v}}$ and then $\mathbb{V}^{-1} \cdot \mathbb{X}_e = \mathbb{X}_{\mathrm{v}}$,
where $\mathbb{V}$ is the matrix having as columns the vectors $\{\mathbb{V}_1, \mathbb{V}_2, ..., \mathbb{V}_n\}$.
And so, since: $\mathbb{X}_e = \mathbb{W} \cdot \mathbb{X}_{\mathrm{w}} = \mathbb{V} \cdot \mathbb{X}_{\mathrm{v}}$ we obtain, premultiplying properly:
$\mathbb{X}_{\mathrm{w}} = \mathbb{W}^{-1} \cdot \mathbb{V} \cdot \mathbb{X}_{\mathrm{v}}$ or $\mathbb{X}_{\mathrm{v}} = \mathbb{V}^{-1} \cdot \mathbb{W} \cdot \mathbb{X}_{\mathrm{w}}$
to obtain directly the transition from the coordinates $\mathbb{X}_{\mathrm{v}}$ to the coordinates $\mathbb{X}_{\mathrm{w}}$ and vice versa.
The matrix $\mathbb{P} = \mathbb{W}^{-1} \cdot \mathbb{V}$ is the transition matrix from the basis $\{\mathbb{V}_1, \mathbb{V}_2, ..., \mathbb{V}_n\}$ to the basis $\{\mathbb{W}_1, \mathbb{W}_2, ..., \mathbb{W}_n\}$.
The matrix $\mathbb{S} = \mathbb{V}^{-1} \cdot \mathbb{W}$ is the transition matrix from the basis $\{\mathbb{W}_1, \mathbb{W}_2, ..., \mathbb{W}_n\}$ to the basis $\{\mathbb{V}_1, \mathbb{V}_2, ..., \mathbb{V}_n\}$.
Obviously $\mathbb{S} = \mathbb{V}^{-1} \cdot \mathbb{W} = \left(\mathbb{W}^{-1} \cdot \mathbb{V}\right)^{-1} = \mathbb{P}^{-1}$.

Example 76 : Let us consider the vector $\mathbb{X}_e = (3, 5)$.
Chosen the new basis $\{\mathbb{W}_1, \mathbb{W}_2\} = \{(1, 1), (1, -1)\}$, it is $\mathbb{W} = \begin{Vmatrix} 1 & 1 \\ 1 & -1 \end{Vmatrix}$ from which we obtain: $\mathbb{W}^{-1} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{Vmatrix}$. It can be easily seen that $\mathbb{X}_{\mathrm{w}} = (4, -1)$.
Chosing another basis $\{\mathbb{V}_1, \mathbb{V}_2\} = \{(2, 1), (1, 2)\}$, it is $\mathbb{V} = \begin{Vmatrix} 2 & 1 \\ 1 & 2 \end{Vmatrix}$ and $\mathbb{X}_{\mathrm{v}} = \left(\frac{1}{3}, \frac{7}{3}\right)$.

Performing the products we verify that: $\begin{Vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{Vmatrix} \cdot \begin{Vmatrix} 2 & 1 \\ 1 & 2 \end{Vmatrix} \cdot \begin{Vmatrix} \frac{1}{3} \\ \frac{7}{3} \end{Vmatrix} = \begin{Vmatrix} 4 \\ -1 \end{Vmatrix}$, or that: $\mathbb{W}^{-1} \cdot \mathbb{V} \cdot \mathbb{X}_{\mathrm{v}} = \mathbb{X}_{\mathrm{w}}$.

CHANGE OF BASIS AND LINEAR MAPS

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \mathbb{A}_n \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{n,1}$ be a linear map, and let $\mathbb{X} = \mathbb{X}_e$ and $\mathbb{Y} = \mathbb{Y}_e$ be expressed under the standard basis. We change the basis of $\mathbb{R}^n$ using the vectors $\{\mathbb{W}_1, \mathbb{W}_2, ..., \mathbb{W}_n\}$, so $\mathbb{X}$ and $\mathbb{Y}$ will have, under this new basis, the coordinates $\mathbb{X}_{\mathrm{w}}$ and $\mathbb{Y}_{\mathrm{w}}$.
We want to see if there exists a matrix $\mathbb{B}$ such that $\mathbb{B} \cdot \mathbb{X}_{\mathrm{w}} = \mathbb{Y}_{\mathrm{w}}$, that is, such that it realizes a linear map in which the image of $\mathbb{X}_{\mathrm{w}}$ is $\mathbb{Y}_{\mathrm{w}}$, as well as $\mathbb{Y}_e$ was the image of $\mathbb{X}_e$.
With $\mathbb{A}$ and $\mathbb{B}$ we want to represent the same linear map under two different bases.
As previously seen, it is $\mathbb{X}_{\mathrm{w}} = \mathbb{W}^{-1} \cdot \mathbb{X}_e$ and $\mathbb{Y}_{\mathrm{w}} = \mathbb{W}^{-1} \cdot \mathbb{Y}_e$, from which, substituting in $\mathbb{B} \cdot \mathbb{X}_{\mathrm{w}} = \mathbb{Y}_{\mathrm{w}}$, we obtain: $\mathbb{B} \cdot \mathbb{W}^{-1} \cdot \mathbb{X}_e = \mathbb{W}^{-1} \cdot \mathbb{Y}_e$ or:
$\mathbb{W} \cdot \mathbb{B} \cdot \mathbb{W}^{-1} \cdot \mathbb{X}_e = \mathbb{Y}_e = \mathbb{A} \cdot \mathbb{X}_e$ from which we obtain $\mathbb{W} \cdot \mathbb{B} \cdot \mathbb{W}^{-1} = \mathbb{A}$ and so:
$\mathbb{B} = \mathbb{W}^{-1} \cdot \mathbb{A} \cdot \mathbb{W}$, which is the matrix of the same linear map but under the basis $\{\mathbb{W}_1, \mathbb{W}_2, ..., \mathbb{W}_n\}$.

Example 77 : Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \mathbb{A}_2 \cdot \mathbb{X}_{2,1} = \mathbb{Y}_{2,1}$, with $\mathbb{A}_2 = \begin{Vmatrix} 1 & 2 \\ 1 & 3 \end{Vmatrix}$. Let $\mathbb{X}_{2,1} = \begin{Vmatrix} 1 \\ 2 \end{Vmatrix}$ and so $\mathbb{Y}_{2,1} = \begin{Vmatrix} 1 & 2 \\ 1 & 3 \end{Vmatrix} \cdot \begin{Vmatrix} 1 \\ 2 \end{Vmatrix} = \begin{Vmatrix} 5 \\ 7 \end{Vmatrix}$.
We use a new basis of $\mathbb{R}^2$ given by the vectors $\{\mathbb{W}_1, \mathbb{W}_2\} = \{(1,1), (1, -1)\}$.
It is $\mathbb{W} = \begin{Vmatrix} 1 & 1 \\ 1 & -1 \end{Vmatrix}$ from which $\mathbb{W}^{-1} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{Vmatrix}$ and so:
$\mathbb{X}_{\mathrm{w}} = \mathbb{W}^{-1} \cdot \mathbb{X}_e = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{Vmatrix} \cdot \begin{Vmatrix} 1 \\ 2 \end{Vmatrix} = \begin{Vmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{Vmatrix}$ and
$\mathbb{Y}_{\mathrm{w}} = \mathbb{W}^{-1} \cdot \mathbb{Y}_e = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{Vmatrix} \cdot \begin{Vmatrix} 5 \\ 7 \end{Vmatrix} = \begin{Vmatrix} 6 \\ -1 \end{Vmatrix}$.
But $\mathbb{B} = \mathbb{W}^{-1} \cdot \mathbb{A} \cdot \mathbb{W} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{Vmatrix} \cdot \begin{Vmatrix} 1 & 2 \\ 1 & 3 \end{Vmatrix} \cdot \begin{Vmatrix} 1 & 1 \\ 1 & -1 \end{Vmatrix} = \begin{Vmatrix} \frac{7}{2} & -\frac{3}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{Vmatrix}$, and so we obtain: $\mathbb{B} \cdot \mathbb{X}_{\mathrm{w}} = \begin{Vmatrix} \frac{7}{2} & -\frac{3}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{Vmatrix} \cdot \begin{Vmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{Vmatrix} = \begin{Vmatrix} 6 \\ -1 \end{Vmatrix} = \mathbb{Y}_{\mathrm{w}}$, and also:
$\mathbb{W} \cdot \mathbb{B} \cdot \mathbb{W}^{-1} \cdot \mathbb{X}_e = \begin{Vmatrix} 1 & 1 \\ 1 & -1 \end{Vmatrix} \cdot \begin{Vmatrix} \frac{7}{2} & -\frac{3}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{Vmatrix} \cdot \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{Vmatrix} \cdot \begin{Vmatrix} 1 \\ 2 \end{Vmatrix} = \begin{Vmatrix} 5 \\ 7 \end{Vmatrix} = \mathbb{Y}_e$.

MATRIX SIMILARITY

Using what we have seen in the previous section, we give the following:
Definition 42 : Two square matrices $\mathbb{A}_n$ and $\mathbb{B}_n$ are called similar if it exists a square non-singular matrix $\mathbb{P}$ such that $\mathbb{B} = \mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P}$, or, equivalently, such that $\mathbb{A} \cdot \mathbb{P} = \mathbb{P} \cdot \mathbb{B}$.
The matrix $\mathbb{P}$ is sometimes called a similarity transformation.
Similar matrices represent the same linear map under two different bases, and $\mathbb{P}$ is the change of basis matrix.
We will denote two similar matrices $\mathbb{A}$ and $\mathbb{B}$ with the symbol $\mathbb{A} \sim \mathbb{B}$.

Similarity between matrices is an equivalence relation, i.e. it is reflexive, symmetric and transitive.
- Reflexivity: $\forall \mathbb{A}$: $\mathbb{A} \sim \mathbb{A}$.
In fact, just take $\mathbb{P} = \mathbb{I}_n$.
- Symmetry: $\mathbb{A} \sim \mathbb{B} \Rightarrow \mathbb{B} \sim \mathbb{A}$.
In fact $\mathbb{B} = \mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P} \;\Rightarrow \mathbb{A} = \mathbb{P} \cdot \mathbb{B} \cdot \mathbb{P}^{-1}$; if $\mathbb{P}^{-1} = \mathbb{Q}$ and then $\mathbb{P} = \mathbb{Q}^{-1}$, we obtain:
$\mathbb{A} = \mathbb{Q}^{-1} \cdot \mathbb{B} \cdot \mathbb{Q}$, i.e. $\mathbb{B} \sim \mathbb{A}$.
- Transitivity: $((\mathbb{A} \sim \mathbb{B})$ and $(\mathbb{B} \sim \mathbb{C})) \Rightarrow \mathbb{A} \sim \mathbb{C}$.
If $\mathbb{B} = \mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P}$ and if $\mathbb{C} = \mathbb{Q}^{-1} \cdot \mathbb{B} \cdot \mathbb{Q}$, substituting, we obtain:
$\mathbb{C} = \mathbb{Q}^{-1} \cdot \mathbb{B} \cdot \mathbb{Q} = \mathbb{Q}^{-1} \cdot \left(\mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P}\right) \cdot \mathbb{Q} = (\mathbb{P} \cdot \mathbb{Q})^{-1} \cdot \mathbb{A} \cdot (\mathbb{P} \cdot \mathbb{Q})$.
If $\mathbb{S} = \mathbb{P} \cdot \mathbb{Q}$ then we obtain: $\mathbb{C} = \mathbb{S}^{-1} \cdot \mathbb{A} \cdot \mathbb{S}$, i.e. $\mathbb{A} \sim \mathbb{C}$.

Another important property is the following:
Theorem 22 : Two similar matrices have the same Determinant.
Proof: If $\mathbb{A} \sim \mathbb{B}$ then $\mathbb{B} = \mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P}$; applying Binet's theorem we obtain:
$|\mathbb{B}| = \left|\mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P}\right| = \left|\mathbb{P}^{-1}\right| \cdot |\mathbb{A}| \cdot |\mathbb{P}|$ and since $\left|\mathbb{P}^{-1}\right| = \dfrac{1}{|\mathbb{P}|}$ the thesis is confirmed. •
And finally the following
Theorem 23 : If $\mathbb{A} \sim \mathbb{B}$ then $\mathbb{A}^k \sim \mathbb{B}^k, k \in \mathbb{N}$.
Proof: From $\mathbb{B} = \mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P}$ we obtain:
$\mathbb{B}^k = \left(\mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P}\right) \cdot \left(\mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P}\right) \cdot ... \cdot \left(\mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P}\right) =$; for the associative property
$\mathbb{B}^k = \mathbb{P}^{-1} \cdot \mathbb{A} \cdot \left(\mathbb{P} \cdot \mathbb{P}^{-1}\right) \cdot \mathbb{A} \cdot ... \cdot \mathbb{A} \cdot \left(\mathbb{P} \cdot \mathbb{P}^{-1}\right) \cdot \mathbb{A} \cdot \mathbb{P} = \mathbb{P}^{-1} \cdot \mathbb{A}^k \cdot \mathbb{P}$ so the thesis is confirmed.

EIGENVALUES AND EIGENVECTORS
Given a linear map $f : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbb{A}_n \cdot \mathbb{X}_{n,1} = \mathbb{Y}_{n,1}$, we want to check if there are vectors $\mathbb{X}$ whose image in such linear map is a scalar multiple of the vector $\mathbb{X}$ itself, i.e. such that $\mathbb{A} \cdot \mathbb{X} = \lambda\, \mathbb{X}$, where λ is a scalar (real or complex).
So an eigenvector of a square matrix is a non-zero vector that, when multiplied by the matrix, yields a vector that is parallel to the original.
The system $\mathbb{A} \cdot \mathbb{X} = \lambda\, \mathbb{X}$ is equivalent to the system $(\mathbb{A} - \lambda\, \mathbb{I}) \cdot \mathbb{X} = \mathbb{O}$, i.e. a linear homogeneous system in the unknowns $\mathbb{X}$ having, between its coefficients, the parameter λ.
The null vector $\mathbb{X} = \mathbb{O}$ is always a solution of this system, but we are interested of course to the presence of other solutions in addition to the null one.
For this purpose, since $\mathbb{A} - \lambda\, \mathbb{I}$ is a square matrix, its Determinant must be equal to 0: $|\mathbb{A} - \lambda\, \mathbb{I}| = 0$, otherwise, by Cramer's rule, this homogeneous system will have only one solution, i.e. the null one.
The equation $\mathcal{P}_n(\lambda) = |\mathbb{A} - \lambda\, \mathbb{I}| = 0$ is a polynomial equation of degree n in the unknown λ; for the fundamental theorem of algebra it admits exactly n roots, which may be real or complex, simple (of multiplicity 1) or multiple.
If the matrix $\mathbb{A}$ has real entries, complex roots will always be in an even number, each complex root being present with its conjugate.
The roots of the equation $\mathcal{P}_n(\lambda) = |\mathbb{A} - \lambda\, \mathbb{I}| = 0$ are the eigenvalues of the matrix $\mathbb{A}$; a vector $\mathbb{X} \neq \mathbb{O}$ such that $\mathbb{A} \cdot \mathbb{X} = \lambda\, \mathbb{X}$ is called an eigenvector with the eigenvalue λ.
The polynomial $\mathcal{P}_n(\lambda)$ is the characteristic polynomial; the spectrum of a matrix is the set of its eigenvalues: $\{\lambda_1, \lambda_2, ..., \lambda_m\}$, with $m \leq n$ due to the possible multiple roots.

The spectral radius of a square matrix is the supremum among the absolute values of the elements in its spectrum: $\rho = \text{Max}\ \{|\lambda_1|, |\lambda_2|, ..., |\lambda_m|\}$, where $|\lambda_i|$ represents the absolute value if λ_i is a real number, and represents the modulus if λ_i is a complex number.

The algebraic multiplicity of an eigenvalue λ is defined as the multiplicity of the corresponding root λ of the characteristic polynomial and will be denoted by m^a_λ.
Obviously, if $\lambda_1, \lambda_2, ..., \lambda_m,\ m \leq n$, are the eigenvalues (simple or multiple) of the matrix $\mathbb{A}$, it is: $\sum_{i=1}^{m} m^a_{\lambda_i} = n$.

Example 78 : Given the matrix $\mathbb{A} = \begin{Vmatrix} 1 & 2 \\ 3 & 2 \end{Vmatrix}$, let us determine its eigenvalues and the corresponding eigenvectors.
From $|\mathbb{A} - \lambda\,\mathbb{I}| = \begin{vmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{vmatrix} = 0$ we obtain:
$\mathcal{P}_2(\lambda) = (1-\lambda)(2-\lambda) - 6 = \lambda^2 - 3\lambda - 4 = 0$, so the roots are $\lambda_1 = -1$ and $\lambda_2 = 4$.
The spectral radius of the matrix is equal to 4.
To find the eigenvectors with $\lambda_1 = -1$ we solve the system: $(\mathbb{A} - (-1) \cdot \mathbb{I}) \cdot \mathbb{X} = \mathbb{O}$, and so: $\begin{Vmatrix} 2 & 2 \\ 3 & 3 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix} = \mathbb{O}$, i.e. the equation $2x_1 + 2x_2 = 0$, satisfied when $x_2 = -x_1$.
All the vectors $\mathbb{X}_1 = (k, -k)$ are thus such that: $\begin{Vmatrix} 1 & 2 \\ 3 & 2 \end{Vmatrix} \cdot \begin{Vmatrix} k \\ -k \end{Vmatrix} = (-1) \cdot \begin{Vmatrix} k \\ -k \end{Vmatrix}$.
To find the eigenvectors with $\lambda_1 = 4$ we solve the system $(\mathbb{A} - 4 \cdot \mathbb{I}) \cdot \mathbb{X} = \mathbb{O}$, and so: $\begin{Vmatrix} -3 & 2 \\ 3 & -2 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix} = \mathbb{O}$, i.e. the equation $3x_1 - 2x_2 = 0$, satisfied when $x_2 = \frac{3}{2}x_1$.
All the vectors $\mathbb{X}_2 = \left(k, \frac{3}{2}k\right)$ are thus such that: $\begin{Vmatrix} 1 & 2 \\ 3 & 2 \end{Vmatrix} \cdot \begin{Vmatrix} k \\ \frac{3}{2}k \end{Vmatrix} = 4 \cdot \begin{Vmatrix} k \\ \frac{3}{2}k \end{Vmatrix}$.

Example 79 : Given the matrix $\mathbb{A} = \begin{Vmatrix} -1 & 0 & -5 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{Vmatrix}$ let us determine its eigenvalues and the corresponding eigenvectors. From $|\mathbb{A} - \lambda\,\mathbb{I}| = \begin{vmatrix} -1-\lambda & 0 & -5 \\ 0 & 1-\lambda & 2 \\ 1 & 0 & 3-\lambda \end{vmatrix} = 0$ we obtain:
$\mathcal{P}_3(\lambda) = (1-\lambda)\,[(-1-\lambda)\,(3-\lambda) + 5] = (1-\lambda)\left(\lambda^2 - 2\lambda + 2\right) = 0$ whose roots are: $\lambda_1 = 1,\ \lambda_2 = 1 + i,\ \lambda_3 = 1 - i$. We have found three simple solutions, one is real and two are complex and conjugate.
Since $|1+i| = |1-i| = \sqrt{2}$, this is the spectral radius of the matrix.
Let us find the eigenvectors with $\lambda_1 = 1$.
We have to solve the system $(\mathbb{A} - 1 \cdot \mathbb{I}) \cdot \mathbb{X} = \mathbb{O}$, i.e. $\begin{vmatrix} -2 & 0 & -5 \\ 0 & 0 & 2 \\ 1 & 0 & 2 \end{vmatrix} \cdot \begin{Vmatrix} x \\ y \\ z \end{Vmatrix} = \mathbb{O}$, from which we obtain: $\begin{cases} -2x - 5z = 0 \\ 2z = 0 \\ x + 2z = 0 \end{cases}$ which is satisfied if $\begin{cases} x = 0 \\ \forall\, y \\ z = 0 \end{cases}$.

All the vectors $\mathbb{X}_1 = (0, k, 0)$ are thus such that: $\left\|\begin{matrix} -1 & 0 & -5 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{matrix}\right\| \cdot \left\|\begin{matrix} 0 \\ k \\ 0 \end{matrix}\right\| = 1 \cdot \left\|\begin{matrix} 0 \\ k \\ 0 \end{matrix}\right\|$.

Let us find now the eigenvectors with $\lambda_2 = 1 + i$.

We have to solve the system $\left|\begin{matrix} -2-i & 0 & -5 \\ 0 & -i & 2 \\ 1 & 0 & 2-i \end{matrix}\right| \cdot \left\|\begin{matrix} x \\ y \\ z \end{matrix}\right\| = \mathbb{O}$, from which we obtain:

$\begin{cases} (-2-i)\,x - 5z = 0 \\ -i\,y + 2z = 0 \\ x + (2-i)z = 0 \end{cases}$ which is satisfied if $\begin{cases} x = (i-2)\,z \\ y = -2i\,z \end{cases}$.

Therefore all the vectors $\mathbb{X}_2 = k\,(i-2, -2i, 1)$ satisfy the equation:

$\left\|\begin{matrix} -1 & 0 & -5 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{matrix}\right\| \cdot \left\|\begin{matrix} i-2 \\ -2i \\ 1 \end{matrix}\right\| = (1+i)\left\|\begin{matrix} i-2 \\ -2i \\ 1 \end{matrix}\right\|$.

Let us finally find the eigenvectors with $\lambda_3 = 1 - i$.

We have to solve the system $\left|\begin{matrix} -2+i & 0 & -5 \\ 0 & i & 2 \\ 1 & 0 & 2+i \end{matrix}\right| \cdot \left\|\begin{matrix} x \\ y \\ z \end{matrix}\right\| = \mathbb{O}$, from which we obtain:

$\begin{cases} (-2+i)\,x - 5z = 0 \\ i\,y + 2z = 0 \\ x + (2+i)\,z = 0 \end{cases}$ which is satisfied if $\begin{cases} x = -(i+2)\,z \\ y = 2i\,z \end{cases}$.

Therefore all the vectors $\mathbb{X}_3 = k\,(-(i+2), 2i, 1)$ satisfy the equation:

$\left\|\begin{matrix} -1 & 0 & -5 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{matrix}\right\| \cdot \left\|\begin{matrix} -(i+2) \\ 2i \\ 1 \end{matrix}\right\| = (1-i)\left\|\begin{matrix} -(i+2) \\ 2i \\ 1 \end{matrix}\right\|$.

THE CHARACTERISTIC POLYNOMIAL

Definition 43 : The trace of a square matrix $\mathbb{A}_n$, tr($\mathbb{A}$), is defined to be the sum of the entries on the main diagonal: $\mathrm{tr}(\mathbb{A}) = \sum_{i=1}^{n} a_{ii}$.

Definition 44 : Given a square matrix $\mathbb{A}_n$, a Principal Minor MP is the Determinant of a square submatrix having as entries of its main diagonal only entries of the main diagonal of $\mathbb{A}_n$, i.e. the Determinant of a square submatrix built from $\mathbb{A}_n$ by choosing rows and columns with the same indexes.

Example 80 : The matrix $\mathbb{A}_3 = \left\|\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix}\right\|$ has only one Principal Minor of the third order, the Determinant of the matrix itself: $MP_{123} = \left|\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix}\right|$; the matrix has three Principal Minors of the second order, which are: $MP_{12} = \left|\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix}\right|$,

$MP_{13} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$, $MP_{23} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$; the matrix has three Principal Minors of the first order, which are $MP_1 = |a_{11}|$, $MP_2 = |a_{22}|$, $MP_3 = |a_{33}|$.
The indexes refer to the rows (and columns) used to construct the Principal Minor.

Let us now see the relationship between the coefficients of the characteristic polynomial and the Principal Minors of the matrix.

Given the matrix $\mathbb{A}_2 = \left\| \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix} \right\|$, it is $\mathbb{A} - \lambda\,\mathbb{I} = \left\| \begin{matrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{matrix} \right\|$, and so:
$|\mathbb{A} - \lambda\,\mathbb{I}| = \lambda^2 - (a_{11} + a_{22})\,\lambda + (a_{11}\,a_{22} \; - \; a_{12}\,a_{21}) =$
$|\mathbb{A} - \lambda\,\mathbb{I}| = \lambda^2 - (MP_1 + MP_2)\,\lambda + MP_{12} = \lambda^2 - \text{tr}(\mathbb{A})\,\lambda + \det(\mathbb{A})\,.$

Let us consider now the third order matrix $\mathbb{A}_3 = \left\| \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} \right\|$.
It is $\mathbb{A} - \lambda\,\mathbb{I} = \left\| \begin{matrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{matrix} \right\|$ and so we obtain:
$|\mathbb{A} - \lambda\,\mathbb{I}| = \; - \lambda^3 + (a_{11} + a_{22} + a_{33})\,\lambda^2 +$
$- \left[(a_{11}\,a_{22} - a_{12}\,a_{21}) + (a_{11}\,a_{33} - a_{13}\,a_{31}) + (a_{22}\,a_{33} - a_{23}\,a_{32})\right]\lambda +$
$+ (a_{11}\,a_{22}\,a_{33} - a_{11}\,a_{23}\,a_{32} + a_{12}\,a_{23}\,a_{31} - a_{12}\,a_{21}\,a_{33} + a_{13}\,a_{21}\,a_{32} - a_{13}\,a_{22}\,a_{31}) =$
$= \; - \lambda^3 + (MP_1 + MP_2 + MP_3)\,\lambda^2 - (MP_{12} + MP_{13} + MP_{23})\,\lambda + MP_{123} =$
$= \; - \lambda^3 + \text{tr}(\mathbb{A})\,\lambda^2 - (MP_{12} + MP_{13} + MP_{23})\,\lambda + \det(\mathbb{A})\,.$

As seen in the two previous examples, let us construct the general expression of the characteristic polynomial of a square matrix of any order n.
It is:
$$\mathcal{P}_n(\lambda) = |\mathbb{A} - \lambda\,\mathbb{I}| = \sum_{i=0}^{n} (-1)^{n-i} \cdot \left(\sum MP^i\right) \lambda^{n-i}\,,$$
where $\sum MP^i$ is the sum of all the Principal Minors of order i of the matrix, having put $MP^0 = 1$. We see that $\sum MP^i$ is a sum of $\binom{n}{i}$ terms, that $\sum MP^1 = \text{tr}(\mathbb{A})$ and that the constant term is equal to $MP^n = \det(\mathbb{A})$.

Factoring the characteristic polynomial into its roots, we obtain:
$$\mathcal{P}_n(\lambda) = (-1)^n \cdot \prod_{i=1}^{n} (\lambda - \lambda_i)\,.$$
From the algebraic theory we know that: $\mathcal{P}_n(\lambda) = \sum_{i=0}^{n} (-1)^{n-i} \cdot \mathrm{S}_i \cdot \lambda^{n-i}$, where S_i is the sum of all the possible products of i roots.
The coefficient of λ^{n-1} is $(-1)^{n-1} \cdot \sum_{i=1}^{n} \lambda_i$ while the constant term is $\prod_{i=1}^{n} \lambda_i$, and from this we obtain the two equalities: $\sum_{i=1}^{n} \lambda_i = \text{tr}(\mathbb{A})$ and $\prod_{i=1}^{n} \lambda_i = \det(\mathbb{A})$.
From this we see immediately that $\lambda = 0$ is an eigenvalue if and only if $\det(\mathbb{A}) = 0$.

The Determinant $\det(\mathbb{A})$ is the only Principal Minor of order n and it is equal to the product of the n eigenvalues: $\prod_{i=1}^{n} \lambda_i = \det(\mathbb{A})$.

The trace $\mathrm{tr}(\mathbb{A})$ is the sum of all the Principal Minors of order 1 and it corresponds to the sum of the products of the eigenvalues 1 to 1, i.e. to the sum of the eigenvalues: $\sum_{i=1}^{n} \lambda_i = \mathrm{tr}(\mathbb{A})$.

Let us formulate what we have stated into the following:
Theorem 24 : The sum of all the Principal Minors of order $n-i$ is equal to the sum of all the products of $n-i$ eigenvalues.

Example 81 : Given the matrix $\mathbb{A} = \left\| \begin{matrix} 1 & 0 & -1 \\ 0 & 1 & m \\ 1 & 0 & k \end{matrix} \right\|$, let us determine, varying the parameters m and k, its eigenvalues and their algebraic multiplicity. It is:

$$|\mathbb{A} - \lambda\,\mathbb{I}| = \begin{vmatrix} 1-\lambda & 0 & -1 \\ 0 & 1-\lambda & m \\ 1 & 0 & k-\lambda \end{vmatrix} = (1-\lambda)\left[(1-\lambda)(k-\lambda)+1\right] =$$
$$= (1-\lambda)\left(\lambda^2 - \lambda\,(k+1) + k + 1\right) = 0\,.$$

And so the matrix has the eigenvalue $\lambda = 1\ \forall\, m$ and $\forall\, k$.
In order for the eigenvalue $\lambda = 1$ to have its algebraic multiplicity equal to 2, it must be $\lambda^2 - (k+1)\,\lambda + k + 1 = 0$ if $\lambda = 1$, and so: $1 - (k+1) + k + 1 = 0$, which is not satisfied for any value of k.
So we study the roots of $\lambda^2 - (k+1)\,\lambda + k + 1 = 0$.
We obtain $\lambda = \dfrac{(k+1) \pm \sqrt{(k+1)^2 - 4(k+1)}}{2} = \dfrac{(k+1) \pm \sqrt{k^2 - 2k - 3}}{2}$.
If $k^2 - 2k - 3 > 0$ we have two real and distinct eigenvalues, both different from 1 as seen before;
if $k^2 - 2k - 3 < 0$ we have two complex and conjugate eigenvalues;
if $k^2 - 2k - 3 = 0$, i.e. if $k = 3$ or if $k = -1$ we have a real eigenvalue whose algebraic multiplicity is equal to 2; we see with easy calculations that it is $\lambda = 2$ if $k = 3$, while it is $\lambda = 0$ if $k = -1$.

EIGENVALUES OF SPECIAL MATRICES
Given a matrix $\mathbb{A}_n$, $\lambda_1, \lambda_2, ..., \lambda_n$ are its eigenvalues, real or complex, simple or multiple.

The following properties are valid:
A1) The transpose matrix $\mathbb{A}^{\mathrm{T}}$ has the same eigenvalues of $\mathbb{A}$: $\lambda_1, \lambda_2, ..., \lambda_n$.
A2) If $|\mathbb{A}| \neq 0$, the eigenvalues of the inverse matrix $\mathbb{A}^{-1}$ are the reciprocal $\dfrac{1}{\lambda_1}, \dfrac{1}{\lambda_2}, ..., \dfrac{1}{\lambda_n}$.
In fact, from $\mathbb{A} \cdot \mathbb{X} = \lambda\,\mathbb{X}$ it follows: $\mathbb{X} = \mathbb{A}^{-1} \cdot \lambda\,\mathbb{X}$, or $\mathbb{A}^{-1} \cdot \mathbb{X} = \dfrac{1}{\lambda}\,\mathbb{X}$.
A3) The eigenvalues of the matrix $k\,\mathbb{A}$ are the multiples $k\,\lambda_1, k\,\lambda_2, ..., k\,\lambda_n$.
In fact, if $\mathbb{A} \cdot \mathbb{X} = \lambda\,\mathbb{X}$, immediately it follows that: $k\,\mathbb{A} \cdot \mathbb{X} = (k\lambda)\,\mathbb{X}$.
A4) The eigenvalues of a diagonal or triangular matrix are the entries of the main diagonal.

A5) Similar matrices have the same spectrum, the same determinant and the same trace.
In fact, if $\mathbb{A} = \mathbb{P}^{-1} \cdot \mathbb{B} \cdot \mathbb{P}$, from Binet's theorem, we have:
$|\mathbb{A} - \lambda\,\mathbb{I}| = \left|\mathbb{P}^{-1} \cdot \mathbb{B} \cdot \mathbb{P} - \lambda\,\mathbb{I}\right| = |\mathbb{P}| \cdot \left|\mathbb{P}^{-1} \cdot \mathbb{B} \cdot \mathbb{P} - \lambda\,\mathbb{I}\right| \cdot \left|\mathbb{P}^{-1}\right| =$
$= \left|\mathbb{P} \cdot \left(\mathbb{P}^{-1} \cdot \mathbb{B} \cdot \mathbb{P} - \lambda\,\mathbb{I}\right) \cdot \mathbb{P}^{-1}\right| = \left|\mathbb{P} \cdot \mathbb{P}^{-1} \cdot \mathbb{B} \cdot \mathbb{P} \cdot \mathbb{P}^{-1} - \mathbb{P} \cdot \lambda \cdot \mathbb{I} \cdot \mathbb{P}^{-1}\right| = |\mathbb{B} - \lambda\,\mathbb{I}|$.
Therefore two similar matrices have the same characteristic polynomial, and then they have the same eigenvalues, and since the trace and the Determinant are the second and the last coefficient of the characteristic polynomial, they are equal, i.e. $\text{tr}(\mathbb{A}) = \text{tr}(\mathbb{B})$ and $\det(\mathbb{A}) = \det(\mathbb{B})$.
A6) The orthogonal matrices have all their eigenvalues, if real, equal to ± 1.
In fact, since for orthogonal matrices it is $\mathbb{A}^{\text{T}} = \mathbb{A}^{-1}$, for the property A1), the transpose $\mathbb{A}^{\text{T}}$ and the inverse $\mathbb{A}^{-1}$ of an orthogonal matrix have the same eigenvalues of the matrix $\mathbb{A}$. If λ is an eigenvalue of the matrix $\mathbb{A}$, for the property A2) it should also be $\lambda = \dfrac{1}{\lambda}$ or $\lambda^2 = 1$ and so $\lambda = \pm 1$.
A7) Given two square matrices of the same order $\mathbb{A}_n$ and $\mathbb{B}_n$, the matrices $\mathbb{A} \cdot \mathbb{B}$ and $\mathbb{B} \cdot \mathbb{A}$ have the same eigenvalues.

If $\mathbb{A}$ is a rectangular matrix $(m \cdot n)$ and $\mathbb{B}$ is a rectangular matrix $(n \cdot m)$, with $m > n$, the square matrix $\mathbb{C}_m = \mathbb{A} \cdot \mathbb{B}$ has the same eigenvalues of the matrix $\mathbb{D}_n = \mathbb{B} \cdot \mathbb{A}$ adding $m - n$ eigenvalues equal to 0.

Finally, if $\mathbb{A}$ and $\mathbb{B}$ are square matrices, even of a different order, having eigenvalues respectively $\lambda_1, \lambda_2, ..., \lambda_n$ and $\mu_1, \mu_2, ..., \mu_m$, the matrix $\mathbb{A} \otimes \mathbb{B}$, resulting from the Kronecker product, has eigenvalues $\nu_{i,j} = \lambda_i\, \mu_j$, $1 \leq i \leq n$ and $1 \leq j \leq m$.

PROPERTIES OF THE EIGENVECTORS
Regarding the eigenvectors, the following properties are valid:
B1) If $\mathbb{X}$ is an eigenvector with the eigenvalue λ, also $k\,\mathbb{X}, \forall\, k \in \mathbb{R}^*$ is an eigenvector of $\mathbb{A}$ with λ.
In fact: $\mathbb{A} \cdot (k\,\mathbb{X}) = k\,\mathbb{A} \cdot \mathbb{X} = k\,\lambda\,\mathbb{X} = \lambda\,(k\,\mathbb{X})$;
B2) The same eigenvector cannot correspond to two different eigenvalues.
In fact, if $\mathbb{A} \cdot \mathbb{X} = \lambda_1\,\mathbb{X}$ and also $\mathbb{A} \cdot \mathbb{X} = \lambda_2\,\mathbb{X}$, with $\lambda_1 \neq \lambda_2$, subtracting we obtain:
$\mathbb{A} \cdot \mathbb{X} - \mathbb{A} \cdot \mathbb{X} = \mathbb{O} = (\lambda_1 - \lambda_2)\,\mathbb{X}$, and since $\mathbb{X} \neq \mathbb{O}$ this is possible only if $\lambda_1 = \lambda_2$, against the hypothesis.
B3) Eigenvectors with distinct eigenvalues are linearly independent vectors.
Let $\lambda_1, \lambda_2, ..., \lambda_m$, $m \leq n$, be the distinct eigenvalues of the matrix $\mathbb{A}_n$ and let $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_m$ be the corresponding eigenvectors. If these eigenvectors were linearly dependent, at least one of them, $\mathbb{X}_0$, can be expressed as a linear combination of some of the other eigenvectors.
So we suppose: $\mathbb{X}_0 = \alpha_1\,\mathbb{X}_1 + \alpha_2\,\mathbb{X}_2 + ... + \alpha_k\,\mathbb{X}_k$, with $k \leq m - 1$.
Premultiplying both terms of the equality by the matrix $\mathbb{A}$ we obtain:
$\mathbb{A} \cdot \mathbb{X}_0 = \mathbb{A} \cdot (\alpha_1\,\mathbb{X}_1 + \alpha_2\,\mathbb{X}_2 + ... + \alpha_k\,\mathbb{X}_k)$ from which:
$\mathbb{A} \cdot \mathbb{X}_0 = \mathbb{A} \cdot \alpha_1\,\mathbb{X}_1 + \mathbb{A} \cdot \alpha_2\,\mathbb{X}_2 + ... + \mathbb{A} \cdot \alpha_k\,\mathbb{X}_k$ or:
$\lambda_0\,\mathbb{X}_0 = \alpha_1\,\lambda_1\,\mathbb{X}_1 + \alpha_2\,\lambda_2\,\mathbb{X}_2 + \; + \alpha_k\,\lambda_k\,\mathbb{X}_k$,
since $\mathbb{X}_0, \mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_k$ are eigenvectors.
On the contrary, multiplying both terms of the same equality by λ_0 we obtain:
$\lambda_0\,\mathbb{X}_0 = \alpha_1\,\lambda_0\,\mathbb{X}_1 + \alpha_2\,\lambda_0\,\mathbb{X}_2 + \; + \alpha_k\,\lambda_0\,\mathbb{X}_k$.
Subtracting member to member we finally obtain:
$\mathbb{O} = \alpha_1\,(\lambda_1 - \lambda_0)\,\mathbb{X}_1 + \alpha_2\,(\lambda_2 - \lambda_0)\,\mathbb{X}_2 + \; + \alpha_k\,(\lambda_k - \lambda_0)\,\mathbb{X}_k$,
from which, since $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_k$ are linearly independent vectors, it follows that:

$\lambda_1 = \lambda_0,\ \lambda_2 = \lambda_0, ..., \lambda_k = \lambda_0$ against the hypothesis that the eigenvalues are distinct.
B4) If $\mathbb{A} = \mathbb{P} \cdot \mathbb{B} \cdot \mathbb{P}^{-1}$ and if $\mathbb{X}$ is an eigenvector with the eigenvalue λ for the matrix $\mathbb{A}$, then $\mathbb{P}^{-1} \cdot \mathbb{X}$ is an eigenvector with the eigenvalue λ for the similar matrix $\mathbb{B}$.
In fact, given two similar matrices $\mathbb{A}$ and $\mathbb{B}$, and if $\mathbb{A} \cdot \mathbb{X} = \lambda\, \mathbb{X}$, from $\mathbb{A} = \mathbb{P} \cdot \mathbb{B} \cdot \mathbb{P}^{-1}$ we obtain: $\lambda\, \mathbb{X} = \mathbb{A} \cdot \mathbb{X} = \mathbb{P} \cdot \mathbb{B} \cdot \mathbb{P}^{-1} \cdot \mathbb{X}$ from which, premultiplying by $\mathbb{P}^{-1}$, we obtain:
$\mathbb{P}^{-1} \cdot \lambda\, \mathbb{X} = \lambda\, \mathbb{P}^{-1} \cdot \mathbb{X} = \mathbb{P}^{-1} \cdot \mathbb{P} \cdot \mathbb{B} \cdot \mathbb{P}^{-1} \cdot \mathbb{X} = \mathbb{B} \cdot \mathbb{P}^{-1} \cdot \mathbb{X}$,
and so, if $\mathbb{P}^{-1} \cdot \mathbb{X} = \mathbb{Y}$, we obtain:
$\lambda \left(\mathbb{P}^{-1} \cdot \mathbb{X}\right) = \lambda\, \mathbb{Y} = \mathbb{B} \cdot \mathbb{Y} = \mathbb{B} \cdot \left(\mathbb{P}^{-1} \cdot \mathbb{X}\right)$
i.e. $\mathbb{P}^{-1} \cdot \mathbb{X}$ is an eigenvector for the matrix $\mathbb{B}$.
B5) If $\mathbb{X}_0$ is an eigenvector for the non-singular matrix $\mathbb{A}$ with the eigenvalue λ_0, then $\mathbb{X}_0$ is also an eigenvector for the inverse matrix $\mathbb{A}^{-1}$ with the eigenvalue $\dfrac{1}{\lambda_0}$.
In fact, from $\mathbb{A} \cdot \mathbb{X}_0 = \lambda_0 \cdot \mathbb{X}_0$ and since $\mathbb{A}$ is invertible, we obtain:
$\mathbb{X}_0 = \mathbb{A}^{-1} \cdot \lambda_0 \cdot \mathbb{X}_0 = \lambda_0 \cdot \mathbb{A}^{-1} \cdot \mathbb{X}_0$ and then: $\dfrac{1}{\lambda_0} \cdot \mathbb{X}_0 = \mathbb{A}^{-1} \cdot \mathbb{X}_0$, that is the thesis.

Example 82 : Given a matrix $\mathbb{A}$, from $\mathbb{A}^2 = \mathbb{A} \cdot \mathbb{A}$, if $\mathbb{A} \cdot \mathbb{X}_0 = \lambda_0\, \mathbb{X}_0$ it is also:
$\mathbb{A}^2 \cdot \mathbb{X}_0 = \mathbb{A} \cdot \mathbb{A} \cdot \mathbb{X}_0 = \mathbb{A} \cdot \lambda_0\, \mathbb{X}_0 = \lambda_0\, \mathbb{A} \cdot \mathbb{X}_0 = \lambda_0^2\, \mathbb{X}_0$, i.e. the eigenvalues of the matrix $\mathbb{A}^2$ are the squares of the eigenvalues of $\mathbb{A}$, while the corresponding eigenvectors are the same as $\mathbb{A}$.

Example 83 : Let us consider an idempotent matrix, i.e. such that $\mathbb{A}^2 = \mathbb{A}$.
From $\mathbb{A} \cdot \mathbb{X}_0 = \lambda_0\, \mathbb{X}_0$ and from $\mathbb{A}^2 \cdot \mathbb{X}_0 = \mathbb{A} \cdot \mathbb{X}_0$ we obtain: $\lambda_0^2\, \mathbb{X}_0 = \lambda_0\, \mathbb{X}_0$, which can be satisfied only if $\lambda_0 = 0$ or if $\lambda_0 = 1$.
Then only the values 0 and 1 can be the eigenvalues of an idempotent matrix.

THE ASSOCIATED EIGENSPACE OF AN EIGENVALUE
Suppose that $\mathbb{A}$ is a square matrix and λ is an eigenvalue of $\mathbb{A}$, simple or multiple.
Theorem 25 : The eigenvectors with the same eigenvalue λ, with the inclusion of the null vector $\mathbb{O}$, form a vector subspace, called the associated eigenspace of the eigenvalue λ: $\mathcal{ES}_\lambda$.
Proof : If $\mathbb{X}_1$ and $\mathbb{X}_2$ are eigenvectors with the same eigenvalue λ, it is:
$\mathbb{A} \cdot \mathbb{X}_1 = \lambda\, \mathbb{X}_1$ and $\mathbb{A} \cdot \mathbb{X}_2 = \lambda\, \mathbb{X}_2$; then:
$\mathbb{A} \cdot (\alpha\, \mathbb{X}_1 + \beta\, \mathbb{X}_2) = \alpha\, \mathbb{A} \cdot \mathbb{X}_1 + \beta\, \mathbb{A} \cdot \mathbb{X}_2 = \alpha\, \lambda\, \mathbb{X}_1 + \beta\, \lambda\, \mathbb{X}_2 = \lambda\, (\alpha\, \mathbb{X}_1 + \beta\, \mathbb{X}_2)$
i.e. also $\alpha\, \mathbb{X}_1 + \beta\, \mathbb{X}_2$ is an eigenvector with λ. ●

Therefore we must determine, for each eigenvalue λ_i, the dimension of its associated eigenspace. The geometric multiplicity of an eigenvalue λ_i is defined as the dimension of the associated eigenspace, i.e. the number of linearly independent eigenvectors with that eigenvalue; the geometric multiplicity will be denoted by m_i^g.
If λ_0 is an eigenvalue for the matrix $\mathbb{A}$, to find all the eigenvectors with λ_0 we have to solve the linear homogeneous system $(\mathbb{A} - \lambda_0\, \mathbb{I}) \cdot \mathbb{X} = \mathbb{O}$, i.e. we have to determine the Kernel of the linear map $(\mathbb{A} - \lambda_0\, \mathbb{I}) \cdot \mathbb{X}$.
So the following equality applies: $m_{\lambda_0}^g = \mathrm{Dim}(\mathrm{Ker}(\mathbb{A} - \lambda_0\, \mathbb{I}))$.
For the "rank-nullity theorem" or "Sylvester's theorem", we know that:
$\mathrm{Dim}(\mathrm{Ker}(\mathbb{A} - \lambda_0\, \mathbb{I})) = n - \mathrm{Rank}(\mathbb{A} - \lambda_0\, \mathbb{I})$
and then we obtain the final equality:
$m_{\lambda_0}^g = n - \mathrm{Rank}(\mathbb{A} - \lambda_0\, \mathbb{I})$.
Since $|\mathbb{A} - \lambda_0\, \mathbb{I}| = 0$, it follows that $\mathrm{Rank}(\mathbb{A} - \lambda_0\, \mathbb{I}) \leq n - 1$ and so $m_{\lambda_0}^g \geq 1$, i.e.:

Theorem 26 : The associated eigenspace of every eigenvalue λ is a vector subspace whose dimension is $m^g_\lambda \geq 1$.
I.e.: no eigenvalue can have an eigenspace reduced to a single point, that is the null vector $\mathbb{O}$.

From the property B3) we know that the eigenvectors with distinct eigenvalues are linearly independent vectors; from the fundamental theorem of Algebra we know that a polynomial equation of degree n admits at most n distinct roots.
Since the eigenvectors of a matrix $\mathbb{A}_n$ are vectors of $\mathbb{R}^n$, we have the following:
Theorem 27 : A matrix $\mathbb{A}_n$ has at most n linearly independent eigenvectors.
Theorem 28 : If a matrix $\mathbb{A}_n$ has n distinct eigenvalues, then it has n linearly independent eigenvectors.

Example 84 : For the null matrix, it is: $|\mathbb{O}_n - \lambda\,\mathbb{I}_n| = (-\lambda)^n = 0$ and so the multiple solution $\lambda = 0$ with algebric multiplicity $m^a_0 = n$.
Since $\mathbb{O} \cdot \mathbb{X} = 0 \cdot \mathbb{X} = \mathbb{O}\,,\ \forall\, \mathbb{X} \in \mathbb{R}$, the associated eigenspace of the eigenvalue 0 is the whole $\mathbb{R}^n$, and so $m^g_0 = n = m^a_0$.

Similarly for the unit matrix $\mathbb{I}_n$, it is $|\mathbb{I}_n - \lambda\,\mathbb{I}_n| = (1-\lambda)^n = 0$, and so the multiple solution $\lambda = 1$ with algebraic multiplicity $m^a_1 = n$.
Since $\mathbb{I}_n \cdot \mathbb{X} = 1 \cdot \mathbb{X} = \mathbb{X}\,,\ \forall\, \mathbb{X} \in \mathbb{R}$, the associated eigenspace of the eigenvalue 1 is the whole $\mathbb{R}^n$, and so $m^g_1 = n = m^a_1$.

Example 85 : Given the matrix $\mathbb{A} = \left\|\begin{matrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{matrix}\right\|$, let us determine the associated eigenspace of its eigenvalues.
It is $|\mathbb{A} - \lambda\,\mathbb{I}| = \left\|\begin{matrix} 1-\lambda & 0 & -1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{matrix}\right\| = -\lambda^3 + 4\lambda^2 - 5\lambda + 2 = 0$; this equation has solutions $\lambda_1 = 1\,,\ \lambda_2 = 1\,,\ \lambda_3 = 2$. So $\lambda = 2$ is a simple solution while $\lambda = 1$ is a double solution.
To determine the eigenspace associated to $\lambda_1 = \lambda_2 = 1$ we solve the homogeneous system:
$|\mathbb{A} - 1 \cdot \mathbb{I}| = \mathbb{O}$, whose matrix is $(\mathbb{A} - 1 \cdot \mathbb{I}) = \left\|\begin{matrix} 0 & 0 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{matrix}\right\|$.
Since $\mathrm{Rank}(\mathbb{A} - 1 \cdot \mathbb{I}) = 2$, the dimension of the eigenspace associated to $\lambda_1 = \lambda_2 = 1$ is equal to $m^g_1 = 3 - \mathrm{Rank}(\mathbb{A} - 1 \cdot \mathbb{I}) = 1$, and so $1 = m^g_1 < m^a_1 = 2$.
Discarding the second row, we have the system $\begin{cases} -z = 0 \\ x + y = 0 \end{cases}$ that gives us the eigenvectors $(k, -k, 0) = k\,(1, -1, 0), k \in \mathbb{R}$.
To determine the eigenspace associated to $\lambda_3 = 2$ we solve the homogeneous system:
$|\mathbb{A} - 2\,\mathbb{I}| = \mathbb{O}$, whose matrix is $(\mathbb{A} - 2 \cdot \mathbb{I}) = \left\|\begin{matrix} -1 & 0 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & -1 \end{matrix}\right\|$.
Since $\mathrm{Rank}(\mathbb{A} - 2\,\mathbb{I}) = 2$, the dimension of the eigenspace associated to $\lambda_3 = 2$ is equal to $m^g_2 = 3 - \mathrm{Rank}(\mathbb{A} - 2 \cdot \mathbb{I}) = 1$, and so $m^g_2 = m^a_2 = 1$.

Discarding the first row, we have the system $\begin{cases} x+z=0 \\ x+y-z=0 \end{cases}$ that gives us the eigenvectors $(k,\,-2k,\,-k)=k\,(1,\,-2,\,-1), k\in\mathbb{R}$.

Example 86 : Given the matrix $\mathbb{A}=\begin{Vmatrix} 3 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & 3 \end{Vmatrix}$, let us determine the associated eigenspace of its eigenvalues.

It is $|\mathbb{A}-\lambda\,\mathbb{I}|=\begin{Vmatrix} 3-\lambda & -1 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & -1 & 3-\lambda \end{Vmatrix}=(2-\lambda)\big(\lambda^2-6\lambda+8\big)=0$; this equation has solutions $\lambda_1=2$, $\lambda_2=2$, $\lambda_3=4$. So $\lambda=4$ is a simple solution while $\lambda=2$ is a double solution.

To determine the eigenspace associated to $\lambda_1=\lambda_2=2$ we solve the homogeneous system: $|\mathbb{A}-2\,\mathbb{I}|=\mathbb{O}$, whose matrix is $(\mathbb{A}-2\,\mathbb{I})=\begin{Vmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 1 & -1 & 1 \end{Vmatrix}$.

Since $\mathrm{Rank}(\mathbb{A}-2\,\mathbb{I})=1$, the dimension of the eigenspace associated to $\lambda=2$ is equal to $m_2^g=3-\mathrm{Rank}(\mathbb{A}-2\,\mathbb{I})=2$, and so $m_2^g=m_2^a=2$.

From the only remaining equation: $x-y+z=0$ we obtain the eigenvectors $(y-z,y,z)$.

Since $m_2^g=2$, we must determine from this two independent vectors, which may be, for example, $(1,1,0)$ and $(-1,0,1)$.

Every eigenvector with $\lambda=2$ can be written as: $k\,(1,1,0)+h\,(-1,0,1), k,h\in\mathbb{R}$.

To determine the eigenspace associated to $\lambda_3=4$ we solve the homogeneous system: $|\mathbb{A}-4\,\mathbb{I}|=\mathbb{O}$, whose matrix is $(\mathbb{A}-4\,\mathbb{I})=\begin{Vmatrix} -1 & -1 & 1 \\ 0 & -2 & 0 \\ 1 & -1 & -1 \end{Vmatrix}$.

Since $\mathrm{Rank}(\mathbb{A}-4\,\mathbb{I})=2$, the dimension of the eigenspace associated to $\lambda=4$ is equal to $m_4^g=3-\mathrm{Rank}(\mathbb{A}-4\,\mathbb{I})=1$, and so $m_4^g=m_4^a=1$.

Finally we solve the system $\begin{cases} -x-y+z=0 \\ -2y=0 \\ x-y-z=0 \end{cases}$ from which we obtain $\begin{cases} x=z \\ y=0 \end{cases}$ and so the eigenvectors $(k,0,k)=k\,(1,0,1), k\in\mathbb{R}$.

The relation between geometric multiplicity and algebraic multiplicity of any eigenvalue is the following:

$1\le m_i^g=n-\mathrm{Rank}(\mathbb{A}-\lambda_0\,\mathbb{I})\le m_i^a\le n$, and this results from the following:

Theorem 29 : For every eigenvalue λ_i it is: $m_i^g\le m_i^a$.

Proof : By hypothesis, $\mathbb{A}_n$ is a matrix with the eigenvalue λ_0, whose geometric multiplicity is $m_{\lambda_0}^g=k$, while we do not know its algebraic multiplicity $m_{\lambda_0}^a$.

The eigenspace associated to λ_0 has dimension equal to k and therefore there are k linearly independent eigenvectors $\mathbb{X}_1,\mathbb{X}_2,\dots,\mathbb{X}_k$.

Let us form a basis of $\mathbb{R}^n$ with the k eigenvectors $\mathbb{X}_1,\mathbb{X}_2,\dots,\mathbb{X}_k$ together with $n-k$ vectors $\mathbb{V}_{k+1},\mathbb{V}_{k+2},\dots,\mathbb{V}_n$ freely chosen.

Let $\mathbb{W}=\big[\mathbb{X}_1\big|\mathbb{X}_2\big|\dots\big|\mathbb{X}_k\big|\mathbb{V}_{k+1}\big|\mathbb{V}_{k+2}\big|\dots\big|\mathbb{V}_n\big]$ be the matrix having as columns the vectors of the basis we have constructed.

Let's suppose $\mathbb{B}_n=\big[\lambda_0\,\mathbf{e}_1\big|\lambda_0\,\mathbf{e}_2\big|\dots\big|\lambda_0\,\mathbf{e}_k\big|\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{V}_{k+1}\big|\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{V}_{k+2}\big|\dots\big|\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{V}_n\big]$.

We see that $\mathbb{A}\cdot\mathbb{W}=\mathbb{W}\cdot\mathbb{B}$, or that $\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{W}=\mathbb{B}$. In fact:
$\mathbb{A}\cdot\mathbb{W}=\mathbb{A}\cdot\left[\mathbb{X}_1|\mathbb{X}_2|...|\mathbb{X}_k|\mathbb{V}_{k+1}|\mathbb{V}_{k+2}|...|\mathbb{V}_n\right]=$
$=\left[\mathbb{A}\cdot\mathbb{X}_1|\mathbb{A}\cdot\mathbb{X}_2|...|\mathbb{A}\cdot\mathbb{X}_k|\mathbb{A}\cdot\mathbb{V}_{k+1}|\mathbb{A}\cdot\mathbb{V}_{k+2}|...|\mathbb{A}\cdot\mathbb{V}_n\right]=$
$=\left[\lambda_0\,\mathbb{X}_1|\lambda_0\,\mathbb{X}_2|...|\lambda_0\,\mathbb{X}_k|\mathbb{A}\cdot\mathbb{V}_{k+1}|\mathbb{A}\cdot\mathbb{V}_{k+2}|...|\mathbb{A}\cdot\mathbb{V}_n\right]$, while
$\mathbb{W}\cdot\mathbb{B}=\mathbb{W}\cdot\left[\lambda_0\,\mathbf{e}_1|\lambda_0\,\mathbf{e}_2|...|\lambda_0\,\mathbf{e}_k|\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{V}_{k+1}|\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{V}_{k+2}|...|\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{V}_n\right]=$
$=\left[\mathbb{W}\cdot\lambda_0\,\mathbf{e}_1|\mathbb{W}\cdot\lambda_0\,\mathbf{e}_2|...|\mathbb{W}\cdot\lambda_0\,\mathbf{e}_k|\mathbb{W}\cdot\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{V}_{k+1}|...|\mathbb{W}\cdot\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{V}_n\right]=$
$=\left[\lambda_0\,\mathbb{W}\cdot\mathbf{e}_1|\lambda_0\,\mathbb{W}\cdot\mathbf{e}_2|...|\lambda_0\,\mathbb{W}\cdot\mathbf{e}_k|\mathbb{A}\cdot\mathbb{V}_{k+1}|\mathbb{A}\cdot\mathbb{V}_{k+2}|...|\mathbb{A}\cdot\mathbb{V}_n\right]=$
$=\left[\lambda_0\,\mathbb{X}_1|\lambda_0\,\mathbb{X}_2|...|\lambda_0\,\mathbb{X}_k|\mathbb{A}\cdot\mathbb{V}_{k+1}|\mathbb{A}\cdot\mathbb{V}_{k+2}|...|\mathbb{A}\cdot\mathbb{V}_n\right]=\mathbb{A}\cdot\mathbb{W}$.
So the matrices $\mathbb{A}$ and $\mathbb{B}$ are similar, and so they have the same characteristic polynomial and the same eigenvalues. But $\mathbb{B}_n=\left\|\begin{matrix}\lambda_0 & 0 & ... & 0 & v_{1,k+1} & \cdots & v_{1,n}\\ 0 & \lambda_0 & ... & 0 & v_{2,k+1} & \cdots & v_{2,n}\\ ... & ... & ... & ... & \cdots & \cdots & \cdots\\ 0 & 0 & ... & \lambda_0 & v_{k,k+1} & \cdots & v_{k,n}\\ 0 & 0 & ... & 0 & v_{k+1,k+1} & \cdots & v_{k+1,n}\\ ... & ... & ... & ... & \cdots & \cdots & \cdots\\ 0 & 0 & ... & 0 & v_{n,k+1} & \cdots & v_{n,n}\end{matrix}\right\|$, and so:
$|\mathbb{B}-\lambda\cdot\mathbb{I}|=(\lambda_0-\lambda)^k\cdot\mathrm{P}_{n-k}(\lambda)$.
If $\mathrm{P}_{n-k}(\lambda_0)=0$, the root λ_0 will have an algebraic multiplicity greater than k, and therefore it is $m^g_{\lambda_0}<m^a_{\lambda_0}$. If on the contrary it is $\mathrm{P}_{n-k}(\lambda_0)\neq 0$, we will have $m^g_{\lambda_0}=m^a_{\lambda_0}$. •

Example 87 : Given the matrix $\mathbb{A}=\left\|\begin{matrix}3 & 1 & 1\\ 2 & 4 & 2\\ 1 & 1 & 3\end{matrix}\right\|$, let us study its multiple eigenvalue and its algebraic and geometric multiplicity. It is:
$\begin{vmatrix}3-\lambda & 1 & 1\\ 2 & 4-\lambda & 2\\ 1 & 1 & 3-\lambda\end{vmatrix}=(2-\lambda)((4-\lambda)(3-\lambda)-2)-(2-\lambda)(2-4+\lambda)=$
$=(2-\lambda)\left(\lambda^2-8\lambda+12\right)=(2-\lambda)(\lambda-2)(\lambda-6)=0$.
And then its eigenvalues are: $\lambda_1=\lambda_2=2$ and $\lambda_3=6$.
To determine the eigenspace associated to $\lambda_1=\lambda_2=2$ we solve the homogeneous system:
$|\mathbb{A}-2\,\mathbb{I}|=\mathbb{O}$, whose matrix is $(\mathbb{A}-2\,\mathbb{I})=\left\|\begin{matrix}1 & 1 & 1\\ 2 & 2 & 2\\ 1 & 1 & 1\end{matrix}\right\|$.
Since $\mathrm{Rank}(\mathbb{A}-2\,\mathbb{I})=1$, the dimension of the eigenspace associated to $\lambda=2$ is equal to $m^g_2=3-\mathrm{Rank}(\mathbb{A}-2\,\mathbb{I})=2$, and so $m^g_2=m^a_2=2$. From the only remaining equation: $x+y+z=0$ we obtain the eigenvectors $(x,y,-x-y)$.
Since $m^g_2=2$, we must determine from this two independent vectors, which may be, for example, $(1,0,-1)$ and $(0,1,-1)$.
Every eigenvector with $\lambda=2$ can be written as: $k\,(1,0,-1)+h\,(0,1,-1), k,h\in\mathbb{R}$.
To complete a basis of $\mathbb{R}^3$ we choose the vector $(1,0,0)$.
If $\mathbb{W}=\left\|\begin{matrix}1 & 0 & 1\\ 0 & 1 & 0\\ -1 & -1 & 0\end{matrix}\right\|$ and since $\mathbb{W}^{-1}=\left\|\begin{matrix}0 & -1 & -1\\ 0 & 1 & 0\\ 1 & 1 & 1\end{matrix}\right\|$ we obtain:

$$\mathbb{B}=\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{W}=\begin{Vmatrix}0&-1&-1\\0&1&0\\1&1&1\end{Vmatrix}\cdot\begin{Vmatrix}3&1&1\\2&4&2\\1&1&3\end{Vmatrix}\cdot\begin{Vmatrix}1&0&1\\0&1&0\\-1&-1&0\end{Vmatrix}=\begin{Vmatrix}2&0&-3\\0&2&2\\0&0&6\end{Vmatrix}.$$

The matrix $\mathbb{B}$ too has the eigenvalue $\lambda=2$ whose algebraic multiplicity is equal to 2.
Also for the matrix $\mathbb{A}$ it is $m_2^g=m_2^a$.

Instead, let us consider the matrix $\mathbb{A}=\begin{Vmatrix}1&0&-1\\1&2&1\\1&1&1\end{Vmatrix}$. Let us study its multiple eigenvalue and its algebraic and geometric multiplicity. It is:

$$\begin{vmatrix}1-\lambda&0&-1\\1&2-\lambda&1\\1&1&1-\lambda\end{vmatrix}=(1-\lambda)((2-\lambda)(1-\lambda)-1)-1(1-2+\lambda)=$$

$$=(1-\lambda)\left(\lambda^2-3\lambda+2\right)=(1-\lambda)(\lambda-1)(\lambda-2)=0.$$

Then its eigenvalues are: $\lambda_1=\lambda_2=1$ and $\lambda_3=2$.
To determine the eigenspace associated to $\lambda_1=\lambda_2=1$ we solve the homogeneous system:
$|\mathbb{A}-1\cdot\mathbb{I}|=\mathbb{O}$, whose matrix is $(\mathbb{A}-1\cdot\mathbb{I})=\begin{Vmatrix}0&0&-1\\1&1&1\\1&1&0\end{Vmatrix}$.

Since $\text{Rank}(\mathbb{A}-1\cdot\mathbb{I})=2$, the dimension of the eigenspace associated to $\lambda=1$ is equal to $m_1^g=3-\text{Rank}(\mathbb{A}-1\cdot\mathbb{I})=1$, and so $m_1^g=1<m_1^a=2$.

If we had followed the outline of the proof of Teorema 29, we would not have had to find that $m_1^a=2$ while $m_1^g=1$; we would have had to start from the assumption that the dimension of the eigenspace with $\lambda=1$ was equal to 1 and then we had to find out its algebraic multiplicity. But this approach is a purely theoretical and not practical one.

From equations $\begin{cases}-z=0\\x+y=0\end{cases}$ we find the eigenvector $(1,-1,0)$.
To complete a basis of $\mathbb{R}^3$ we choose the two vectors $(1,0,0)$ and $(0,0,1)$.

If $\mathbb{W}=\begin{Vmatrix}1&1&0\\-1&0&0\\0&0&1\end{Vmatrix}$ and since $\mathbb{W}^{-1}=\begin{Vmatrix}0&-1&0\\1&1&0\\0&0&1\end{Vmatrix}$, we obtain:

$$\mathbb{B}=\mathbb{W}^{-1}\cdot\mathbb{A}\cdot\mathbb{W}=\begin{Vmatrix}0&-1&0\\1&1&0\\0&0&1\end{Vmatrix}\cdot\begin{Vmatrix}1&0&-1\\1&2&1\\1&1&1\end{Vmatrix}\cdot\begin{Vmatrix}1&1&0\\-1&0&0\\0&0&1\end{Vmatrix}=\begin{Vmatrix}1&-1&-1\\0&2&0\\0&1&1\end{Vmatrix}.$$

So: $|\mathbb{B}-\lambda\,\mathbb{I}|=(1-\lambda)^2(2-\lambda)$, and then $m_1^a=2>m_1^g=1$.

Theorem 30 : If λ_i is a simple root of the characteristic polynomial, then $m_i^g=m_i^a=1$.
Theorem 31 : A matrix $\mathbb{A}_n$ has exactly n linearly independent eigenvectors if and only if for every eigenvalue λ_i it is: $m_i^g=m_i^a$.

MATRIX DIAGONALIZABILITY AND TRIANGULARISABILITY

The search for the roots of the characteristic polynomial can be very difficult, if not impossible unless stopping to approximate values, when the order of the matrix is quite high.

Since it is very easy to find the eigenvalues of diagonal or triangular matrices, and since similar matrices have the same spectrum, it may be useful to determine when a matrix is similar to a diagonal or to a triangular matrix.
In these cases we deal with diagonalizable or triangularisable matrices.

Definition 45 : A square matrix $\mathbb{A}_n$ is called diagonalizable if it is similar to a diagonal matrix $\mathbb{D}$; a square matrix $\mathbb{A}_n$ is called triangularisable if it is similar to a triangular matrix $\mathbb{T}$.

From the definition of similar matrices, we have the following:
Theorem 32 : A square matrix $\mathbb{A}_n$ is a diagonalizable one if there exists a non-singular matrix $\mathbb{P}$ such that $\mathbb{A} \cdot \mathbb{P} = \mathbb{P} \cdot \mathbb{D}$ or, equivalently, such that $\mathbb{A} = \mathbb{P} \cdot \mathbb{D} \cdot \mathbb{P}^{-1}$, where $\mathbb{D}$ is a diagonal matrix.
And also
Theorem 33 : A square matrix $\mathbb{A}_n$ is a triangularisable one if there exists a non-singular matrix $\mathbb{P}$ such that $\mathbb{A} \cdot \mathbb{P} = \mathbb{P} \cdot \mathbb{T}$ or, equivalently, such that $\mathbb{A} = \mathbb{P} \cdot \mathbb{T} \cdot \mathbb{P}^{-1}$, where $\mathbb{T}$ is a triangular (upper or lower) matrix.

For an orthogonal matrix it is $\mathbb{A}^{\mathrm{T}} = \mathbb{A}^{-1}$, and so, if diagonalizability or triangularisability were performed with an orthogonal matrix $\mathbb{P}$, we have evident saving in the calculations if we can simply calculate $\mathbb{P}^{\mathrm{T}}$ instead of $\mathbb{P}^{-1}$.

Let us see firstly when a matrix is diagonalizable. The following is valid:
Theorem 34 : A matrix $\mathbb{A}_n$ is diagonalizable if and only if it has exactly n linearly independent eigenvectors.
Proof: Let us verify firstly that the condition is sufficient.
If the matrix $\mathbb{A}$ is diagonalizable then there exists, by definition, a non-singular matrix $\mathbb{P}$ such that $\mathbb{A} \cdot \mathbb{P} = \mathbb{P} \cdot \mathbb{D}$. If $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_n$ are the columns of the non-singular matrix $\mathbb{P}$, they are linearly independent vectors.
From $\mathbb{A} \cdot \mathbb{P} = \mathbb{P} \cdot \mathbb{D}$, if $\lambda_1, \lambda_2, ..., \lambda_n$ are the entries of the main diagonal of the matrix $\mathbb{D}$, we obtain:
$\mathbb{A} \cdot \left[\mathbb{X}_1 \middle| \mathbb{X}_2 \middle| ... \middle| \mathbb{X}_n\right] = \left[\mathbb{X}_1 \middle| \mathbb{X}_2 \middle| ... \middle| \mathbb{X}_n\right] \cdot \mathbb{D}$, which can be written as:
$\left[\mathbb{A} \cdot \mathbb{X}_1 \middle| \mathbb{A} \cdot \mathbb{X}_2 \middle| ... \middle| \mathbb{A} \cdot \mathbb{X}_n\right] = \left[\lambda_1 \mathbb{X}_1 \middle| \lambda_2 \mathbb{X}_2 \middle| ... \middle| \lambda_n \mathbb{X}_n\right]$
from which, by equating the columns, we obtain:
$\mathbb{A} \cdot \mathbb{X}_1 = \lambda_1 \mathbb{X}_1,\ \mathbb{A} \cdot \mathbb{X}_2 = \lambda_2 \mathbb{X}_2, ...,\ \mathbb{A} \cdot \mathbb{X}_n = \lambda_n \mathbb{X}_n$.
From this we see that $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_n$ are eigenvectors with the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, and they are also linearly independent eigenvectors since the matrix $\mathbb{P}$ is non-singular.
Then we verify that the condition is necessary.
If $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_n$ are n linearly independent eigenvectors with the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, let $\mathbb{P}$ be the matrix having $\mathbb{X}_1, \mathbb{X}_2, ..., \mathbb{X}_n$ as columns.
The n equalities:
$\mathbb{A} \cdot \mathbb{X}_1 = \lambda_1 \mathbb{X}_1$, $\mathbb{A} \cdot \mathbb{X}_2 = \lambda_2 \mathbb{X}_2$,..., $\mathbb{A} \cdot \mathbb{X}_n = \lambda_n \mathbb{X}_n$
can be written in block-matrix form as:
$\left[\mathbb{A} \cdot \mathbb{X}_1 \middle| \mathbb{A} \cdot \mathbb{X}_2 \middle| ... \middle| \mathbb{A} \cdot \mathbb{X}_n\right] = \left[\lambda_1 \mathbb{X}_1 \middle| \lambda_2 \mathbb{X}_2 \middle| ... \middle| \lambda_n \mathbb{X}_n\right]$
that we can rewrite, if $\mathbb{D}$ is the diagonal matrix having $\lambda_1, \lambda_2, ..., \lambda_n$ as its main diagonal entries, as:
$\mathbb{A} \cdot \left[\mathbb{X}_1 \middle| \mathbb{X}_2 \middle| ... \middle| \mathbb{X}_n\right] = \left[\mathbb{X}_1 \middle| \mathbb{X}_2 \middle| ... \middle| \mathbb{X}_n\right] \cdot \mathbb{D}$ i.e. $\mathbb{A} \cdot \mathbb{P} = \mathbb{P} \cdot \mathbb{D}$.
And so, being $\mathbb{P}$ an invertible matrix since it has linearly independent columns, we obtain:
$\mathbb{A} = \mathbb{P} \cdot \mathbb{D} \cdot \mathbb{P}^{-1}$, i.e. the matrix $\mathbb{A}$ is diagonalizable. •

The matrix $\mathbb{P}$ having the eigenvectors as columns is called modal matrix (in Italian literature "matrice modale"), and provides, by means of the $\mathbb{D} = \mathbb{P}^{-1} \cdot \mathbb{A} \cdot \mathbb{P}$, the procedure for the diagonalization of $\mathbb{A}$.

From the previous theorems, we deduce also the following:
Theorem 35 : A matrix $\mathbb{A}_n$ is diagonalizable if and only if $m_i^g = m_i^a$, $\forall$ eigenvalue λ_i.
Theorem 36 : If a matrix $\mathbb{A}_n$ has n distinct eigenvalues then the matrix is a diagonalizable one.

TRIANGULARISABLE MATRICES
As previously seen, matrices having multiple eigenvalues may not be diagonalizable. For at least one multiple eigenvalue in fact it may happen that $m_i^g < m_i^a$.
For those matrices $\mathbb{A}_n$ that don't have exactly n independent eigenvectors, we can use the similarity to a triangular matrix, since the following is valid:
Theorem 37 (Schur decomposition) : Every matrix $\mathbb{A}_n$ is similar to an upper triangular matrix $\mathbb{T}$. Indeed, there is an orthogonal matrix $\mathbb{U}$ such that: $\mathbb{A} \cdot \mathbb{U} = \mathbb{U} \cdot \mathbb{T}$ or $\mathbb{U}^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{U} = \mathbb{T}$.

SYMMETRIC MATRICES
Let us examine diagonalizability for symmetric matrices, i.e. the matrices for which $\mathbb{A}^{\mathrm{T}} = \mathbb{A}$.
A first important property is the following:
Theorem 38 : The eigenvalues of a symmetric matrix are always real numbers.
Proof: Let $\mathbb{A} \cdot \mathbb{X} = \lambda \mathbb{X}$. Turning to the conjugate we have : $\overline{\mathbb{A}} \cdot \overline{\mathbb{X}} = \overline{\lambda}\, \overline{\mathbb{X}}$.
Since the entries of $\mathbb{A}$ are real numbers, it is $\overline{\mathbb{A}} \cdot \overline{\mathbb{X}} = \mathbb{A} \cdot \overline{\mathbb{X}} = \overline{\lambda}\, \overline{\mathbb{X}}$.
Taking the transpose we obtain: $\left(\mathbb{A} \cdot \overline{\mathbb{X}}\right)^{\mathrm{T}} = \left(\overline{\lambda}\, \overline{\mathbb{X}}\right)^{\mathrm{T}}$ or $\overline{\mathbb{X}}^{\mathrm{T}} \cdot \mathbb{A}^{\mathrm{T}} = \overline{\lambda}\, \overline{\mathbb{X}}^{\mathrm{T}}$.
Since $\mathbb{A}$ is symmetric, it is $\mathbb{A}^{\mathrm{T}} = \mathbb{A}$, so we obtain: $\overline{\mathbb{X}}^{\mathrm{T}} \cdot \mathbb{A} = \overline{\lambda}\, \overline{\mathbb{X}}^{\mathrm{T}}$.
From $\mathbb{A} \cdot \mathbb{X} = \lambda \mathbb{X}$, multiplying on the left by $\overline{\mathbb{X}}^{\mathrm{T}}$ we obtain: $\overline{\mathbb{X}}^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{X} = \lambda\, \overline{\mathbb{X}}^{\mathrm{T}} \cdot \mathbb{X}$;
from $\overline{\mathbb{X}}^{\mathrm{T}} \cdot \mathbb{A} = \overline{\lambda}\, \overline{\mathbb{X}}^{\mathrm{T}}$, multiplying on the right by $\mathbb{X}$ we obtain: $\overline{\mathbb{X}}^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{X} = \overline{\lambda}\, \overline{\mathbb{X}}^{\mathrm{T}} \cdot \mathbb{X}$.
And so $\lambda\, \overline{\mathbb{X}}^{\mathrm{T}} \cdot \mathbb{X} = \overline{\lambda}\, \overline{\mathbb{X}}^{\mathrm{T}} \cdot \mathbb{X}$, and since $\overline{\mathbb{X}}^{\mathrm{T}} \cdot \mathbb{X} \neq 0$, it follows $\lambda = \overline{\lambda}$, i.e. $\lambda \in \mathbb{R}$. •

With respect to the eigenvectors, we have the following:
Theorem 39 : In a symmetric matrix, any two eigenvectors with distinct eigenvalues are orthogonal.
I.e. to distinct eigenvalues orthogonal eigenvectors correspond.
Proof: Let $\mathbb{X}_1$ and $\mathbb{X}_2$ be two eigenvectors with distinct eigenvalues λ_1 and λ_2.
Since this product $\mathbb{X}_1^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{X}_2$ gives a real number (and $k = k^{\mathrm{T}}$), it is:
$\mathbb{X}_1^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{X}_2 = \left(\mathbb{X}_1^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{X}_2\right)^{\mathrm{T}} = \mathbb{X}_2^{\mathrm{T}} \cdot \mathbb{A}^{\mathrm{T}} \cdot \mathbb{X}_1 = \mathbb{X}_2^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{X}_1$,
because the matrix is symmetric, and then, as $\mathbb{X}_1$ and $\mathbb{X}_2$ are eigenvectors, it follows that:
$\mathbb{X}_1^{\mathrm{T}} \cdot \lambda_2 \mathbb{X}_2 = \lambda_2 \mathbb{X}_1^{\mathrm{T}} \cdot \mathbb{X}_2 = \mathbb{X}_2^{\mathrm{T}} \cdot \lambda_1 \mathbb{X}_1 = \lambda_1 \mathbb{X}_2^{\mathrm{T}} \cdot \mathbb{X}_1$.
But $\mathbb{X}_1^{\mathrm{T}} \cdot \mathbb{X}_2 = \mathbb{X}_2^{\mathrm{T}} \cdot \mathbb{X}_1$, because the scalar product of the vectors $\mathbb{X}_1$ and $\mathbb{X}_2$ gives as result a real number k, and so we obtain: $\lambda_1 \mathbb{X}_2^{\mathrm{T}} \cdot \mathbb{X}_1 = \lambda_2 \mathbb{X}_2^{\mathrm{T}} \cdot \mathbb{X}_1$, and this equality is possible, since $\lambda_1 \neq \lambda_2$, if and only if $\mathbb{X}_2^{\mathrm{T}} \cdot \mathbb{X}_1 = 0$, i.e. if and only if $\mathbb{X}_1$ and $\mathbb{X}_2$ are orthogonal.•

As a first consequence, from the previous theorems, we have:
Theorem 40 : If a symmetric matrix $\mathbb{A}$ has all distinct eigenvalues, then there exists an orthogonal matrix $\mathbb{U}$ that diagonalizes $\mathbb{A}$.
Proof : If the eigenvalues are all distinct, the corresponding eigenvectors are not only linearly independent, but, for the previous theorem, are also two by two orthogonal vectors.

When constructing the modal matrix it is enough to normalize these eigenvectors to obtain an orthogonal matrix $\mathbb{U}$ such that: $\mathbb{A} \cdot \mathbb{U} = \mathbb{U} \cdot \mathbb{D}$. •

However, applying to a symmetric matrix the Schur's decomposition theorem, we obtain a more general result, namely that every symmetric matrix is always diagonalizable by means of an orthogonal matrix. Indeed the following is valid:
Theorem 41 (spectral theorem) : Every symmetric real matrix can be diagonalized by an orthogonal matrix.
I.e.: For every symmetric real matrix $\mathbb{A}$ there exists a real orthogonal matrix $\mathbb{U}$ such that $\mathbb{D} = \mathbb{U}^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{U}$ is a diagonal matrix.
Proof: From Schur's theorem every matrix $\mathbb{A}$ is similar to an upper triangular matrix $\mathbb{T}$, i.e. there is an orthogonal matrix $\mathbb{U}$ such that $\mathbb{T} = \mathbb{U}^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{U}$.
But, since $\mathbb{A}$ is symmetric, we obtain:
$\mathbb{T}^{\mathrm{T}} = \left(\mathbb{U}^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{U}\right)^{\mathrm{T}} = \mathbb{U}^{\mathrm{T}} \cdot \mathbb{A}^{\mathrm{T}} \cdot \mathbb{U} = \mathbb{U}^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{U} = \mathbb{T}$
and so $\mathbb{T}^{\mathrm{T}} = \mathbb{T}$, that is, the triangular matrix $\mathbb{T}$ is a symmetric one, so it is a diagonal matrix.
Therefore $\mathbb{U}^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{U} = \mathbb{T} = \mathbb{D}$, i.e. $\mathbb{A}$ is diagonalizable by an orthogonal matrix. •

Example 88 : Given the symmetric matrix $\mathbb{A} = \left\| \begin{matrix} 2 & 1 & -2 \\ 1 & 2 & 2 \\ -2 & 2 & -1 \end{matrix} \right\|$, let us determine the orthogonal matrix that diagonalizes $\mathbb{A}$.
Searching for the roots of the characteristic polynomial, we have:
$$|\mathbb{A} - \lambda\,\mathbb{I}| = \left\| \begin{matrix} 2-\lambda & 1 & -2 \\ 1 & 2-\lambda & 2 \\ -2 & 2 & -1-\lambda \end{matrix} \right\| = -\lambda^3 + 3\lambda^2 + 9\lambda - 27 = 0$$
and so the roots: $\lambda_1 = 3$, $\lambda_2 = 3$, $\lambda_3 = -3$, therefore three real roots, one is a multiple (double) root and one is a simple one.
To determine the eigenspace associated to $\lambda_1 = \lambda_2 = 3$ we solve the homogeneous system:
$$|\mathbb{A} - 3\,\mathbb{I}| = \mathbb{O}\text{, or } \left\| \begin{matrix} -1 & 1 & -2 \\ 1 & -1 & 2 \\ -2 & 2 & -4 \end{matrix} \right\| \cdot \left\| \begin{matrix} x \\ y \\ z \end{matrix} \right\| = \mathbb{O}\,.$$
Since $\mathrm{Rank}(\mathbb{A} - 3\,\mathbb{I}) = 1$, the dimension of the eigenspace associated to $\lambda = 3$ is equal to $m_3^g = 3 - \mathrm{Rank}(\mathbb{A} - 3\,\mathbb{I}) = 2$, and so $m_3^g = m_3^a = 2$.
Obviously, since the matrix is a symmetric one.
From the only remaining equation: $x - y + 2z = 0$ we obtain $x = y - 2z$ and then the eigenvectors $(y - 2z, y, z)$.
Since $m_3^g = 2$, we must determine from these two independent eigenvectors; if we choose $y = 1$ and $z = 0$ the first vector is $\mathbb{X}_1 = (1, 1, 0)$.
To determine the second we must remember that we want orthogonal eigenvectors with the same eigenvalue.
Requiring that $(1, 1, 0) \cdot (y - 2z, y, z) = 0$ we derive the condition $z = y$, for which the other eigenvector will be, choosing $y = 1$, $\mathbb{X}_2 = (-1, 1, 1)$.
To determine the eigenspace associated to $\lambda_3 = -3$ we solve the homogeneous system:
$$|\mathbb{A} + 3\,\mathbb{I}| = \mathbb{O}\text{, or } \left\| \begin{matrix} 5 & 1 & -2 \\ 1 & 5 & 2 \\ -2 & 2 & 2 \end{matrix} \right\| \cdot \left\| \begin{matrix} x \\ y \\ z \end{matrix} \right\| = \mathbb{O}\,.$$

Since $\lambda = -3$ is a simple root, the dimension of the eigenspace associated to $\lambda = -3$ is equal to 1. The homogeneous system becomes:

$$\begin{cases} 5x + y - 2z = 0 \\ x + 5y + 2z = 0 \\ -2x + 2y + 2z = 0 \end{cases} \quad \text{whose solutions are } \begin{cases} x = -y \\ z = -2y \end{cases}.$$

From the eigenvectors $(-y, y, -2y)$ we choose $\mathbb{X}_3 = (1, -1, 2)$, and we see that this vector is orthogonal to both the eigenvectors with $\lambda = 3$.

To have the modal orthogonal matrix that diagonalizes the symmetric matrix we should finally normalize the three eigenvectors we have found, and so we obtain:

$$\mathbb{P} = \begin{Vmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{Vmatrix} \text{ the modal orthogonal matrix.}$$

Therefore:

$$\mathbb{U}^{\mathrm{T}} \cdot \mathbb{A} \cdot \mathbb{U} = \begin{Vmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{Vmatrix} \cdot \begin{Vmatrix} 2 & 1 & -2 \\ 1 & 2 & 2 \\ -2 & 2 & -1 \end{Vmatrix} \cdot \begin{Vmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{Vmatrix} =$$

$$= \begin{Vmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{Vmatrix} = \mathbb{D}.$$