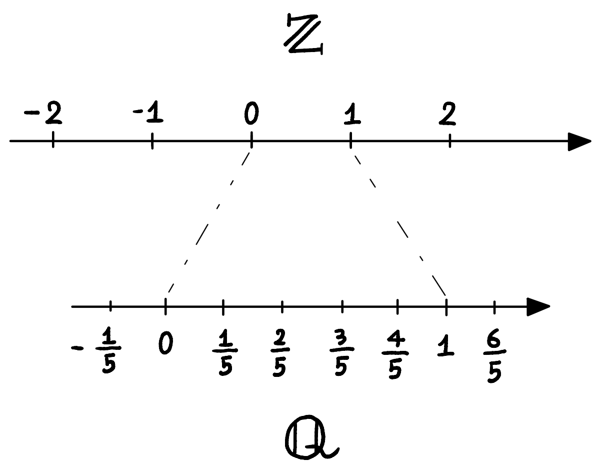
Lecture 1: Real numbers

In this lecture we briefly discuss the set of real numbers $\mathbb{R}$, and show how such set can be constructed based on successive extensions of the set of natural numbers

$$\mathbb{N} = \{1, 2, 3, \ldots\}.$$

The main steps are:

1. Construct $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, ...\}$ (integer numbers).
2. Construct $\mathbb{Q} = \left\{ \dfrac{p}{q} : \quad p, q \in \mathbb{Z}, \quad q \neq 0 \right\}$ (rational numbers).
3. Define the set of irrational numbers and add it to the set of rational numbers to obtain the set of real numbers.



Of course we can go on and look for four more rational numbers between 0 and 1/5, i.e.,

$$\left\{ \frac{1}{25}, \frac{2}{25}, \frac{3}{25}, \frac{4}{25} \right\}, \tag{1}$$

etc. Clearly, we have[1]

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}. \tag{2}$$

Rational numbers either have a finite number of decimal digits or an infinite number of decimal digits repeating periodically. For example:

$$\begin{aligned}
\frac{3}{4} &= 0.75 \qquad \text{(finite number of decimals)},\\
\frac{1}{3} &= 0.333333\ldots = 0.\overline{3} \qquad \text{(infinite decimals repeating periodically)},\\
\frac{1}{7} &= 0.\underbrace{142857}\underbrace{142857}\ldots = 0.\overline{142857} \qquad \text{(infinite decimals repeating periodically)}.
\end{aligned}$$

[1] In mathematics, the symbol "$\subset$" means "subset of". Note that $\mathbb{N}$ is a subset of $\mathbb{Z}$ because $\mathbb{Z}$ includes all integer numbers $\{1, 2, 3, \ldots\}$. In addition, $\mathbb{Z}$ includes the negative of all integer numbers and the element zero $\{0\}$. Of course, the set of rational numbers $\mathbb{Q}$ includes, by definition, the set of natural numbers as well as the set of integer numbers.

Moreover, the sum or the products of two rational numbers is still a rational number. For example,

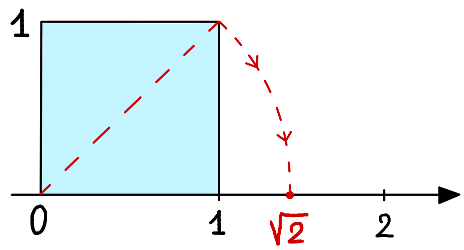
$$\frac{1}{3} + \frac{2}{15} = \frac{7}{15}, \qquad \frac{1}{3} \times \frac{2}{15} = \frac{2}{45}. \tag{3}$$

You may have heard that the set of rational numbers is not "complete". In other words, there are numbers on the (continuum) horizontal lines sketched above that cannot be represented as a ratio between two integers, i.e., as a rational number. One of such numbers is the square root of 2

$$\sqrt{2} = 1.4142135623730950488016887242096980785696718 7\cdots \tag{4}$$

which has an infinite number of decimals that do not repeat periodically as in the case of rational numbers.

The number $\sqrt{2}$ is indeed an *irrational* number that can be visualized by rotating the diagonal of a unit square by 45 degrees clockwise as follows



Recall, in fact, that by the Pythagorean theorem, the length of the diagonal of the unit square is

$$\sqrt{1^2 + 1^2} = \sqrt{2}. \tag{5}$$

Remarkably, no matter how hard we try to cover the continuum line with rational numbers we find out that the number of "holes" left to be filled is *infinite* and *uncountable* (cardinality larger than $\mathbb{N}$). Hereafter we rigorously show that $\sqrt{2}$ is indeed not a rational number. To this end, we formulate the following Theorem[2]:

Theorem. There is no rational number the square of which equals 2. Equivalently, there is no rational number that equals $\sqrt{2}$.

Proof. Suppose that there exists a rational number p/q (irreducible fraction[3]) the square of which equals two:

$$\left(\frac{p}{q}\right)^2 = 2 \quad \Rightarrow \quad p^2 = 2q^2. \tag{6}$$

Clearly, $p^2 = 2q^2$ is an even natural number (2 times the natural number q^2 is necessarily even). This implies that p is an even integer number, and therefore can be written as

$$p = 2s \qquad \text{for some } s \in \mathbb{Z}. \tag{7}$$

[2]A Theorem is a statement that is not self-evident but can be proved (or disproved) by a sequence of logical or mathematical operations.

[3]An irreducible fraction is a fraction that cannot be simplified any further. For example, 3/2 is an irreducible fraction while 6/4 is not an irreducible fraction as both the numerator and the denominator can be divided by 2 to obtain 3/2.

Substituting equation (7) into (6) yields

$$\frac{2^2 s^2}{q^2} = 2 \quad \Rightarrow \quad q^2 = 2s^2 \quad \Rightarrow \quad q \quad \text{is an even number.} \tag{8}$$

But this contradicts the fact that p/q is an irreducible fraction. Indeed, we just concluded that both p and q are divisible by 2 since they are both even numbers. In summary, the assumption that there exists a rational number p/q equals to $\sqrt{2}$ yields a contradiction, and therefore the assumption must be wrong, meaning that such a rational number cannot exist.

□

The technique we just used to prove the Theorem above is known as "proof by contradiction" ("Reductio ad absurdum" in Latin). Essentially it is a form of argument that establishes a statement by arriving at a contradiction or at something that is impossible or absurd, even when the initial assumption is the negation of the statement to be proved. For example, let us prove the following statement

"There isn't a smallest positive rational number"

by using the proof by contradiction technique. To this end, we first assume that there is a smallest positive rational number (negation of the statement) and immediately notice that dividing such rational number by 2 (or any other integer number larger than 2) yields another rational number that is smaller than the one we started with. This contradicts the hypothesis that there is a smallest rational number. Therefore the statement "There isn't a smallest rational number" must be true.

There are many other examples of numbers that cannot be represented as a ratio between two integer numbers (i.e., rationals). Such numbers are called *irrational numbers*, and they have an infinite number of decimals (non-repeating). Moreover, there are infinite irrational numbers (uncountably many!). Well known examples of irrational numbers are: $\pi = 3.141592653\cdots$, $\sqrt{2} = 1.41421356\cdots$, $\sqrt{3} = 1.73205080\cdots$, $e = 2.71828182845\cdots$ (Napier number).

Remarkably, all irrational numbers can be obtained as limits of suitable sequences of rational numbers. For example,

$$\pi = 4\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots\right) \qquad \text{(Leibnitz sequence)}$$

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \cdots$$

Similarly, $\sqrt{2}$ can be obtained by iterating (an infinite number of times) the following sequence of rational numbers:

$$S_{n+1} = \frac{S_n}{2} + \frac{1}{S_n} \qquad S_0 = 1 \tag{9}$$

i.e.,

$$S_0 = 1, \qquad S_1 = \frac{3}{2} = 1.5, \qquad S_2 = \frac{17}{12} = 1.41\overline{6}, \qquad S_3 = \frac{577}{408} = 1.414\overline{2156862745098039}, \quad \ldots$$

The sequences above are not unique, meaning that there are many other sequences of rationals converging to the same irrationals. In any case, such sequences do exist, and they can represent (in their limit) all irrational numbers[4]. By adding (formally) the set of irrational numbers (denoted by $\mathbb{I}$), which is uncountable, to the set of the rationals $\mathbb{Q}$ we obtain the set of *real numbers*[5]:

$$\mathbb{R} = \mathbb{Q} \cup \mathbb{I} \tag{10}$$

A remarkable result of number theory says that both $\mathbb{Q}$ and $\mathbb{I}$ are "dense" in $\mathbb{R}$. This means that any real number can be obtained as limit point of sequences of rational numbers or sequences of irrational numbers. Moreover, between any two distinct real numbers there always exists a rational number and an irrational one.

Axiomatic definition of $\mathbb{R}$

The set of real numbers can be defined in an axiomatic way. An axiom is statement or a proposition which is regarded as being established, accepted, or self-evidently true. Hence, by defining $\mathbb{R}$ in terms of axioms we specify properties of $\mathbb{R}$ that are self-evidently true.

Field axioms. The set of real number $\mathbb{R}$ is a (algebraic) field, i.e., it is a set in which we can define two operations (addition "+" and multiplication[6]) with the following properties:

1. Associative property[7]:
$$\forall x, y, z \in \mathbb{R}, \qquad (x+y)+z = x+(y+z) \quad \text{and} \quad (xy)z = x(yz).$$
2. Commutative property:
$$\forall x, y \in \mathbb{R}, \;\; x+y = y+x \quad \text{and} \quad xy = yx.$$
3. Additive neutral element:
$$\text{There exists an element of } \mathbb{R}, \text{ denoted by } 0, \text{ such that } \forall x \in \mathbb{R},\; x+0 = x.$$
4. Multiplicative neutral element:
$$\text{There exists an element of } \mathbb{R}, \text{ denoted by } 1, \text{ such that } \forall x \in \mathbb{R},\; 1x = x.$$
5. Inverse with respect to addition:
$$\forall x \in \mathbb{R}, \text{ there exists } y \in \mathbb{R} \text{ such that } x+y = 0\, (y = -x) \qquad (y \text{ is the opposite of } x).$$
6. Inverse with respect to multiplication:
$$\forall x \in \mathbb{R} \setminus \{0\}, \text{ there exists } y \in \mathbb{R} |\; x \cdot y = 1 \qquad (y \text{ is the inverse of } x).$$

[4] The sequences are actually used in practice to compute approximations of irrational numbers. For example, in 2016 Ron Watkins used the sequence (9) to compute 10 trillion digits of $\sqrt{2}$ (see http://www.numberworld.org/digits/Sqrt(2)).

[5] Note that $\mathbb{R}$ can also be thought of as $\mathbb{Q}$ plus all limit points of converging sequences of rational numbers.

[6] The multiplication operation between two elements $x, y \in \mathbb{R}$ is denoted simply as xy.

[7] In mathematics, the symbol $\forall$ means "for all", while the symbol $\in$ means "in". Hence, writing $\forall x, y, z \in \mathbb{R}$ can be spelled out as follows: "for all x, y and z in the set of real numbers".

7. Distributive property of multiplication:
$$\forall x, y, z \in \mathbb{R} \quad x(y+z) = xy + xz.$$

Remark: The set $\mathbb{R}$ is closed under addition and multiplication. This means that addition and product of real numbers is still a real number. In general, any set satisfying properties 1.-7. is a called a (algebraic) *field*. In particular, it can be verified that: 1) $\mathbb{N}$ is not a field; 2) $\mathbb{Z}$ is not a field (the inverse with respect to multiplication is not in $\mathbb{Z}$); 3) $\mathbb{Q}$ is a field; 4) $\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$ is not a field (the product of two irrationals can be an integer $\sqrt{2}\sqrt{2} = 2$).

By using the field axioms it is easy define subtraction and division between two real numbers as:

Subtraction: $\forall x, y \in \mathbb{R},\ x - y = x + (-y)$ (subtraction seen as adding the opposite of y).

Division: $\forall x, y \in \mathbb{R}$, $y \neq 0$, $x/y = xy^{-1}$ (division seen as multiplying by the inverse of y).

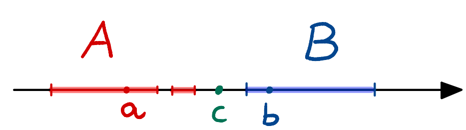
Ordering axioms. The field of real numbers $\mathbb{R}$ is totally ordered, i.e., we can define in $\mathbb{R}$ an ordering relation $\leq$ such that for all $x, y \in \mathbb{R}$:

1. $x \leq y \quad \Rightarrow \quad x + z \leq y + z, \qquad \forall z \in \mathbb{R}.$
2. $x \leq y \quad \Rightarrow \quad xz \leq yz, \qquad \forall z \geq 0.$

The mathematical symbol $\leq$ means "less or equal". Similarly, "$\geq$" means "greater or equal". So the ordering axiom number 2. can be phrased as follows: "Let x and y be two arbitrary real numbers; if x is smaller or equal than y, and z is any non-negative real number, then xz is smaller or equal than yz."

Remark: The ordering axioms say that all elements of $\mathbb{R}$ are ordered, i.e., we can always tell which element is bigger or smaller than any other element. That is why the lines sketched at Page 1 and Page 2 have one arrow (not two!) that indicates the direction in which the numbers are increasing.

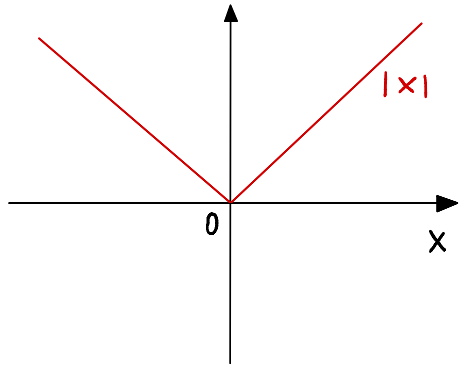
Completeness axiom. $\mathbb{R}$ is a field that is totally ordered and complete. "Complete" means that for every subsets $A, B \subseteq \mathbb{R}$ not empty and separated (i.e., such that $a \leq b$ $\forall a \in A$ and $\forall b \in B$) there exists at least one $c \in \mathbb{R}$ such that $a \leq c \leq b$.



The completeness axiom assures that the set $\mathbb{R}$ has no "holes" in it. Also, it can be shown that between two real numbers there is always an irrational and a rational, and between two rational numbers there is always a real number and an irrational number.

Absolute value. The absolute value of a real number is a function defined as
$$\begin{aligned} |\cdot| &: \mathbb{R} \to \mathbb{R}^+ \\ x \to |x| &= \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases} \end{aligned} \tag{11}$$

Here, $\mathbb{R}^+$ denotes the set of non-negative real numbers, i.e., $\mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\}$.



The absolute value function satisfies a certain number of properties summarized in the following Theorem. Each of the properties can be proved based on the definition of $|\cdot|$.

Theorem (Properties of the absolute value)**.** Let $a, b \in \mathbb{R}$. Then we have

1. $|a| \geq 0$
2. $|a| = 0 \Leftrightarrow a = 0$
3. $|a| = |-a|$
4. $-|a| \leq a \leq |a|$
5. $|a|^2 = a^2$
6. $|ab| = |a||b|$
7. $|a+b| \leq |a| + |b|$ (Triangle inequality)
8. $|a+b| \geq ||a| - |b||$ (Reverse triangle inequality)

Proof. Let us prove the triangle inequality and the reverse triangle inequality. For every $a, b \in \mathbb{R}$ we have[8]

$$a + b \leq |a| + |b| \quad \text{and} \quad a + b \geq -|a| - |b|. \tag{12}$$

This implies that

$$-\left(|a| + |b|\right) \leq a + b \leq |a| + |b| \tag{13}$$

i.e.,

$$|a+b| \leq |a| + |b| \quad \text{for all} \quad a, b \in \mathbb{R}. \tag{14}$$

The last step follows from the fact that if c is any non-negative real number then

$$|a| \leq c \quad \Leftrightarrow \quad -c \leq a \leq c. \tag{15}$$

[8]To prove (12) we notice that for all $a, b \in \mathbb{R}$ we have $a \leq |a|$ and $b \leq |b|$. Therefore $a + b \leq |a| + |b|$. Similarly, we have $-|a| \leq a$ and $-|b| \leq b$, which imply that $-|a| - |b| \leq a + b$. Multiplying the last inequality by -1 yields $-(a + b) \leq |a| + |b|$.

Similarly, for the reverse triangle inequality we observe that

$$a = a + b - b \quad \Rightarrow |a| \leq |a + b| + |b|\,, \tag{16}$$
$$b = b + a - a \quad \Rightarrow |b| \leq |a + b| + |a|\,. \tag{17}$$

Therefore

$$-|a + b| \leq |a| - |b| \leq |a + b| \tag{18}$$

i.e.,

$$||a| - |b|| \leq |a + b|\,. \tag{19}$$

□

By using the definition of absolute value we can define closed and open intervals of the real line with endpoint a and b ($a, b \in \mathbb{R}$, $a < b$) as

$$[a, b] = \{x \in \mathbb{R} |\ a \leq x \leq b\} = \left\{x \in \mathbb{R} : \left| x - \frac{a+b}{2} \right| \leq \frac{b-a}{2} \right\} \quad \text{(closed interval)}, \tag{20}$$

$$]a, b[= \{x \in \mathbb{R} |\ a < x < b\} = \left\{x \in \mathbb{R} : \left| x - \frac{a+b}{2} \right| < \frac{b-a}{2} \right\} \quad \text{(open interval)}. \tag{21}$$

Solution to linear and nonlinear equations

It is good practice to specify in which space we are looking for solutions of a certain equation. For instance, the following linear equation

$$2x = 1 \tag{22}$$

has no solution in $\mathbb{N}$ and no solution in $\mathbb{Z}$, but it has a unique solution in $\mathbb{Q}$ equal to $x = 1/2$, and of course a unique solution in $\mathbb{R}$ (since $\mathbb{Q} \subset \mathbb{R}$). Many nonlinear equations, however, do not admit a solution in $\mathbb{R}$. For example, the following polynomial (quadratic) equation

$$x^2 + 1 = 0 \tag{23}$$

has no solution in $\mathbb{R}$. In fact, the square of any real number x is non-negative, i.e., $x^2 \geq 0$ for all $x \in \mathbb{R}$. Hence, there is no element in $\mathbb{R}$ such that $x^2 = -1$, and therefore (23) has no solution in $\mathbb{R}$.

If we are interested in defining a solutions to equation (23), then we need to utilize a different set of numbers. In particular, such a set should include particular type of numbers the square of which is negative and real. As we will see such numbers are called *imaginary numbers*, and will be described in detail in the next lecture. Imaginary numbers are a subset of a more general set of numbers which is *complex numbers* and denoted as $\mathbb{C}$. Complex numbers were historically developed to make sense of solutions of polynomial equations (i.e., zeros of polynomials). For instance it was shown that:

Theorem (Fundamental theorem of algebra)**.** Every non-constant polynomial of the form

$$p_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 \tag{24}$$

with real or complex coefficients $\{a_n, \ldots, a_0\}$ has at least one complex root.

By applying this theorem recursively, it can be shown that every polynomial of degree n has exactly n complex roots (which may not be all distinct).

Complex numbers and complex functions pay a fundamental role in a variety of applications, e.g., series expansions of periodic functions, signal processing, solution to PDEs via Fourier series/transforms, quantum mechanics (e.g., Schrödinger equation), fluid dynamics (e.g., Joukowsky transformations for airfoil design), conformal maps, nonlinear dynamics and control, image processing, wave propagation, etc.

Lecture 2: Complex numbers

The quadratic equation

$$z^2 + 1 = 0 \tag{1}$$

has no solution in $\mathbb{R}$. In fact, there is no real number such that $z^2 = -1$ (recall that the square of any real number is either positive or equal to zero). However, we can still define solutions of equation (1), but we have to seek them in a different set of numbers. In particular, such a set must include new types of numbers the square of which is a negative real number. These numbers are called *imaginary numbers*.

Let "i" be one of such numbers, i.e., an imaginary number defined as

$$i = \sqrt{-1}. \tag{2}$$

Clearly, $z = i$ is a solution of equation (1). In fact,

$$z^2 + 1 = i^2 + 1 = -1 + 1 = 0. \tag{3}$$

Moreover, $z = -i$ is another solution of equation (1) since

$$z^2 + 1 = (-i)^2 + 1 = (-1)^2 i^2 + 1 = -1 + 1 = 0. \tag{4}$$

Next, consider the polynomial equation

$$z^2 + z + 1 = 0. \tag{5}$$

We can rearrange such polynomial equation as

$$\left(z + \frac{1}{2}\right)^2 + \frac{3}{4} = 0. \tag{6}$$

Upon definition of

$$u = z + \frac{1}{2} \tag{7}$$

we can write (6) as

$$u^2 = -\frac{3}{4} \quad \Leftrightarrow \quad u_{1,2} = \pm i\frac{\sqrt{3}}{2}. \tag{8}$$

Substituting $u_{1,2}$ back into (7) yields the following two solutions to equation (5)

$$z_{1,2} = -\frac{1}{2} \pm i\frac{\sqrt{3}}{2}. \tag{9}$$

This suggests that the new set of numbers we are interested in (at lest from the viewpoint of solving quadratic polynomial equations) has the form

$$z = x + iy \qquad \text{(complex number)}, \tag{10}$$

where x and y are real numbers and i is the imaginary unit defined in equation (2). Specifically, x is called the *real part* of z, and y is called the *imaginary part* of z. The real and imaginary parts of z are often denotes as

$$\text{Re}(z) = x \qquad \text{and} \qquad \text{Im}(z) = y. \tag{11}$$

Numbers of the form (10) are called *complex numbers*. The set of all complex numbers will be denoted by

$$\mathbb{C} = \{z = x + iy : \; x, y \in \mathbb{R}\}. \tag{12}$$

As we shall see hereafter, $\mathbb{C}$ is an algebraic field, i.e., is possible to define in $\mathbb{C}$ addition and multiplication operations satisfying the same field axioms we have seen in Lecture 1 for $\mathbb{R}$ (field axioms for real numbers).

Addition and multiplication. Consider the following complex numbers

$$z_1 = x_1 + iy_1 \qquad z_2 = x_2 + iy_2. \tag{13}$$

It is natural to define addition and multiplication in $\mathbb{C}$ by using the addition and multiplication operations we defined in $\mathbb{R}$ (see Lecture 1). Specifically, we define

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2) \qquad \text{(addition operation)} \tag{14}$$

Note that $z_1 + z_2$ is still of the form $x + iy$ (with $x = (x_1 + x_2)$ and $y = (y_1 + y_2)$). Therefore $\mathbb{C}$ is closed[1] under the addition operation "+" defined in (14). Similarly,

$$\begin{aligned} z_1 z_2 =& (x_1 + iy_1)(x_2 + iy_2) \\ =& (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1) \qquad \text{(multiplication operation)} \end{aligned} \tag{15}$$

Again, $z_1 z_2$ is a complex number, i.e., a number of the form $x + iy$ (with $x = (x_1 x_2 - y_1 y_2)$ and $y = (x_1 y_2 + x_2 y_1)$). This means that $\mathbb{C}$ is closed under the multiplication operation defined in (15)

It is easy to show that the set of complex numbers $\mathbb{C}$, with the addition and multiplication operations defined in (14) and (15) is a field. In other words, for all $z_1, z_2, z_3 \in \mathbb{C}$ we have that:

1. Addition and multiplication are *commutative*

$$z_1 + z_2 = z_2 + z_1 \qquad z_1 z_2 = z_2 z_1 \tag{16}$$

2. Addition and multiplication in are *associative*

$$(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3) \qquad (z_1 z_2) z_3 = z_1 (z_2 z_3) \tag{17}$$

3. The *distributive property* of multiplication relative to addition holds

$$z_1 (z_2 + z_3) = z_1 z_2 + z_1 z_3 \tag{18}$$

4. There exists *neutral elements* for both addition and multiplication

$$\begin{aligned} &z_1 + z_2 = z_1 \quad \Rightarrow \quad z_2 = 0 + i0 \qquad \text{(additive neutral element)} \\ &z_1 z_2 = z_1 \qquad \Rightarrow \quad z_2 = 1 + i0 \qquad \text{(multiplicative neutral element)} \end{aligned} \tag{19}$$

[1]We say that $\mathbb{C}$ is closed under the addition operation + defined in (14) if for all $z_1, z_2 \in \mathbb{C}$ we have that $(z_1 + z_2) \in \mathbb{C}$.

Based on the definition of additive and multiplicative neutrals it is straightforward to define the *opposite* and the *inverse* of a complex number. To this end, let $z = x + y$

$$
\begin{aligned}
&z + z_1 = 0 + 0i \quad \Rightarrow \quad z_1 = -x - iy \qquad \text{(opposite of } z\text{)}\\
&zz_2 = 1 + 0i \qquad \Rightarrow \quad z_2 = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2} \qquad \text{(inverse of } z\text{)}
\end{aligned} \tag{20}
$$

Let us denote z_1 as $-z$ and z_2 as $1/z$. Note that equations (20) allow us to define *subtraction* and *division* between two complex numbers in terms of addition and multiplication. In fact, subtraction of z_2 from z_1 is the same as adding the opposite of z_2 to z_1. Similarly, z_1/z_2 is the same as multiplying z_1 by the inverse of z_2.

Remark: We have now set up all the machinery to perform any type of algebraic calculation between complex numbers, including addition, subtraction, multiplication and division.

Remark: From what has been said, it is clear that $\mathbb{C}$ includes $\mathbb{R}$ as a subset, i.e., $\mathbb{R} \subset \mathbb{C}$. This can be seen by noting that real numbers are simply complex numbers with zero imaginary part. Moreover, the addition and multiplication operations we defined in $\mathbb{C}$, i.e., equations (14)-(15), reduce to addition and multiplication between real numbers if we set to zero the imaginary parts. Hence $\mathbb{R} \subset \mathbb{C}$.

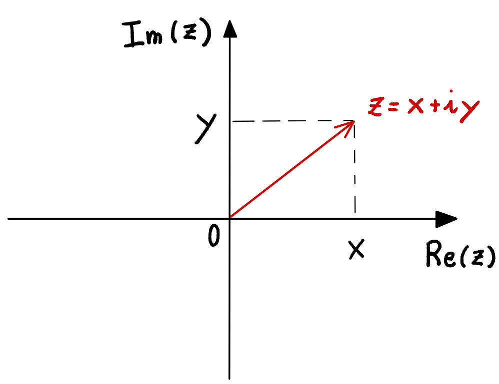
Remark: $\mathbb{C}$ is not an ordered field. In other words, it does not make sense to write inequalities between complex numbers.

Graphical representation of complex numbers. There is a one-to-one correspondence between the complex number

$$z = x + iy$$

and the pair of real numbers $x, y \in \mathbb{R}$. This means that z identifies uniquely x and y, and conversely the pair (x, y) identifies uniquely the complex number z. This suggests that we could represent (x, y) as a point (or a vector) in the Cartesian plane.

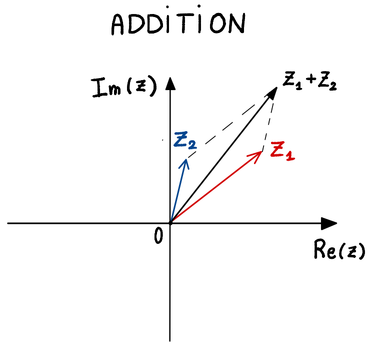
When the Cartesian plane is used to represent complex numbers, it is usually called *complex plane*. In this setting, the x-axis is called *real axis* while the y-axis is called *imaginary axis*.

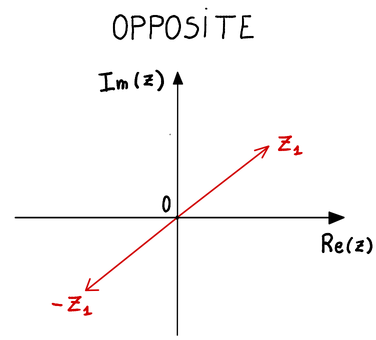


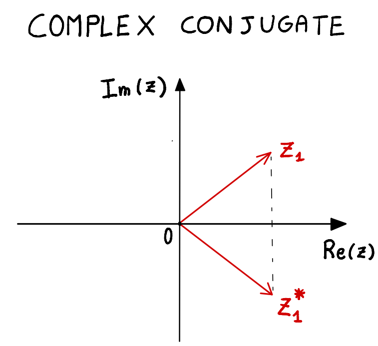
Recall that for the complex number $z = x + iy$ we defined

$$\text{Re}(z) = x \quad \text{(real part of } z\text{)} \qquad \text{Im}(z) = y \quad \text{(imaginary part of } z\text{)}. \tag{21}$$

The complex plane allows us to easily visualize addition between complex numbers (parallelogram rule), and other operations such as the opposite of a complex number, and the complex conjugate (reflections with respect to the real axis).







Complex conjugate. Let $z = x + iy$ be a complex number. The complex conjugate of z is the complex number

$$z^* = x - iy \qquad \text{(complex conjugate).} \tag{22}$$

Note that z^* has the same real part of z, but opposite imaginary part. With this notation, we have the following characterization of the complex conjugate.

Theorem 1 (Properties of the complex conjugate)**.** Let $z, w \in \mathbb{C}$ be two arbitrary complex numbers. Then

1. $(z^*)^* = z$
2. $(z + w)^* = z^* + w^*$
3. $(zw)^* = z^* w^*$
4. $z + z^* = 2\,\mathrm{Re}(z)$
5. $z - z^* = 2i\,\mathrm{Im}(z)$
6. $zz^* = \mathrm{Re}(z)^2 + \mathrm{Im}(z)^2$
7. $z = z^* \Leftrightarrow z \in \mathbb{R}$

Proof. Let us prove property 3, property 4, and property 6. The proof of the other properties is left as exercise. Let $z = x + iy$ and $w = a + ib$ be two arbitrary complex numbers.

Property 3.

$$(zw)^* = ((x+iy)(a+ib))^* = (xa-yb+i(xb+ya))^* = xa-yb-i(xb+ya) = (x-iy)(a-ib) = z^*w^*.$$

Property 4.

$$z + z^* = (x + iy) + (x + iy)^* = 2x + iy - iy = 2\,\mathrm{Re}(z).$$

Property 6.

$$zz^* = (x + iy)(x + iy)^* = (x + iy)(x - iy) = x^2 + y^2 = \mathrm{Re}(z)^2 + \mathrm{Im}(z)^2.$$

□

Remark: By using the complex conjugate, it is easy to express the quotient between two complex numbers, e.g.,

$$\frac{3-2i}{-1+2i} \tag{23}$$

in an standard algebraic form. We know that such ratio is a complex number[2], and therefore it can written is the form $x+iy$. The question is what is x and what is y? There is a shortcut to answer this question. In practice, given a quotient between two complex numbers z_1 and z_2 (i.e., z_1/z_2) we can multiply the numerator and the denominator by z_2^* to obtain the algebraic form

$$\frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*}. \tag{24}$$

The denominator in (24) a real number (by property 6. in Theorem 1).

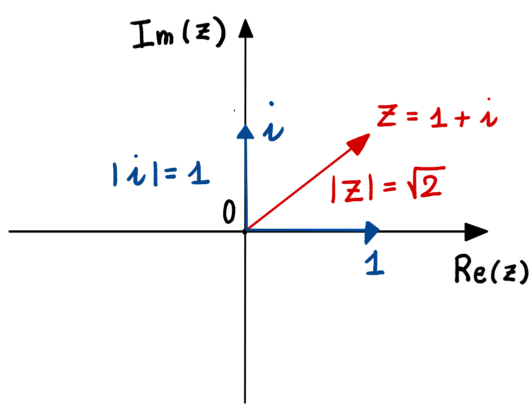
Example: Let $z_1 = 3-2i$ and $z_2 = -1+2i$. Compute the algebraic form of the complex number z_1/z_2. We have

$$\frac{3-2i}{-1+2i} = \frac{(3-2i)(-1-2i)}{\underbrace{(-1+2i)}_{z_1}\underbrace{(-1-2i)}_{z_1^*}} = \frac{(3-2i)(-1-2i)}{5} = -\frac{7}{5} - \frac{4}{5}i.$$

Modulus of a complex number. The modulus of a complex number $z = x+iy$ is a real number defined as

$$|z| = \sqrt{x^2+y^2} = \sqrt{zz^*} \qquad \text{(modulus of } z\text{)}$$

The modulus of z represents the *length* of the vector defined by the point (x,y) in the complex plane.



[2] In fact for any $z_1, z_2 \in \mathbb{C}$ we have that z_1/z_2 is the multiplication of z_1 by the inverse of z_2 (which is a complex number). Recall that multiplication between two complex numbers is a complex number. Therefore z_1/z_2 is a complex number that can be written in the algebraic form $x+iy$.

Clearly, the modulus of the imaginary number i is

$$|i| = |0 + 1i| = \sqrt{0^2 + 1^2} = 1. \tag{25}$$

Similarly, the modulus of the complex number $z = 1 + i$ is

$$|z| = |1 + 1i| = \sqrt{1^2 + 1^2} = \sqrt{2}. \tag{26}$$

The modulus of a complex number satisfies a certain number of properties which are summarized in the following Theorem.

Theorem 2 (Properties of the modulus)**.** Let $z, w \in \mathbb{C}$ be two arbitrary complex numbers. Then,

1. $|z| = 0 \Leftrightarrow z = 0$
2. $|z^*| = |z|$
3. $\{|\operatorname{Re}(z)|, |\operatorname{Im}(z)|\} \leq |z| \leq |\operatorname{Re}(z)| + |\operatorname{Im}(z)|$
4. $|zw| = |z||w|$
5. $|z + w| \leq |z| + |w|$ (triangle inequality)
6. $||z| - |w|| \leq |z + w|$ (reverse triangle inequality)

Proof. Let us prove property 4, and property 5. The proof of the other properties is left as exercise. Let z and w be two arbitrary complex numbers.

Property 4.

$$|zw|^2 = zwz^*w^* = zz^*ww^* = |z|^2|w|^2 \quad \Rightarrow \quad |zw| = |z||w|.$$

Property 5.

$$\begin{aligned} |z + w|^2 &= (z + w)(z^* + w^*) \\ &= zz^* + ww^* + wz^* + w^*z \\ &= |z|^2 + |w|^2 + 2\operatorname{Re}(wz^*) \\ &\leq |z|^2 + |w|^2 + 2|\operatorname{Re}(wz^*)|. \end{aligned} \tag{27}$$

At this point we notice that[3]

$$\operatorname{Re}(wz^*)^2 = |wz^*|^2 - \operatorname{Im}(wz^*)^2 \leq |wz^*|^2 = |w||z^*|^2 = |w|^2|z|^2, \tag{28}$$

i.e.,

$$|\operatorname{Re}(wz^*)| \leq |w||z|. \tag{29}$$

A substitution of this equation into (27) yields

$$|z + w|^2 \leq |z|^2 + |w|^2 + 2|w|^2|z|^2 = (|z| + |w|)^2. \tag{30}$$

By taking the square root of (30) we obtain Property 5.

[3]Equation (28) follows from property 6 in Theorem 1, and property 2 and 4 in Theorem 2.

□

Polar form of a complex number. We have seen in Theorem 2 (property 4) that given two arbitrary complex numbers the norm of their product is equal to the product of their norms, i.e.,

$$|zw| = |z||w| \qquad \forall z, w \in \mathbb{C}. \tag{31}$$

This implies implies that the product of two complex numbers with modulus one is still a complex number with modulus one. In other words, the set

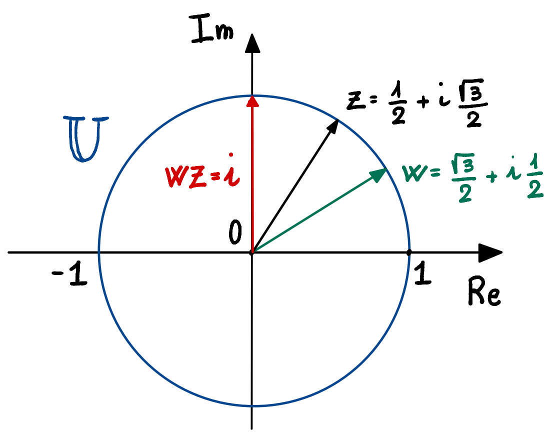
$$\mathbb{U} = \{z \in \mathbb{C} : \ |z| = 1\} \qquad \text{(unit circle in the complex plane)} \tag{32}$$

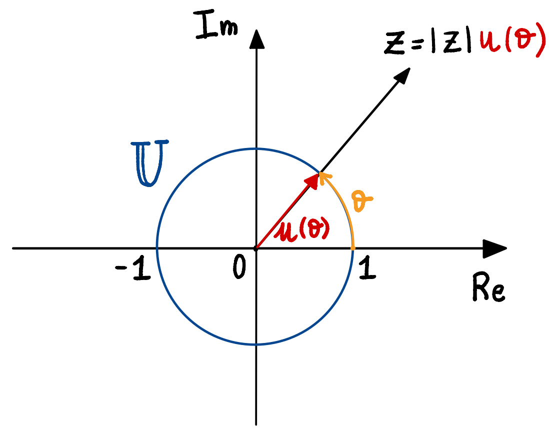
is closed under multiplication. For example, consider the following two complex numbers z and w

$$z = \frac{1}{2} + i\frac{\sqrt{3}}{2} \qquad w = \frac{\sqrt{3}}{2} + i\frac{1}{2} \tag{33}$$

both of which have modulus equal to one (verify it!). Clearly, we have

$$|zw| = |i| = 1. \tag{34}$$





The inverse of a complex number on the unit circle (32) coincides with the complex conjugate. In fact,

$$|z|^2 = 1 \quad \Rightarrow \quad zz^* = 1 \quad \Rightarrow \quad z^* = \frac{1}{z}. \tag{35}$$

Clearly, by using elements of the set $\mathbb{U}$ defined in (32) we can represent any complex number as

$$z = |z|u(\vartheta) \qquad u(\vartheta) \in \mathbb{U}. \tag{36}$$

Note that $u(\vartheta)$ depends only one parameter, i.e., the angle ϑ (arclength on the unit circle). Moreover, by using well-known results of trigonometry we can write the complex number $u(\vartheta)$ as

$$u(\vartheta) = \cos(\vartheta) + i\sin(\vartheta). \tag{37}$$

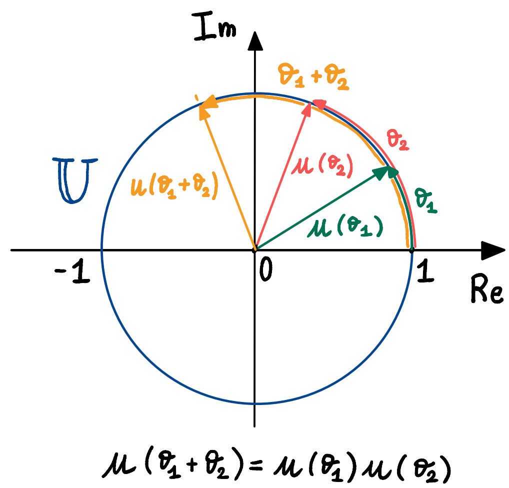
Complex exponential function. Consider two arbitrary complex numbers on the unit circle (32)

$$u(\vartheta_1) = \cos(\vartheta_1) + i\sin(\vartheta_1) \qquad u(\vartheta_2) = \cos(\vartheta_2) + i\sin(\vartheta_2) \tag{38}$$

and take their product

$$\begin{aligned} u(\vartheta_1)u(\vartheta_2) &= \cos(\vartheta_1)\cos(\vartheta_2) - \sin(\vartheta_1)\sin(\vartheta_2) + i\left[\sin(\vartheta_1)\cos(\vartheta_2) + i\cos(\vartheta_1)\sin(\vartheta_2)\right] \\ &= \cos(\vartheta_1+\vartheta_2) + i\sin(\vartheta_1+\vartheta_2) \\ &= u(\vartheta_1+\vartheta_2) \end{aligned} \tag{39}$$

Remark: This means that the function $u(\vartheta)$ defined in (37) transforms sums into products, i.e., $u(\vartheta_1+\vartheta_2) = u(\vartheta_1)u(\vartheta_2)$.



The similarity between the function $u(\vartheta)$ and the real exponential function e^x ($x \in \mathbb{R}$) is quite remarkable. In fact, we have

$$e^{x_1+x_2} = e^{x_1}e^{x_2}, \qquad \text{for all} \quad x_1, x_2 \in \mathbb{R}. \tag{40}$$

This suggests the following definition of *complex exponential function*

$$e^{i\vartheta} = \cos(\vartheta) + i\sin(\vartheta). \tag{41}$$

Remark: There are several other reasons supporting the definition of complex exponential function (41). For instance, consider the Taylor series of the real exponential function

$$e^x = \sum_{k=1}^{\infty} \frac{x^k}{k!}. \tag{42}$$

It is known that such series converges for all $x \in \mathbb{R}$. By substituting x with $i\vartheta$ in (42) we obtain

$$\begin{aligned} e^{i\vartheta} &= \sum_{k=1}^{\infty} \frac{i^k\vartheta^k}{k!} \\ &= \left(1 - \frac{\vartheta^2}{2} + \frac{\vartheta^4}{24} - \cdots\right) + i\left(\vartheta - \frac{\vartheta^3}{6} + \frac{\vartheta^5}{120} - \cdots\right) \\ &= \cos(\vartheta) + i\sin(\vartheta). \end{aligned} \tag{43}$$

In fact, recall that the Taylor series of $\cos(\vartheta)$ and $\sin(\vartheta)$ are

$$\cos(\vartheta) = 1 - \frac{\vartheta^2}{2!} + \frac{\vartheta^4}{4!} + \cdots \qquad \sin(\vartheta) = \vartheta - \frac{\vartheta^3}{3!} + \frac{\vartheta^5}{5!} - \cdots . \tag{44}$$

Another reason why it makes sense to define the complex exponential as in (41) is that

$$\frac{de^{i\vartheta}}{d\vartheta} = ie^{i\vartheta}. \tag{45}$$

This can be verified by calculating the derivatives of the right hand side of (41) with respect to ϑ.

In summary, the complex exponential function (41) has the same properties of the real exponential function, e.g., Taylor expansion, derivatives, and the product rule

$$e^{i(\vartheta_1+\vartheta_2)} = e^{i\vartheta_1}e^{i\vartheta_2} \tag{46}$$

Euler's formulas. By using equation (41) is straightforward to express $\sin(\vartheta)$ and $\cos(\vartheta)$ in terms of complex exponential functions. To this end, we first evaluate (41) at $-\vartheta$

$$e^{-i\vartheta} = \cos(\vartheta) - i\sin(\vartheta). \tag{47}$$

Then we add and subtract (47) to (41) to obtain

$$\cos(\vartheta) = \frac{e^{i\vartheta} + e^{-i\vartheta}}{2} \qquad \text{and} \qquad \sin(\vartheta) = \frac{e^{i\vartheta} - e^{-i\vartheta}}{2i}. \tag{48}$$

Argument of a complex number. We have seen that an arbitrary complex number $z \in \mathbb{C}$ can be written in three equivalent forms:

1. $z = x + iy$ (algebraic form)
2. $z = |z|e^{i\vartheta}$ (polar form)
3. $z = |z|\left(\cos(\vartheta) + i\sin(\vartheta)\right)$ (trigonometric form)

The real number ϑ is called *argument* of the complex number z, and it represents the arclength (in radiants) identified by the point $z/|z|$ on the unit circle $\mathbb{U}$ (see Eq. (32)). To calculate the argument of z, consider the following relations between algebraic form of z and the trigonometric form

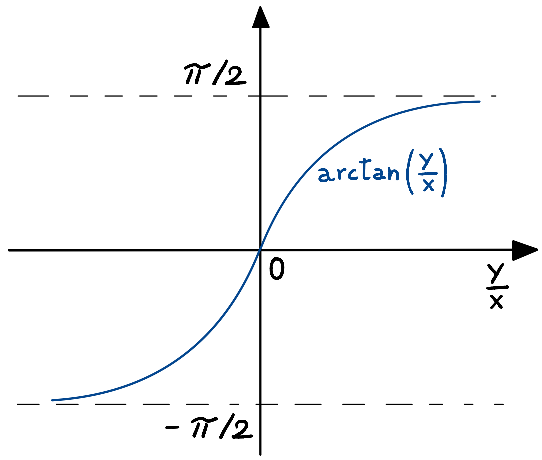
$$x = |z|\cos(\vartheta) \qquad y = |z|\sin(\vartheta) \tag{49}$$

The ratio y/x coincides with the tangent of ϑ

$$\tan(\vartheta) = \frac{y}{x} \tag{50}$$

How do we extract the angle ϑ from the previous equation? One possibility is to use the inverse of the tangent function, i.e., $\arctan(\cdot)$, and write

$$\vartheta = \arctan\left(\frac{y}{x}\right) \tag{51}$$

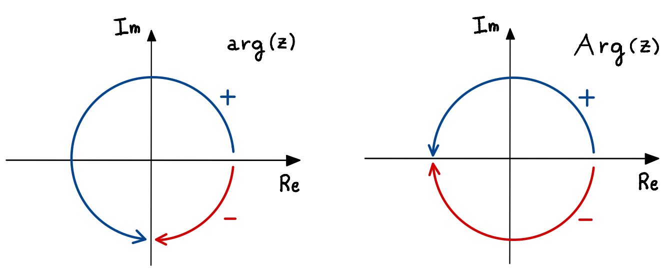


The problem with this simple approach is that function $\arctan(x)$ is defined only in the open interval $]-\pi/2, \pi/2[$. Hence, the expression (51) can be used only to compute the argument of complex number with strictly positive real part[4] (first and fourth quadrants of the complex plane).

To compute the argument of arbitrary complex number $x = x + iy$ we need to shift $\arctan(y/x)$ by π if the real part x is negative

$$\arg(z) = \begin{cases} \arctan\left(\dfrac{y}{x}\right) & x > 0 \\ \dfrac{\pi}{2}\mathrm{sign}(y) & x = 0 \\ \arctan\left(\dfrac{y}{x}\right) + \pi & x < 0 \end{cases} \tag{52}$$

With this definition $\vartheta = \arg(z)$ is unique for all $z \in \mathbb{C}$ and it ranges in $[-\pi/2, 3\pi/2[$.



Alternatively, we can define the argument as (note that here we use capitalized $\mathrm{Arg}(\cdot)$ to distinguish it from (52))

$$\mathrm{Arg}(z) = \begin{cases} \arctan\left(\dfrac{y}{x}\right) & x > 0 \\ \dfrac{\pi}{2}\mathrm{sign}(y) & x = 0 \\ \arctan\left(\dfrac{y}{x}\right) + \pi & x < 0, y \geq 0 \\ \arctan\left(\dfrac{y}{x}\right) - \pi & x < 0, y < 0 \end{cases} \tag{53}$$

With this definition $\vartheta = \mathrm{Arg}(z)$ is unique for all $z \in \mathbb{C}$ and it ranges in $[-\pi, \pi[$.

[4]Complex numbers with argument $\vartheta \in]-\pi/2, \pi/2[$ are either in first quadrant ($\vartheta \in [0, \pi/2[$) or in the fourth quadrant ($\vartheta \in]-\pi/2, 0]$) of the complex plane.

Remark: If we shift the argument of a complex number by $2k\pi$ ($k \in \mathbb{Z}$, the number is not going to change. Hence, the following complex numbers

$$z = 3e^{i\pi/3} \qquad z = 3e^{13i\pi/3} \qquad z = 3e^{-5i\pi/3} \tag{54}$$

are actually the same complex number. This is due to the 2π-periodicity of the circular functions defining the complex exponential (42).

Integer powers of a complex number (De Moivre's formula). Consider a complex number z expressed in a polar form

$$z = |z|e^{i\vartheta}, \tag{55}$$

where $|z|$ is the modulus of z and ϑ denotes its argument. By multiplying z recursively by itself we obtain

$$z^2 = |z|^2 e^{2i\vartheta}, \qquad z^3 = |z|^3 e^{3i\vartheta}, \quad \dots. \tag{56}$$

Similarly,

$$z^{-1} = \frac{1}{z} = \frac{z^*}{zz^*} = \frac{|z|}{|z|^2}e^{-i\vartheta} = \frac{1}{|z|}e^{-i\vartheta} = |z|^{-1}e^{-i\vartheta} \tag{57}$$

By multiplying $1/z$ recursively by itself we obtain

$$z^{-2} = |z|^{-2}e^{-2i\vartheta}, \qquad z^{-3} = |z|^{-3}e^{-3i\vartheta}, \quad \dots. \tag{58}$$

Therefore we proved the following Theorem.

Theorem 3 (De Moivre's formula)**.** Let z be any complex number with modulus $|z|$ and argument ϑ. Then

$$z^n = |z|^n e^{in\vartheta} \qquad \forall n \in \mathbb{Z}. \tag{59}$$

Remark: The powers of a complex number complex are points on a spiral in the complex plane. In fact, that the parametric form of a spiral in the Cartesian plane is

$$x(t) = a^t \cos(bt) \qquad y(t) = a^t \sin(bt), \tag{60}$$

where t is the spiral parameter, and a, b are fixed real numbers. These equations coincide with the real and imaginary parts of the powers of z. In fact,

$$\mathrm{Re}(z^n) = |z|^n \cos(n\vartheta) \qquad \mathrm{Im}(z^n) = |z|^n \sin(n\vartheta) \qquad n \in \mathbb{Z}. \tag{61}$$

Lecture 3: Roots of complex polynomials

To characterize the roots of complex polynomials we first study the roots of a complex number.

Roots of a complex number. Let $z, w \in \mathbb{C}$ be two complex numbers, and $n \in \mathbb{N}$ a natural number. We have seen how to compute the n-th power of z (or w) using De Moivre's formula (see Lecture 2), i.e.,

$$z = |z|e^{i\vartheta} \quad \Rightarrow \quad z^n = |z|^n e^{in\vartheta}. \tag{1}$$

Now we consider the inverse operation, i.e., how to compute the n-th root of a complex number. We say that z is the n-th root[1] of w if

$$z^n = w. \tag{2}$$

This is the simplest polynomial equation involving complex numbers: here $w \in \mathbb{C}$ is *given* while $z \in \mathbb{C}$ is *to be determined*. We shall see hereafter that the polynomial equation (2) has exactly n solutions in $\mathbb{C}$. To compute such solutions it is convenient to first write both z and w in a polar form as

$$w = |w|e^{it} \qquad \text{and} \qquad z = |z|e^{i\vartheta}. \tag{3}$$

Taking the n-th power of z as in (1) and substituting it into (2) yields

$$|z|^n e^{in\theta} = |w|e^{it} \tag{4}$$

This equation is equivalent to the following system of equations

$$|z|^n = |w|, \qquad e^{in\vartheta} = e^{it}. \tag{5}$$

The first one admits the unique solution[2]

$$|z| = \sqrt[n]{|w|}. \tag{6}$$

The second equation $e^{in\vartheta} = e^{it}$ is an equality between two vectors on the unit circle in the complex plane, and it has exactly n distinct solutions $\{\vartheta_0, \dots, \vartheta_{n-1}\}$. To compute such solutions we simply rewrite the equality $e^{in\vartheta} = e^{it}$ in a trigonometric form as

$$\cos(n\vartheta) = \cos(t), \qquad \sin(n\vartheta) = \sin(t). \tag{7}$$

This is a system of two *nonlinear* equations in the unknown variable ϑ.

How do we solve the nonlinear system (7) for ϑ? How many distinct solutions does it have?

- Clearly
$$n\vartheta = t. \tag{8}$$
is a solution to the system (7) since it satisfies both equations. In fact, substituting $n\vartheta = t$ into (7) yields two identities: $\cos(t) = \cos(t)$ and $\sin(t) = \sin(t)$. However, (8) is not the only solution. In fact, by using the periodicity of the cosine and sine functions we have that
$$\cos(t + 2k\pi) = \cos(t) \qquad \sin(t + 2k\pi) = \sin(t) \quad \text{for all } k \in \mathbb{Z} \tag{9}$$

[1] In equation (2) w is the n-th power of z while z is the n-th root of w.

[2] Recall that $|z| \geq 0$ and $|w| \geq 0$. Therefore there exists a unique solution to $|z|^n = |w|$.

This means that

$$n\vartheta = t + 2k\pi \qquad k \in \mathbb{Z} \tag{10}$$

are solutions. These solutions however are not all distinct. In fact, we have seen that $\vartheta \pm 2\pi$ identifies the same complex number on the unit circle. Therefore the only distinct solutions of (7) are

$$n\vartheta_k = t + 2k\pi \qquad k = 0, \ldots, n-1. \tag{11}$$

Therefore, the complex n-the roots of a number $w \in \mathbb{C}$ can be written explicitly as follows:

$$z^n = w \qquad \Leftrightarrow \qquad \boxed{z_k = \sqrt[n]{|w|}e^{i\vartheta_k} \qquad \vartheta_k = \frac{t+2k\pi}{n} \qquad k = 0, \ldots, n-1} \tag{12}$$

where $|w|$ and t are, respectively, the modulus and the argument of the complex number w. Note that all complex roots of a number w lie on a circle with radius $\sqrt[n]{|w|}$ in the complex plane.

Example: Compute the complex 4-th roots the real number $w = -1$. Such roots are defined by the (complex) solutions to the equation

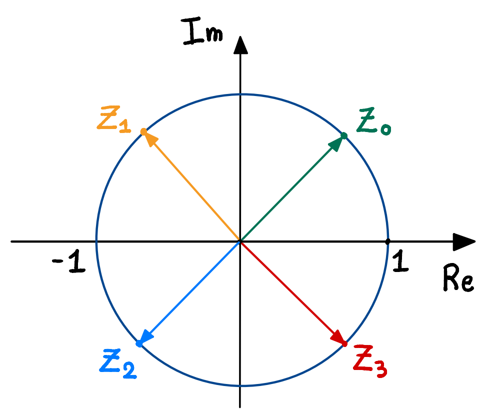
$$z^4 = -1. \tag{13}$$

By applying formula (12) we immediately get

$$z_k = e^{i(\pi+2k\pi)/4} \qquad k = 0, 1, 2, 3.$$

In fact, the modulus of w is equal to 1, and therefore $\sqrt[4]{|w|} = 1$. The four complex 4-th roots of -1 can be written in an algebraic form as follows

$$z_0 = \frac{1+i}{\sqrt{2}}, \qquad z_1 = \frac{-1+i}{\sqrt{2}}, \qquad z_2 = \frac{-1-i}{\sqrt{2}}, \qquad z_3 = \frac{1-i}{\sqrt{2}}.$$



It can be verified that each z_k in (14) satisfies indeed $z_k^4 = -1$. For example, using the algebraic form we have

$$\begin{aligned} z_0^4 =& \frac{(1+i)^4}{4} = \frac{(1+i)^2(1+i)^2}{4} = \frac{(2i)^2}{4} = -1, \\ z_1^4 =& \frac{(-1+i)^4}{4} = \frac{(-1+i)^2(-1+i)^2}{4} = \frac{(-2i)^2}{4} = -1. \end{aligned} \tag{14}$$

Example: Compute the complex cubic roots of the real number $w = 2$. The cubic roots are complex solutions of the polynomial equation

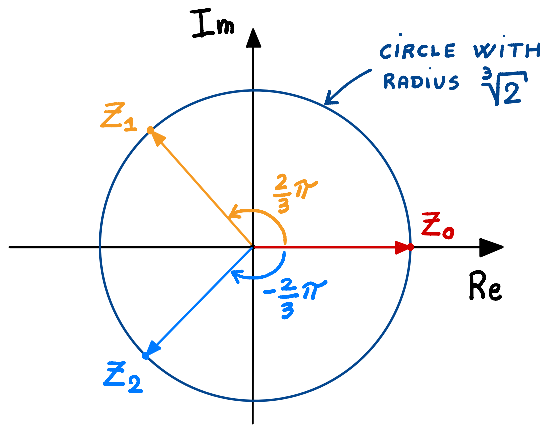
$$z^3 = 2 \tag{15}$$

By using (12) we immediately obtain

$$z_k = \sqrt[3]{2} e^{i2k\pi/3} \qquad k = 0, 1, 2$$

i.e.,

$$z_0 = \sqrt[3]{2}, \qquad z_1 = \frac{\sqrt[3]{2}}{2}\left(-1 + i\sqrt{3}\right), \qquad z_2 = \frac{\sqrt[3]{2}}{2}\left(-1 - i\sqrt{3}\right).$$



We remark that if we solve $z^3 = 1$ in $\mathbb{R}$ instead of $\mathbb{C}$ then we obtain a unique solution, i.e., $z = 1$. On the other hand in $\mathbb{C}$ we have three solutions: one real, and two *complex conjugates*.

Remark: Formula (12) suggests that once the first n-th root z_0 is found, then all others can be obtained by simply dividing the circle with radius $|z| = \sqrt[n]{|w|}$ into n evenly-spaced parts!

Roots of quadratic polynomial equations in $\mathbb{C}$. Consider the following quadratic polynomial[3]

$$az^2 + bz + c = 0, \tag{17}$$

where a, b, and c can be complex numbers. Divide (17) by a

$$z^2 + \frac{b}{a}z + \frac{c}{a} = 0$$

and complete the square

$$\left(z + \frac{b}{2a}\right)^2 + \frac{c}{a} - \frac{b^2}{4a^2} = 0. \tag{18}$$

Upon definition of

$$\delta = z + \frac{b}{2a} \tag{19}$$

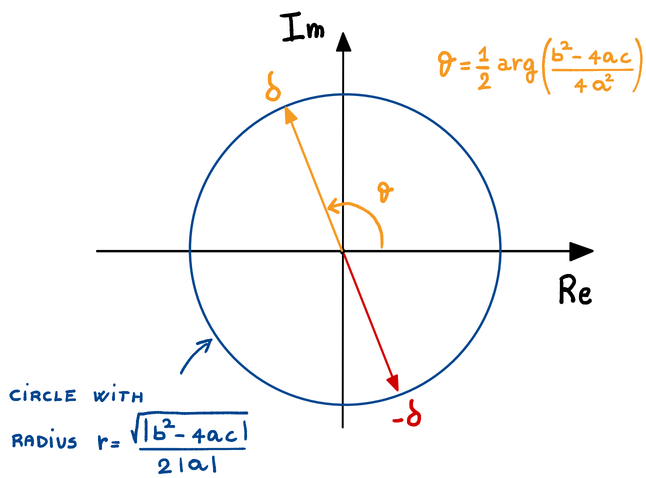
[3]An example of a quadratic polynomial with complex coefficients is

$$(1 + i)z^2 + 5z - i = 0. \tag{16}$$

we can write (18) as

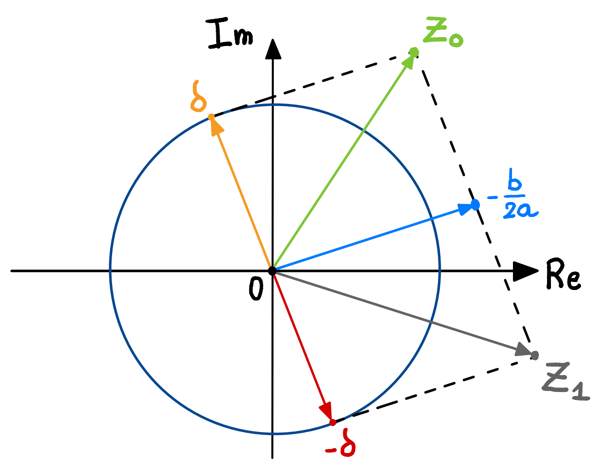
$$\delta^2 = \frac{b^2 - 4ac}{4a^2}.$$

This equation can be solved by taking the square root of the complex number $b^2 - 4ac/4a^2$. As is well-known, this yields two complex numbers δ and $-\delta$ opposite to each other and sitting on a circle with radius $|b^2 - 4ac|^{1/2}/(2|a|)$.



By using (19) we see that the solution of the quadratic polynomial equation (17) is then

$$z_0 = \delta - \frac{b}{2a}, \qquad z_1 = -\delta - \frac{b}{2a}. \tag{20}$$



Example: Consider the polynomial equation

$$z^2 + z + 1 = 0. \tag{21}$$

This equation has real coefficients but no solution in $\mathbb{R}$. By using the mathematical steps discussed above it can be shown that (21) can be written as

$$\delta^2 = -\frac{3}{4} \qquad \text{where} \qquad \delta = z + \frac{1}{2}. \tag{22}$$

Therefore the two complex solutions are

$$z_{0,1} = -\frac{1}{2} \pm i\frac{\sqrt{3}}{2}. \tag{23}$$

Quadratic polynomials with real coefficients have roots that are either real or complex conjugates. The roots can also be computed with the standard quadratic formula in this case.

Example: Consider the polynomial equation

$$z^2 + iz + 1 + i = 0. \tag{24}$$

By using the mathematical steps discussed above it can be shown that (24) can be written as

$$\delta = z + \frac{i}{2} \qquad \delta^2 = -\frac{5}{4} - i. \tag{25}$$

The polar form of the complex number $-5/4 - i$ is

$$-\frac{5}{4} - i = \frac{\sqrt{41}}{4} e^{i(\arctan(4/5)+\pi)}. \tag{26}$$

Therefore, the two solutions of $\delta^2 = -5/4 - i$ are

$$\delta_0 = \frac{\sqrt[4]{41}}{2} e^{i(\arctan(4/5)+\pi)/2}, \qquad \delta_1 = \frac{\sqrt[4]{41}}{2} e^{i(\arctan(4/5)+3\pi)/2}. \tag{27}$$

This implies that the roots of (24) are

$$z_0 = \frac{1}{2}\left(-i + \sqrt[4]{41} e^{i(\arctan(4/5)+\pi)/2}\right), \qquad z_1 = \frac{1}{2}\left(-i + \sqrt[4]{41} e^{i(\arctan(4/5)+3\pi)/2}\right). \tag{28}$$

Roots of complex polynomials. In the previous section we have seen how to compute the roots of quadratic polynomials with complex coefficients. A natural question is whether it is possible to generalize such computations to complex polynomials of degree $n > 2$. These polynomials can be written as

$$p(z) = a_n z^n + \cdots + a_1 z + a_0 \qquad a_i \in \mathbb{C}. \tag{29}$$

An even deeper question is whether polynomials of the form (29) actually have roots. This question was answered in 1799 by Gauss.

Theorem 1 (Fundamental theorem of algebra, Gauss 1799)**.** Every non-constant polynomial of the form (29) has at least one complex root.

By applying Gauss's theorem recursively it is straightforward to conclude that (29) has exactly n complex roots. This is summarized in the following Corollary.

Corollary 1. Every non-constant polynomial of the form (29) has exactly n complex roots.

Proof. Let $z_1 \in \mathbb{C}$ be a root of (29). We known that such a root exists because of Theorem 1. Let us first transform (29) to a monic polynomial (just divide by $a_n \neq 0$)

$$\hat{p}(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0, \qquad b_i = \frac{a_i}{a_n} \quad i = 0, \ldots, n-1. \tag{30}$$

Obviously we can factor out z_1 as

$$\hat{p}(z) = (z - z_1)\hat{p}_1(z) \tag{31}$$

where $\hat{p}_1(z)$ is a polynomial of degree $n-1$ obtained by dividing $\hat{p}(x)$ by $(z - z_1)$. The reminder of such polynomial division is zero because z_1 is a root of $\hat{p}(z)$. At this point we apply Theorem 1 again to $\hat{p}_1(x)$ to conclude that there exists another root z_2 and a polynomial $\hat{p}_2(x)$ of degree $n-2$ such that

$$\hat{p}(z) = (z - z_1)(z - z_2)\hat{p}_2(z). \tag{32}$$

Proceeding recursively we conclude that the polynomial (29) can be factorized as

$$\hat{p}(z) = (z - z_1)(z - z_2)\cdots(z - z_n). \tag{33}$$

This means that $\hat{p}(z)$ has exactly n roots in $\mathbb{C}$ (not necessarily distinct).

□

Regarding the computation of the roots, Ruffini (1799) and Abel (1824) proved that it is impossible to obtain closed form expressions for the roots of arbitrary polynomials of degree $n \geq 5$. This means that if we are interested in computing the roots of a given polynomial with degree $n \geq 5$ then we have to proceed numerically.

Remark: Effective algorithms to compute the roots of (29) are based on eigenvalue solvers. In fact, it can be shown that eigenvalues of the following *companion matrix*

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -b_0 \\ 1 & 0 & \cdots & 0 & -b_1 \\ 0 & 1 & \cdots & 0 & -b_2 \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \cdots & 1 & -b_{n-1} \end{bmatrix} \tag{34}$$

coincide with the roots of the polynomials (30) and (29).

The following theorem characterizes the roots of polynomials with *real* coefficients.

Theorem 2. Let $z_0 \in \mathbb{C}$ be a root of a polynomial with real coefficients. Then z_0^* (complex conjugate of z_0) is also a root.

Proof. Let

$$p(z) = \sum_{k=0}^{n} a_k z^k \tag{35}$$

be a polynomial with real coefficients $\{a_n, \ldots, a_0\}$. If z_0 is a root of $p(z)$ then

$$\sum_{k=0}^{n} a_k z_0^k = 0. \tag{36}$$

By taking the complex conjugate of (36) and recalling that[4]

$$(z_0^k)^* = (z_0^*)^k \qquad a_k^* = a_k \tag{37}$$

we obtain

$$\sum_{k=0}^{n} a_k (z_0^*)^k = 0. \tag{38}$$

Therefore if z_0 is a root of $p(z)$ then z_0^* is also a root of $p(z)$.

□

Theorem 2 essentially states the roots of a polynomial of degree n with real coefficients are either real or complex conjugates. This implies that the number of complex roots is always even for polynomials with real coefficients.

[4]Equation (37) follows from $(z_0^2)^* = (z_0 z_0)^* = z_0^* z_0^* = (z_0^*)^2$.

Lecture 4: Matrices and vectors

A matrix is a rectangular table with entries arranged in rows and columns. The entries can be numbers, functions, operators, matrices, symbols, etc. For example, the following matrix is a 2×3 matrix (2 rows and 3 columns) with real entries

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

Similarly,

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \tag{2}$$

is a matrix of trigonometric functions known as *rotation matrix*[1]. In general, an $m \times n$ matrix with entries in some set V has the form

$$A = \begin{array}{cc} \begin{array}{ccccc} & & j\text{-th column} & & \\ & & \downarrow & & \end{array} & \\ \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} & \leftarrow i\text{-th row} \end{array} \tag{3}$$

Denote the set of $m \times n$ matrices with entries in V as $M_{m\times n}(V)$. For example, we have:

- $M_{m\times n}(\mathbb{R})$: set of $m \times n$ matrices with real entries
- $M_{m\times n}(\mathbb{C})$: set of $m \times n$ matrices with complex entries
- $M_{m\times n}(C_0([0, 2\pi]))$: set of $m\times n$ matrices with entries in the space continuous functions defined on the interval $[0, 2\pi]$. An element of this set for $n = m = 2$ is the matrix defined in (2), i.e., $R(\theta) \in M_{2\times 2}(C_0([0, 2\pi]))$. Indeed, the entries of $R(\theta)$ are continuous functions in $[0, 2\pi]$.

Example (plotting functions and surfaces): Let us provide a simple example of how vectors and matrices can be used to plot one-dimensional and two-dimensional functions. To this end, consider

$$y = \sin(x) + 2 \qquad x \in [0, 2\pi]. \tag{4}$$

We are interested in plotting this function "point-by-point", i.e., map a set of points $\{x_1, \dots, x_n\}$ to $y_i = \sin(x_i) + 2$ $(i = 1, \dots, n)$ one by one. In particular, we choose the set of evenly-spaced points

$$x_{i+1} = \frac{2\pi}{n-1} i \qquad i = 0, \dots, n-1 \tag{5}$$

[1] The rotation matrix (2) defines rigid rotations of the Cartesian plane by an angle θ.

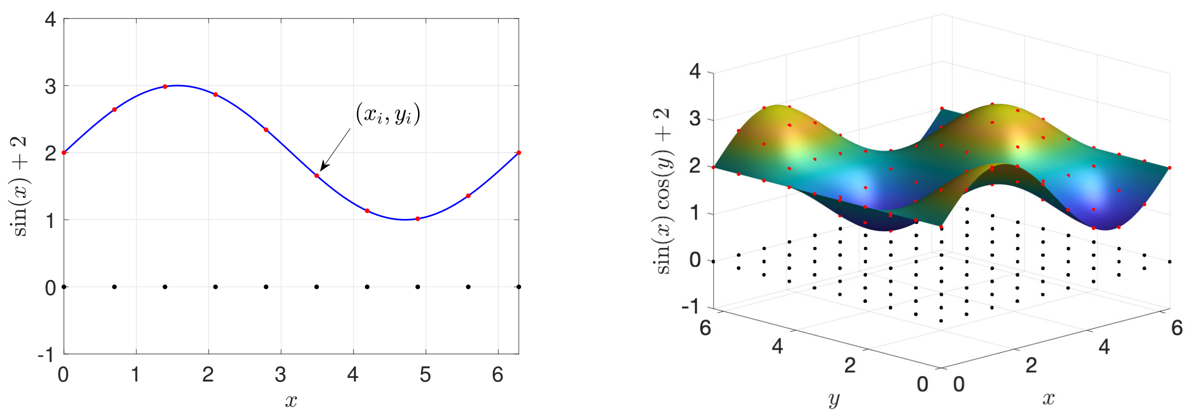
We collect all these points into a matrix with one row and n columns, i.e., a *row vector*

$$x = \begin{bmatrix} 0 & \dfrac{2\pi}{n-1} & \dfrac{4\pi}{n-1} & \cdots & 2\pi \end{bmatrix}. \tag{6}$$

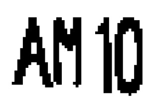
We can then map each entry of the vector x into the corresponding entry of another row vector y as

$$y_i = \sin(x_i) + 2 \qquad i = 1, \ldots, n. \tag{7}$$

Hereafter we show the results of this procedure for $n = 10$ in one and two dimensions.



Example (Matrices representing images): The following 50×50 black and white image "AM 10"



is represented by the matrix (1 means white, 0 means black)

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	0	0	0	1	1	1	1	1
1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1	1	0	0	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	0	0	0	0	0	1	1	1
1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1	1	0	0	0	1	1	1	1	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	1	1
1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	0	0	0	0	1	1	1	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	1	1
1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1	1	0	0	0	0	1	0	0	0	0	1	1
1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1
1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	1	1	1	1	0	0	0	1	1
1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	1	1	1	1	0	0	0	1	1
1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	1	1	1	1	0	0	0	1	1
1	1	1	1	1	0	0	0	1	1	0	0	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1
1	1	1	1	1	0	0	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0	0	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1
1	1	1	1	0	0	0	0	1	1	0	0	0	1	1	1	0	0	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1
1	1	1	1	0	0	0	0	1	1	0	0	0	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1
1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1
1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1
1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1
1	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1
1	1	1	0	0	0	0	1	1	1	1	0	0	0	1	1	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	0	0	0	1	1
1	1	1	0	0	0	1	1	1	1	1	0	0	0	1	1	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	0	0	0	1	1
1	1	0	0	0	0	1	1	1	1	1	0	0	0	1	1	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	0	0	0	1	1
1	1	0	0	0	0	1	1	1	1	1	0	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	0	0	0	1	1
1	1	0	0	0	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	1	1	1	1	0	0	0	1	1
1	1	0	0	0	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	0	1	1
1	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	1	0	0	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Addition between matrices. It makes sense to define addition between matrices with the same number of rows and and the same number of columns. To this end, let A and B are two $n \times m$ matrices

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} \tag{8}$$

We define

$$A + B = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \tag{9}$$

In this way $A+B$ is still an $m \times n$ matrix, i.e., the set of $n \times m$ matrices is closed under the addition operation defined in (9).

Example: Consider the following matrices:

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 4 & 1 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & 3 \end{bmatrix} \quad \Rightarrow \quad A + B = \begin{bmatrix} 1 & 0 & 4 \\ 5 & 3 & 5 \end{bmatrix}. \tag{10}$$

We also define the product between a matrix and number c (real or complex) as

$$cA = \begin{bmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & & \vdots \\ ca_{m1} & \cdots & ca_{mn} \end{bmatrix}. \tag{11}$$

Clearly cA is a $m \times n$ matrix.

Examples:

$$\begin{aligned} A &= \begin{bmatrix} 1 & 0 & 3 \\ 4 & 1 & 2 \end{bmatrix} & &\Rightarrow \; 3A = \begin{bmatrix} 3 & 0 & 9 \\ 12 & 3 & 6 \end{bmatrix}. \\ B &= \begin{bmatrix} i & 0 & 3+2i \\ 1+i & i & 2 \\ 1 & 0 & 6i \end{bmatrix} & &\Rightarrow \; iB = \begin{bmatrix} -1 & 0 & -2+3i \\ -1+i & -1 & 2i \\ i & 0 & -6 \end{bmatrix}. \end{aligned} \tag{12}$$

It is clear that the neutral element for the addition operation (9) is the *zero matrix*

$$0_{n \times m} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}. \tag{13}$$

In fact, for any $m \times n$ matrix A we have $A + 0_{m \times n} = A$. The *opposite* of the matrix A is the matrix[2]

$$-A = \begin{bmatrix} -a_{11} & \cdots & -a_{1n} \\ \vdots & & \vdots \\ -a_{m1} & \cdots & -a_{mn} \end{bmatrix}. \tag{14}$$

[2]The opposite of a matrix A is, by definition, the element B such that $A + B = 0_{m \times n}$.

Vector space of matrices. The addition and multiplication by a number operations defined in (9) and (11) satisfy the following properties

1. $A + B = B + A$ (matrix addition is commutative)
2. $(A + B) + C = A + (B + C)$ (matrix addition is associative)
3. $A + 0_{m\times n} = 0_{m\times n} + A = A$ (additive neutral, i.e., the zero matrix)
4. $A - A = 0_{m\times n}$ (opposite matrix $-A$)
5. $c(A + B) = cA + cB$ $c \in \mathbb{R}$ (or $\mathbb{C}$)
6. $(a + b)A = aA + bA$ $a, b \in \mathbb{R}$ (or $\mathbb{C}$)
7. $(ab)A = a(bA)$ $c \in \mathbb{R}$ (or $\mathbb{C}$)

In technical terms, we say that the space of $n \times m$ matrices over the field of complex numbers forms a *vector space*. More generally, any set in which we define an addition operation "+" and a multiplication by $c \in \mathbb{C}$ satisfying properties 1-7 listed above forms is a *vector space over* $\mathbb{C}$. The set of matrices with positive real entries is *not* a vector space since the opposite of a matrix with positive entries is not a matrix with positive entries.

The elements of a vector space are called *vectors*. Hence, a matrix is a vector in the vector space of matrices. A function $f(x) = \sin(x)^2$ is a vector in the vector space continuous functions from $\mathbb{R}$ into $\mathbb{R}$.

Matrix multiplication. Let us consider two matrices A and B and suppose that the number of columns of A (say p) coincides with the number of rows of B

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mp} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{p1} & \cdots & b_{pn} \end{bmatrix}. \tag{15}$$

The (standard) *matrix product* between A and B is defined as

$$(AB)_{ij} = a_{i1}b_{1j} + \cdots + a_{ip}b_{pj}, \qquad i = 1, \ldots, m, \quad i = 1, \ldots, n. \tag{16}$$

Here, $(AB)_{ij}$ denotes the (ij)-th entry of the matrix $m \times n$ matrix AB. Note that if the matrix A has size $m \times p$ and the matrix B has size $p \times n$ then the matrix AB defined in (16) has size $m \times n$.

The matrix product (16) corresponds to the so-called *row-column rule* in which the entries of the i-th row of the matrix A are multiplied by the entries j-the column of B and the results of all these multiplications are summed up to obtain the ij-entry of AB

$$\begin{array}{c} \\ \\ i \\ \\ \end{array} \underbrace{\begin{array}{c} j \\ \begin{bmatrix} & \vdots & & \\ \cdots & \square & & \\ & & & \end{bmatrix} \end{array}}_{AB} = i \underbrace{\begin{bmatrix} [\cdots & \cdots & \cdots & \cdots] \\ & & & \end{bmatrix}}_{A} \underbrace{\begin{array}{c} j \\ \begin{bmatrix} & \begin{bmatrix} \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} & \end{bmatrix} \end{array}}_{B}.$$

Example: Consider the two matrices

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 1 & 5 & -1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & -1 \\ 2 & -2 \\ 1 & -3 \end{bmatrix}. \tag{17}$$

The matrix product AB is well-defined and it is computed as follows:

$$AB = \begin{bmatrix} (1+6+2) & (-1-6-6) \\ (1+10-1) & (-1-10+3) \end{bmatrix} = \begin{bmatrix} 9 & -13 \\ 10 & -8 \end{bmatrix}. \tag{18}$$

Similarly, the matrix product BA in this case is well-defined[3] and it corresponds to the 3×3 matrix

$$BA = \begin{bmatrix} 0 & -2 & 3 \\ 0 & -4 & 6 \\ -2 & -12 & 5 \end{bmatrix}. \tag{19}$$

Remark: For square matrices A and B (i.e., matrices with size $m = p = n$) both products AB and BA are well-defined and they yield $n \times n$ matrices. However, the matrix product is (in general) not commutative, i.e., $AB \neq BA$. For example

$$A = \begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \tag{20}$$

do not commute. In fact, we have

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix}, \\ BA &= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ 7 & -3 \end{bmatrix}. \end{aligned} \tag{21}$$

The matrix $C = AB - BA$ is called *matrix commutator* of A and B and it often denoted by $C = [A, B]$. If $AB = BA$ then we say that A and B commute. If A and B commute then the commutator $[A, B]$ is the necessarily the zero matrix.

Remark: A very important example of matrix product is the so-called *matrix-vector* product, in which a $m \times n$ matrix A is multiplied by a column vector[4] with n entries

$$Ax = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix}. \tag{22}$$

Clearly, Ax is a column vector with entries

$$(Ax)_i = a_{i1}x_1 + \cdots + a_{in}x_n \qquad i = 1, \ldots m. \tag{23}$$

[3]More generally, if $A \in M_{n\times m}$ and $B \in M_{m\times n}$ then $AB \in M_{n\times n}$ and $BA \in M_{m\times m}$.

[4]A column vector with n entries is a $n \times 1$ matrix.

Remark: The neutral element for the multiplication operation (16) is called *identity matrix*

$$I_n = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \in M_{n\times n}(\mathbb{R}) \tag{24}$$

The identity matrix is a square matrix with ones along the main diagonal and zeros everywhere else. If A is a $m \times n$ matrix then

$$I_m A = A I_n = A \tag{25}$$

Theorem 1 (Properties of the matrix product)**.** Let A, B and C three matrices for which the following products and sums are well-defined. Then:

1. $A(BC) = (AB)C$ (matrix multiplication is associative),
2. $A(B+C) = AB + AC$ (left distributive property),
3. $(A+B)C = AC + BC$ (right distributive property),
4. $c(AB) = A(cB)$, $c \in \mathbb{C}$.

Proof. Properties 1 to 4 can be proved simply by using the definition (9), (11) and (16). Let us prove property 2. To this end, let $B, C \in M_{n\times m}$ and $A \in M_{p\times n}$ so that the matrix multiplication in property 2 is well-defined. The ij entry of the matrix $A(B+C)$ can be written as

$$(A(B+C))_{ij} = \sum_{p=1}^{n} a_{ip}(b_{ij} + c_{pj}) = \sum_{p=1}^{n} a_{ip} b_{ij} + \sum_{p=1}^{n} a_{ip} c_{ij} = (AB)_{ij} + (AC)_{ij}. \tag{26}$$

□

Remark: Consider an arbitrary square matrix A and a positive integer p. The p-th power of A is the matrix

$$A^p = \underbrace{AA\cdots A}_{p \text{ times}} \qquad \text{(matrix power)}. \tag{27}$$

For example, the square of the matrix A defined in equation (20) is

$$A^2 = \begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ -3 & 7 \end{bmatrix}. \tag{28}$$

Remark: It is possible to define other types of matrix products, e.g., the Kronecker product "$\otimes$" or the Hadamard product "$\circ$". These types of products are different from the matrix product (16), and they satisfy different properties. For example, the Hadamard product between the matrices

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

is defined as

$$A \circ B = \begin{bmatrix} a_{11}b_{11} & \cdots & a_{1n}b_{1n} \\ \vdots & & \vdots \\ a_{m1}b_{m1} & \cdots & a_{mn}b_{mn} \end{bmatrix} \qquad \text{(Hadamard product).} \tag{29}$$

and it is clearly commutative[5], i.e., $A \circ B = B \circ A$. On the other hand, given two matrices $A \in M_{n\times m}$ and $B \in M_{p\times q}$ their Kronecker product is defined as defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \qquad \text{(Kronecker product).} \tag{30}$$

Note that that $A \otimes B$ is a block matrix of size $np \times mq$. In fact, each entry of $A \otimes B$ is a matrix of size $p \times q$.

Transpose of a matrix. The transpose of the $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \tag{31}$$

is the matrix obtained by switching the row and column indices of A, i.e.,

$$A^T = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix}. \tag{32}$$

For example,

$$A = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -2 & 4 & 1 \end{bmatrix} \quad \Leftrightarrow \quad A^T = \begin{bmatrix} 1 & 0 \\ 2 & -2 \\ 1 & 4 \\ 3 & 1 \end{bmatrix}. \tag{33}$$

Theorem 2 (Properties of transpose matrix)**.** Let A and B two matrices for which the following operations are well-defined. Then:

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(AB)^T = B^T A^T$
4. $(cA)^T = cA^T \qquad c \in \mathbb{C}$

[5] Recall that the standard matrix product between two square matrices is (in general) not commutative, i.e, $AB \neq BA$.

Proof. Let us prove property 3. To this end, let $A \in M_{n\times m}$ and $B \in M_{m\times p}$. The ij entry of the matrix AB is (see equation (16))

$$(AB)_{ij} = \sum_{k=1}^{m} a_{ik}b_{kj} \tag{34}$$

To obtain the ij entry of $(AB)^T$ we simply need to switch i and j. This yields,

$$\left((AB)^T\right)_{ij} = \sum_{k=1}^{m} a_{jk}b_{ki} = \sum_{k=1}^{m} \left(B^T\right)_{ik}\left(A^T\right)_{kj} = \left(B^T A^T\right)_{ij} \tag{35}$$

All other properties can be proved in a similar way.

□

Remark: Let A, B, C and D four matrices such that the product $ABCD$ is well-defined. Then

$$(ABCD)^T = D^T C^T B^T A^T \tag{36}$$

In fact, by applying property 3 in Theorem 2 recursively we have

$$(ABCD)^T = (CD)^T (AB)^T = D^T C^T B^T A^T \tag{37}$$

Remark (Conjugate transpose): For matrices with complex entries we can also define the conjugate transpose as

$$A^H = \left(A^T\right)^* \tag{38}$$

The conjugate transpose of a matrix A has entries

$$a^H_{ij} = a^*_{ji}. \tag{39}$$

Symmetric and skew-symmetric matrices. Let $A \in M_{n\times n}$ be a square matrix[6].

If $A = A^T$ then we say that A is *symmetric*.

If $A = -A^T$ then we say that A is *skew-symmetric* (or *anti-symmetric*).

Examples of symmetric and skew-symmetric matrices are

$$A = \begin{bmatrix} -1 & 3 & 1 \\ 3 & -3 & 5 \\ 1 & 5 & 0 \end{bmatrix} \quad \text{(symmetric)}, \qquad B = \begin{bmatrix} 0 & 3 & 1 \\ -3 & 0 & -4 \\ -1 & 4 & 0 \end{bmatrix} \quad \text{(skew-symmetric)}. \tag{40}$$

By definition, the entries of a symmetric matrix A satisfy $a_{ij} = a_{ji}$. Similarly, the entries of a skew-symmetric matrix satisfy $a_{ij} = -a_{ji}$. Note that this implies that the diagonal entries of a skew symmetric matrix are necessarily zero

$$a_{ii} = -a_{ii} \quad \Rightarrow \quad a_{ii} = 0. \tag{41}$$

[6]The definition of symmetric and skew-symmetric matrices makes sense only for square matrices. In fact, the statements $A = A^T$ and $A = -A^T$ are legitimate only for square matrix. Otherwise we are saying that, e.g., a 3×2 matrix equals a 2×3 matrix.

Any square matrix $A \in M_{n\times n}$ can be decomposed into a sum of a symmetric matrix and a skew-symmetric matrix as follows

$$A = \underbrace{\frac{1}{2}(A + A^T)}_{\text{symmetric}} + \underbrace{\frac{1}{2}(A - A^T)}_{\text{skew-symmetric}} . \tag{42}$$

The following result holds for arbitrary rectangular matrices.

Theorem 3. Let $A \in M_{n\times m}$ be an arbitrary $n \times m$ matrix. Then AA^T is a $n \times n$ symmetric matrix and $A^T A$ is a $m \times m$ symmetric matrix.

The proof is left as exercise.

Remark: If A is a $n \times n$ square matrix then $A^T A$ and AA^T are both $n \times n$ symmetric matrices. In general, $A^T A \neq AA^T$. However, if A is symmetric then $AA^T = A^T A$ (show it!).

Matrix inverse. Let $A \in M_{n\times n}$ be a square matrix. We say that A is *invertible* if there exists a $n \times n$ matrix A^{-1} such that

$$AA^{-1} = I_n \qquad A^{-1}A = I_n. \tag{43}$$

where I_n is the identity matrix (24).

Theorem 4 (Uniqueness of the inverse matrix)**.** The matrix A^{-1} satisfying (43) is unique.

Proof. Suppose that there are two matrices B_1 and B_2 such that

$$AB_1 = I_n, \quad B_1A = I_n \qquad \text{and} \qquad AB_2 = I_n, \quad B_2A = I_n. \tag{44}$$

Then

$$B_2 = B_2I_n = B_2(AB_1) = (B_2A)B_1 = B_1, \tag{45}$$

i.e., $B_2 = B_1$. This means that for any matrix A, the inverse is unique (if it exists). □

Hence, if A is invertible[7] then there exists a unique matrix A^{-1} that commutes with A such that the matrix product between A and A^{-1} yields the identity matrix (24).

Theorem 5 (Properties of the inverse matrix)**.** Let A and B be two $n \times n$ invertible matrices. Then

1. $\left(A^{-1}\right)^{-1} = A$
2. $(AB)^{-1} = B^{-1}A^{-1}$
3. $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

Proof. Let us prove properties 1, 2 and 3.

1. Let C be the inverse of A^{-1}. Then

$$A^{-1}C = I_n \qquad CA^{-1} = I_n. \tag{46}$$

Theorem 3 says that there exists only one matrix that satisfies (46), and that matrix is A. Thus, the inverse of A^{-1} is A.

[7] We will derive conditions for the invertibility of a matrix A in subsequent lecture notes. As we will see, not every square matrix admits an inverse.

2. The following identities
$$I_n = AB(B^{-1}A^{-1}), \qquad I_n = (B^{-1}A^{-1})AB \tag{47}$$
imply that the inverse of the matrix AB is $B^{-1}A^{-1}$.

3. Consider
$$I_n = (AA^{-1})^T = (A^{-1})^T A^T \qquad I_n = (A^{-1}A)^T = A^T(A^{-1})^T. \tag{48}$$
Therefore the inverse of A^T, i.e. $(A^T)^{-1}$, is equal to $(A^{-1})^T$.

□

Orthogonal and unitary matrices. Let $A \in M_{n\times n}$ be a square matrix with real entries. We say that A is an *orthogonal matrix* [8] if
$$AA^T = A^T A = I_n. \tag{49}$$
Clearly, if A is an orthogonal matrix then (by using the definition of the inverse and its uniqueness)
$$A^T = A^{-1}. \tag{50}$$

Moreover, if A is an orthogonal matrix then the commutator
$$[A, A^T] = AA^T - A^T A = I_n - I_n = 0_{M_{n\times n}}. \tag{51}$$
If the matrix A has complex entries then we say that A is a *unitary matrix* if
$$AA^H = A^H A = I_n \tag{52}$$
where A^H is the conjugate transpose of A.

Linear systems of equations. Consider the linear system of equations (m equations in n unknowns)
$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{53}$$
Upon definition of
$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \tag{54}$$
we can write (53) in a matrix-vector product form as
$$Ax = b. \tag{55}$$
In the particular case where $m = n$ (number equations equals the number of unknowns) we have that if the matrix A is invertible then the system (53) admit the unique solution[9]
$$x = A^{-1}b \tag{56}$$
As we shall see in the next lecture, there is no need to compute the inverse matrix A^{-1} to solve the linear system (53).

[8] As we will see, the reason why we call the matrix A satisfying (49) an orthogonal matrix follows from the fact that the rows (or the columns) of such matrix are orthonormal relative to standard "dot product" in $\mathbb{R}^n$.

[9] By applying A^{-1} to both sides of (55) we obtain $\underbrace{A^{-1}A}_{I_n}x = A^{-1}b$, i.e., $x = A^{-1}b$.

Lecture 5: Linear equations

An equation in n variables is *linear* if it can be written in the form

$$a_n x_n + \ldots + a_1 x_1 = b. \tag{1}$$

The numbers $\{a_1, \ldots, a_n\}$ are the *coefficients* of the equation, while b is usually called *constant term.*

The variables x_j and the constant term b can be elements of rather general vector spaces. For example, x_j can be vectors in $\mathbb{R}^n$, $n \times m$ matrices with real entries, or real-valued continuous functions, while a_i are usually real or complex numbers[1].

Examples: Let us provide a few simple examples of linear equations in the space $\mathbb{R}^n$ for $n = 2$ and $n = 3$. The elements of $\mathbb{R}^n$ are n-tuples of real numbers of the form

$$x = (x_1, \ldots, x_n) \qquad x_i \in \mathbb{R}. \tag{2}$$

In a matrix setting, x can be represented as a row vector or as a column vector

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}. \tag{3}$$

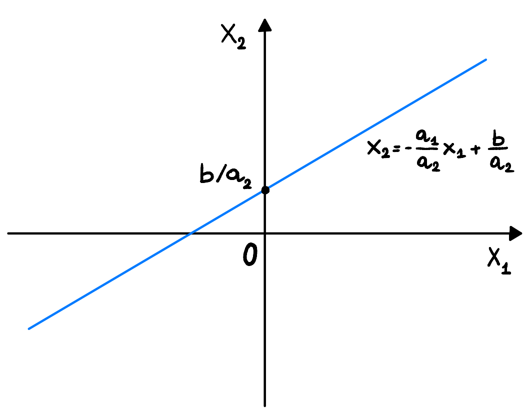
(a) The linear equation

$$a_1 x_1 + a_2 x_2 = b \qquad a_1, a_2, b \in \mathbb{R} \tag{4}$$

represents a *line* in $\mathbb{R}^2$. In fact, if $a_2 \neq 0$ then we can express x_2 in terms of x_1 as

$$x_2 = -\frac{a_1}{a_2} x_1 + \frac{b}{a_2}. \tag{5}$$

The graph x_2 versus x_1 is, e.g.,



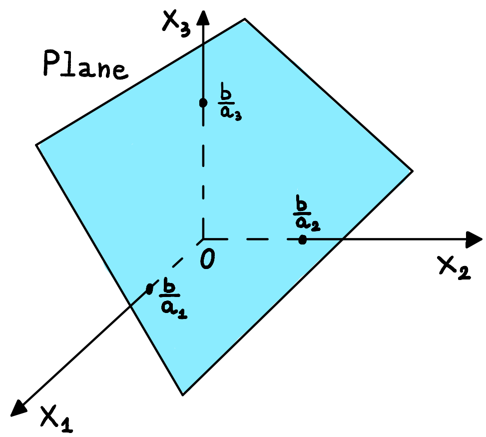
If $a_2 = 0$ and $a_1 \neq 0$ we obtain the vertical line $x_1 = b/a_1$. Lastly, if $a_1 = a_2 = 0$ then we necessarily have $b = 0$ and the linear equation reduces to $0 = 0$, which is uninformative.

[1] The vast majority of vector spaces are constructed over the field $\mathbb{R}$ or $\mathbb{C}$.

(b) The linear equation

$$a_3x_3 + a_2x_2 + a_1x_1 = b \qquad a_i, b \in \mathbb{R} \tag{6}$$

represents a *plane* in $\mathbb{R}^3$. Such a plane is a two-dimensional surface embedded in three dimensional space, which can be sketched as follows



This plane can be also expressed as a linear combination (linear equation) of two 3D vectors lying on the plane, plus a constant 3D vector.

(c) The following linear equation represents a so-called *hyper-plane* in $\mathbb{R}^n$ ($n \geq 4$).

$$a_nx_n + \cdots + a_1x_1 = b \qquad a_i, b \in \mathbb{R} \tag{7}$$

Systems of linear equations. A system of m linear equations of the form (1) can be written as

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_n \end{cases} \tag{8}$$

For example,

$$\begin{cases} 3x_1 + 2x_2 = 1 \\ x_1 - 5x_2 = 0 \end{cases} \qquad \text{2 equations in 2 variables}$$

$$\begin{cases} 5x_1 - x_3 = 3 \\ x_1 + 2x_2 - 8x_3 = 5 \end{cases} \qquad \text{2 equations in 3 variables}$$

A solution to the linear system (8) is a set n variables $(x_1, \ldots, x_n)$ satisfying all equations in (8). In general, linear systems can have

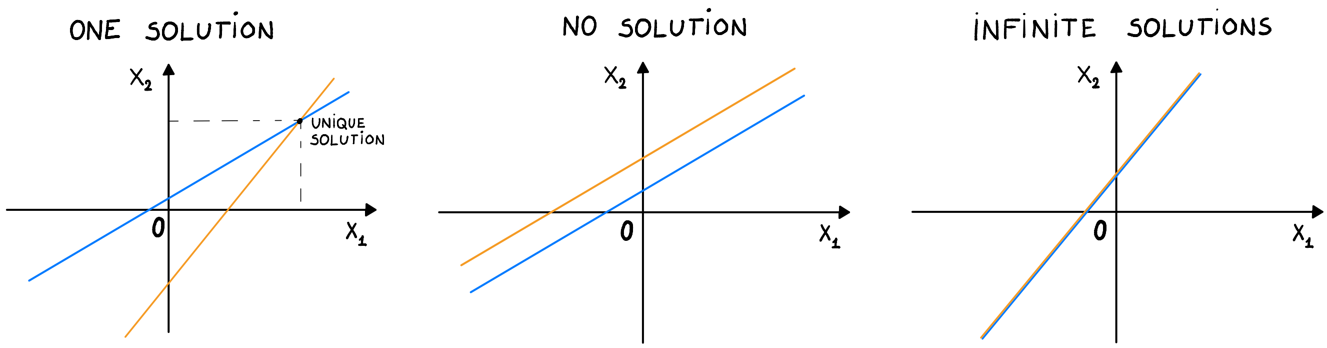
1. Exactly one solution
2. No solution
3. Infinite solutions

Geometric interpretation:

- We have seen that a linear equation in $\mathbb{R}^2$ defines a line in the Cartesian plane. Hence, the following system of two equations in $\mathbb{R}^2$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{9}$$

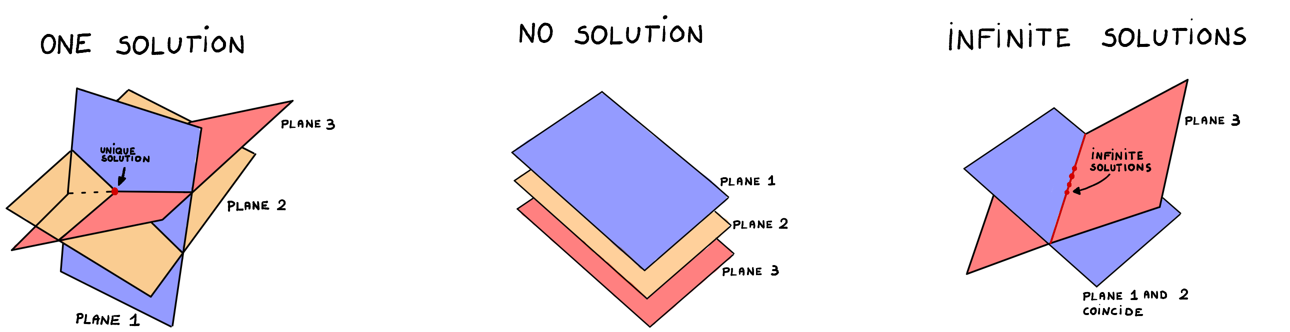
defines two lines. Such lines can intersect at one point (unique solution), can be parallel (no solutions) or they can be superimposed (infinite solutions).



- We have seen that a linear equation in $\mathbb{R}^3$ defines a plane in the three-dimensional space. Hence, the following three equations in $\mathbb{R}^3$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_2 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \tag{10}$$

define three planes. Such planes can intersect at one point (unique solution), can be parallel and distinct (no solution if just two planes are parallel), or they can intersect along one line (infinite solutions, one-dimensional set), or even be the same plane (infinite solutions, two-dimensional set).



Remark: A linear system of m equations in n variables can be written in a compact matrix-vector form as

$$Ax = b \tag{11}$$

where

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}. \tag{12}$$

Solving a linear system of equations. Let us begin with the the following simple example of a system of 2 linear equations in 2 unknowns

$$\begin{cases} x_1 + x_2 = 3 \\ x_1 - 2x_2 = 1 \end{cases} \tag{13}$$

Clearly, we can express x_1 in terms of x_2 by using the second equation,i.e.,

$$x_1 = 1 + 2x_2 \tag{14}$$

and then substitute this result into the first equation to obtain

$$1 + 2x_2 + x_2 = 3 \quad \Rightarrow \quad x_2 = \frac{2}{3} \tag{15}$$

$$x_1 = 1 + 2\left(\frac{2}{3}\right) \quad \Rightarrow \quad x_1 = \frac{7}{3} \tag{16}$$

Note that $x_1 = 2/3$ and $x_2 = 7/3$ satisfy (13). The method we just described, is not very efficient for linear systems in higher dimensions, e.g.,

$$\begin{cases} x_1 + x_2 + x_3 + x_4 - 3x_5 - = 1 \\ x_1 - x_2 + x_3 - x_4 - 12x_5 = 2 \\ 3x_1 - 3x_2 + x_3 + x_4 + x_5 = -2 \\ -x_1 + 2x_2 + x_3 + x_4 + -4x_5 = -2 \\ -4x_1 - x_2 + x_3 + x_4 + x_5 = -2 \end{cases}$$

A more effective method relies on transforming a linear system into an *equivalent* one, i.e., a systems with the same solutions, that is easier to solve. The key observation is the following:

> *The solution of a linear system does not change if we replace one equation with a linear combination of that equation and others in the system (we will see why!).*

Is this true? Let us verify the statement in the simplest possible setting, i.e., for the 2×2 linear system

$$\begin{cases} 2x_1 + x_2 = 1 \\ x_1 + x_2 = 0 \end{cases} . \tag{17}$$

This system can be written in a matrix-vector form as

$$\underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{b}. \tag{18}$$

The solution is clearly $x_1 = 1$ and $x_2 = -1$. Let us now replace the second equation in (17), i.e., $x_1 + x_2 = 0$, with the first equation multiplied by 2 plus the second. This yields

$$\begin{cases} 2x_1 + x_2 = 1 \\ 5x_1 + 3x_2 = 2 \end{cases} \tag{19}$$

which still has the unique solution $x_1 = 1$ and $x_2 = -1$. So the statement seems to be true.

If we replace the second equation in (17) by the second equation multiplied by 2 itself minus the first equation we can *eliminate* the variable x_1 to obtain

$$\begin{cases} 2x_1 + x_2 = 1 \\ x_2 = -1 \end{cases} \tag{20}$$

This system can be written in a matrix-vector form as

$$\underbrace{\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}}_{A_1} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{b_1} \tag{21}$$

The matrix A has an *upper-trianglar* triangular structure which allows us to solve the system by using *backward substitution*, i.e., solving the last equation first and then substituting the result back into into the previous equations.

Remark: Note that the operation we just described, i.e, "subtract the first equation from the second multiplied by 2" can be represented by a *lower-triangular* (invertible) matrix

$$T_1 = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \tag{22}$$

In fact, by applying T_1 to equation (18) we obtain equation (21), i.e.,

$$T_1 A x = T_1 b \quad \Rightarrow \quad A_1 x = b_1 \tag{23}$$

This can be verified by a direct calculation

$$T_1 A = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = A_1 \qquad T_1 b = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = b_1. \tag{24}$$

Gauss elimination method and row echelon forms of a matrix. The method we just described to transform a linear system in an "upper triangular" form is known as *Gauss elimination method*, and it can be applied to linear systems with an arbitrary number of linear equations and an arbitrary number of unknowns.

When performing Gaussian elimination is also convenient to interchange the rows of the augmented matrix so that the row with largest (in absolute value) entry acts as a pivot for the elimination step. This procedure is called *Gauss elimination method with pivoting by row.* In general, the following *elementary row operations* performed on the augmented matrix do not change the solution of the associated linear system of equations:

1. multiplication of one row by a non-zero number,
2. addition of one row to another, and
3. interchange two rows.

All these operations can be represented by invertible matrices. This implies that they do not change the solution of the system. In fact, if T is an invertible matrix then

$$Ax = b \quad \Leftrightarrow \quad TAx = Tb. \tag{25}$$

In orther words, $Ax = b$ and $TAx = Tb$ have the same solution. Note that it is possible to transform $TAx = Tb$ back into $Ax = b$ if and only if T is invertible[2]. On the other hand, if T is *not* invertible then

$$Ax = b \quad \Rightarrow \quad TAx = Tb, \qquad \text{but} \qquad TAx = Tb \not\Rightarrow Ax = b. \tag{26}$$

This means that the systems are not equivalent if T is not invertible. Let us clarify why elementary row operations on a matrix can be represented as multiplications by invertible matrices.

Example: Consider the following 2×4 matrix

$$\begin{bmatrix} 1 & 2 & 1 & 1 \\ -3 & 1 & 2 & -2 \end{bmatrix} \tag{27}$$

The interchange of the first and the second row is represented by the matrix T_1

$$\begin{bmatrix} -3 & 1 & 2 & -2 \\ 1 & 2 & 1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{T_1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ -3 & 1 & 2 & -2 \end{array}\right] \tag{28}$$

Similarly, multiplication of the first row by $-1/3$ is represented by the matrix T_2

$$\begin{bmatrix} 1 & -1/3 & -2/3 & 2/3 \\ 1 & 2 & 1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} -1/3 & 0 \\ 0 & 1 \end{bmatrix}}_{T_2} \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{T_1} \begin{bmatrix} 1 & 2 & 1 & 1 \\ -3 & 1 & 2 & -2 \end{bmatrix} \tag{29}$$

Finally, the subtraction of the first row from the second one is represented by the matrix T_3

$$\begin{bmatrix} 1 & -1/3 & -2/3 & 2/3 \\ 0 & 4/3 & 5/3 & 1/3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}}_{T_3} \underbrace{\begin{bmatrix} -1/3 & 0 \\ 0 & 1 \end{bmatrix}}_{T_2} \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{T_1} \begin{bmatrix} 1 & 2 & 1 & 1 \\ -3 & 1 & 2 & -2 \end{bmatrix}. \tag{30}$$

[2]Just apply T^{-1} to $TAx = Tb$ to obtain $Ax = b$.

The matrices T_1, T_2 and T_3 are all invertible, and therefore their product $T = T_3T_2T_1$ is invertible[3]. The invertibility of T establishes a one-to-one correspondence between the matrix (27) and the matrix at the left hand side of (30).

The matrix (30) is said to be in *row echelon form* A matrix is in row echelon form if:

Whenever two successive rows do not consist entirely of zeros, then the second row starts with a non-zero entry at least one step further to the right than the first row. All the rows consisting entirely of zeros are at the bottom of the matrix. The row echelon form of a matrix is not unique.

Let us now show how to solve a linear system by using Gauss elimination with pivoting by row. To this end, consider the linear system

$$\begin{cases} x_1 + 2x_2 + x_3 = 2 \\ 2x_1 + x_2 + x_3 = 1 \\ x_1 + x_2 + x_3 = 1 \end{cases} \tag{32}$$

This system can be written in a matrix-vector form as

$$\underbrace{\begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}}_{b}. \tag{33}$$

Define the following *augmented matrix* associated with (32) (or equivalently (33))

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array}\right] \tag{34}$$

Note that the augmented matrix is obtained by concatenating the column vector b to the right of the matrix A. As we shall see hereafter, the Gauss elimination method with pivoting by row yields an augmented matrix in row echelon form.

Let us know describe the Gauss elimination method with pivoting by row which will transform the augmented matrix (34) in row echelon form.

1. Pivoting step: We select the equation with the largest absolute value of a_{i1}, i.e., the second equation in (33), and we interchange it with the first to obtain

$$\begin{cases} 2x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 + x_3 = 2 \\ x_1 + x_2 + x_3 = 1 \end{cases}$$

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

(Augmented matrix of the new system)

[3]Recall that the inverse of a production of invertible matrices T_1, T_2 and T_3 is invertible and that

$$(T_3T_2T_1)^{-1} = T_1^{-1}T_2^{-1}T_3^{-1}. \tag{31}$$

2. Elimination step: We multiply the first equation by $-1/2$ and add it to the second and the third equation. This yields,

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

Therefore, we obtain

$$\begin{cases} 2x_1 + x_2 + x_3 = 1 \\ \dfrac{3}{2}x_2 + \dfrac{1}{2}x_3 = \dfrac{3}{2} \\ \dfrac{1}{2}x_2 + \dfrac{1}{2}x_3 = \dfrac{1}{2} \end{cases} \qquad \left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & 3/2 & 1/2 & 3/2 \\ 0 & 1/2 & 1/2 & 1/2 \end{array}\right]$$

(Augmented matrix of the new system)

3. Pivoting step: We look for the equation with the maximum absolute value of the coefficient $a_{j2}, (j \geq 2)$. In this case, it is the second equation. Hence, we do not do any permutation.

4. Elimination step: We multiply the second equation by $-1/3$ and we add it to the last one to eliminate x_2

$$\begin{cases} 2x_1 + x_2 + x_3 = 1 \\ \dfrac{3}{2}x_2 + \dfrac{1}{2}x_3 = \dfrac{3}{2} \\ \dfrac{1}{2}x_2 + \dfrac{1}{2}x_3 - \dfrac{1}{2}x_2 - \dfrac{1}{6}x_3 = \dfrac{1}{2} - \dfrac{1}{2} \end{cases} \Rightarrow \begin{cases} 2x_1 + x_2 + x_3 = 1 \\ \dfrac{3}{2}x_2 + \dfrac{1}{2}x_3 = \dfrac{3}{2} \\ \dfrac{1}{3}x_3 = 0 \end{cases}$$

Thus, we obtained

$$\begin{cases} 2x_1 + x_2 + x_3 = 1 \\ \dfrac{3}{2}x_2 + \dfrac{1}{2}x_3 = \dfrac{3}{2} \\ \dfrac{1}{3}x_3 = 0 \end{cases} \qquad \left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & 3/2 & 1/2 & 3/2 \\ 0 & 0 & 1/3 & 0 \end{array}\right] \tag{35}$$

(Augmented matrix in *row echelon form*)

At this point we can now use backward substitution (i.e. solve the system of equations form the bottom to the top). This yields the following unique solution to the system (33)

$$\begin{cases} x_3 = 0 \\ x_2 = \dfrac{2}{3}\left(\dfrac{3}{2} - \dfrac{1}{2}x_3\right) = \dfrac{2}{3}\left(\dfrac{3}{2} - \dfrac{1}{2}(0)\right) = 1 \\ x_1 = \dfrac{1}{2}(1 - x_2 - x_3) = \dfrac{1}{2}(1 - 1 - 0) = 0 \end{cases}$$

Remark: For a given system of linear equations, the row echelon forms is *not* unique. In fact there is infinite number of ways by which the augmented matrix of a linear system can be transformed in a row echelon form. For example, if we perform Gauss elimination without pivoting in (33), then we obtain the following row echelon form

$$\begin{cases} x_1 + 2x_2 + x_3 = 2 \\ -3x_2 - x_3 = -3 \\ \dfrac{1}{3}x_3 = 0 \end{cases} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & -3 & -1 & -3 \\ 0 & 0 & 1/3 & 0 \end{array}\right] \tag{36}$$

(Augmented matrix in *row echelon form*)

The row echelon forms (35) and (36) are different, but they are both obtained from by apply elementary row operations to the same linear system (33).

Reduced row echelon form. The Gauss elimination method with pivoting by row can be applied to any linear system of equations (e.g., 2 equations in 3 unknowns) to obtain a row echelon form. Once the row echelon form is available, then we can normalize the entries of a certain row by dividing them by the pivot, and then perform *backward elimination* to remove all entries above such pivot. In numerical linear algebra this is known as *Jordan backward elimination.* Let us show how this works. To this end, consider the system

$$\begin{cases} x_1 + 2x_2 + x_3 = 2 \\ x_2 + \dfrac{1}{3}x_3 = 1 \\ x_3 = 0 \end{cases} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & 1/3 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

(row echelon form)

Multiply the third equation by 1/3 and 1, respectively, and subtract it from the second and first equation, respectively. This yields

$$\begin{cases} x_1 + 2x_2 = 2 \\ x_2 = 1 \\ x_3 = 0 \end{cases} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

(still in row echelon form)

Finally, multiply the second equation by 2 and subtract it from the first equation to obtain

$$\begin{cases} x_1 = 0 \\ x_2 = 1 \\ x_3 = 0 \end{cases} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

(reduced row echelon form)

The augmented matrix of a linear system is in a *reduced row echelon form* if: 1) it is in an echelon form; and 2) in every pivot column, the pivot value is 1 and all other entries are 0. The reduced row echelon form of a matrix or linear system is *unique*.

Example: Consider the augmented matrix in row echelon form we obtained by performing Gauss elimination on (33) without pivoting, i.e.,

$$\begin{cases} x_1 + 2x_2 + x_3 = 2 \\ -3x_2 - x_3 = -3 \\ \dfrac{1}{3}x_3 = 0 \end{cases} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & -3 & -1 & -3 \\ 0 & 0 & 1/3 & 0 \end{array}\right]$$

(row echelon form)

To obtain the reduced row echelon form, we first rescale the third equation equation by multiplying it by 3. This yields,

$$\begin{cases} x_1 + 2x_2 + x_3 = 2 \\ -3x_2 - x_3 = -3 \\ x_3 = 0 \end{cases} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & -3 & -1 & -3 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

(row echelon form)

Next, we perform backward elimination of x_3 to obtain

$$\begin{cases} x_1 + 2x_2 = 2 \\ -3x_2 = -3 \\ x_3 = 0 \end{cases} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & -3 & 0 & -3 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

(row echelon form)

At this point, we rescale the second equation by $-1/3$ and use it to eliminate x_2 in the first equation. This yields

$$\begin{cases} x_1 = 0 \\ x_2 = 1 \\ x_3 = 0 \end{cases} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

(reduced row echelon form)

Note that the last column of the reduced-row echelon form is the solution of the system (33).

Example: The following matrices are in a reduced row echelon form

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

Example: The following matrices are not in a reduced row echelon form

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 5 & 7 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 3 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Remark: A linear system is said to be *consistent* if admits a solution. A system admits a solution (and therefore it is consistent) if and only if the row echelon form (or the reduced row echelon form) of the augmented matrix has no row of the form:

$$\left[\begin{array}{cccc|c} 0 & 0 & \dots & 0 & z \end{array}\right], \qquad z \neq 0$$

If the system is consistent then we can have one (unique) solution or infinitely many. An example of a system that is not consistent is the following

$$\begin{cases} x_1 + x_2 - x_3 = 1 \\ x_1 + x_2 - x_3 = 4 \end{cases}$$

This system defines two parallel planes (not intersecting). The reduced row echelon form of the augmented matrix is

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

and therefore the system is not consistent.

Computation of the inverse matrix. Let $A \in M_{n\times n}$ be an invertible matrix. By definition, the inverse of A is a square matrix denoted as A^{-1} with the following properties

$$AA^{-1} = I_n \qquad A^{-1}A = I_n, \tag{37}$$

where I_n is the $n \times n$ identity matrix. Let h_i be the columns of the matrix A^{-1}, i.e.,

$$A^{-1} = [h_1 \quad h_2 \quad \cdots \quad h_n] \qquad h_i \in M_{n\times 1} \quad i = 1, \dots, n. \tag{38}$$

By definition of matrix-vector product we have

$$AA^{-1} = [Ah_1 \quad Ah_2 \quad \cdots \quad Ah_n]. \tag{39}$$

At this point we define the following column vectors $e_i \in M_{n\times 1}$ $(i = 1, ..., n)$

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \qquad \cdots, \qquad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{40}$$

Note that e_i is the i-th column of the identity matrix I_n. With this notation we can write the matrix equation $AA^{-1} = I_n$ as

$$[Ah_1 \quad Ah_2 \quad \cdots \quad Ah_n] = [e_1 \quad e_2 \quad \cdots \quad e_n]. \tag{41}$$

Hence, the n columns of the inverse matrix A^{-1}, i.e., h_1, ..., h_n are solutions to n linear systems

$$Ah_1 = e_1, \qquad Ah_2 = e_2, \qquad \dots, \qquad Ah_n = e_n. \tag{42}$$

To solve these systems we can compute the reduced row echelon form of the following augmented matrices

$$\left[A \mid e_1\right], \qquad \left[A \mid e_2\right], \qquad \dots, \qquad \left[A \mid e_n\right] \tag{43}$$

If A is invertible, then A can be row-reduced to I_n. This means that the reduced row echelon form of the systems (43) is

$$\left[I_n \,\middle|\, h_1\right], \qquad \left[I_n \,\middle|\, h_2\right], \qquad \ldots, \qquad \left[I_n \,\middle|\, h_n\right], \tag{44}$$

where h_i is the i-the column of the inverse matrix.

More compactly, we can compute the reduced row echelon form of the matrix

$$\left[A \,\middle|\, I_n\right] \qquad \text{to obtain} \qquad \left[I_n \,\middle|\, A^{-1}\right] \tag{45}$$

Example: Compute the inverse of the following 2×2 matrix

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}. \tag{46}$$

We begin by constructing the augmented matrix $[A|I_2]$

$$\left[A \,\middle|\, I_2\right] = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right] \tag{47}$$

Then we transform the augmented matrix into row-reduced echelon form as

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right] \xrightarrow{R_2 : R_2 - R_1} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{array}\right] \xrightarrow{R_2 : -R_2} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & -1 \end{array}\right] \tag{48}$$

$$\xrightarrow{R_1 : R_1 - 2R_2} \left[\begin{array}{cc|cc} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & -1 \end{array}\right]. \qquad \underbrace{\begin{matrix} -1 & 2 \\ 1 & -1 \end{matrix}}_{A^{-1}} \tag{49}$$

Hence, the inverse of the matrix A defined in (46) is

$$A^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}. \tag{50}$$

It is good practice to verify that A^{-1} is indeed the inverse of A. To this end, we just need to check that $AA^{-1} = I_2$

$$AA^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \tag{51}$$

Example: Compute the inverse of the following 3×3 matrix

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}. \tag{52}$$

As before,

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3 : R_3 - R_1]{R_2 : R_2 - R_1} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & -2 & 1 & -1 & 0 & 1 \end{array}\right] \xrightarrow{R_2 : -R_2} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & -2 & 1 & -1 & 0 & 1 \end{array}\right]$$

$$
\xrightarrow{R_3\,:\,R_3+2R_2}
\left[\begin{array}{ccc|ccc}
1 & 2 & 0 & 1 & 0 & 0\\
0 & 1 & -1 & 1 & -1 & 0\\
0 & 0 & -1 & 1 & -2 & 1
\end{array}\right]
\xrightarrow{R_3\,:\,-R_3}
\left[\begin{array}{ccc|ccc}
1 & 2 & 0 & 1 & 0 & 0\\
0 & 1 & -1 & 1 & -1 & 0\\
0 & 0 & 1 & -1 & 2 & -1
\end{array}\right]
\xrightarrow{R_2\,:\,R_2+R_3}
$$

$$
\left[\begin{array}{ccc|ccc}
1 & 2 & 0 & 1 & 0 & 0\\
0 & 1 & 0 & 0 & 1 & -1\\
0 & 0 & 1 & -1 & 2 & -1
\end{array}\right]
\xrightarrow{R_1\,:\,R_1-2R_2}
\left[\begin{array}{ccc|ccc}
1 & 0 & 0 & 1 & -2 & 2\\
0 & 1 & 0 & 0 & 1 & -1\\
0 & 0 & 1 & -1 & 2 & -1
\end{array}\right]. \tag{53}
$$

Therefore, the inverse of the matrix A defined in (52) is

$$
A^{-1} = \begin{bmatrix}
1 & -2 & 2\\
0 & 1 & -1\\
-1 & 2 & -1
\end{bmatrix}. \tag{54}
$$

Lecture 6: Vector spaces

Vector spaces are sets in which we define an addition operation and a multiplication by a scalar satisfying certain number of properties. Let us first give a formal definition of vector space and then provide a few examples. Consider a nonempty set V in which define an addition operation "+" satisfying the following properties[1]:

1. $\forall u, v \in V \;\; (u+v) \in V$ $\qquad$ (V is closed under the addition operation)
2. $\forall u, v \in V \;\; u+v = v+u$ $\qquad$ (addition is commutative)
3. $\forall u, v, w \in V \;\; (u+v)+w = u+(v+w) \in V$ $\qquad$ (addition is associative)
4. $\exists\, 0_V \in V$ such that $u + 0_V = u \quad \forall u \in V$ $\qquad$ (additive neutral)
5. $\forall u \in V$, $\exists v \in V$ such that $u + v = 0_V$ $\qquad$ (opposite element)

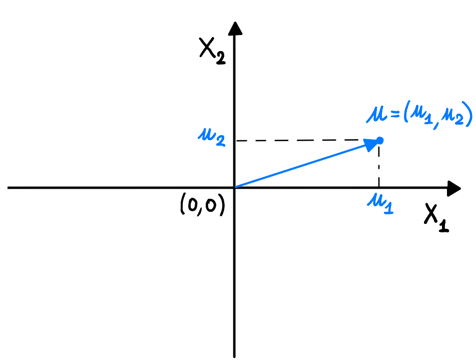
We also define the multiplication operation between an element of the set V and an element of a field K (e.g., $\mathbb{R}$ or $\mathbb{C}$) with the following properties:

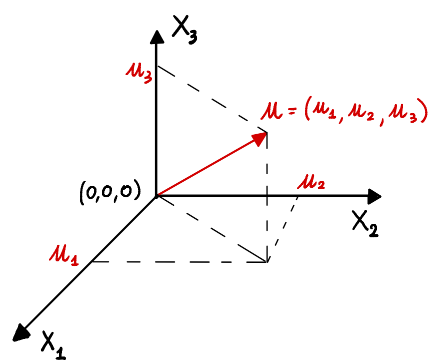
1. $av \in V \quad \forall a \in K, \quad \forall v \in V$
2. $(a+b)v = av + bv \quad \forall a, b \in K, \quad \forall v \in V$
3. $a(v+w) = av + aw \quad \forall a \in K, \quad \forall v, w \in V$
4. $(ab)v = a(bv) \quad \forall a, b \in K, \quad \forall v \in V$
5. $1v = v \quad 1 \in K, \forall v \in V$

Definition (Vector space)**.** A nonempty set V in which we define an addition operation and a multiplication operation satisfying the properties listed above is called *vector space* over K.

Let us provide a few examples of vector spaces over the real or complex numbers.

- The space $\mathbb{R}^n$ (n-tuples of real numbers) with the addition operation defined as $u + v = (u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n)$ is a vector space over $\mathbb{R}$. The neutral element with respect to the addition operation is $0_{\mathbb{R}^n} = (0, 0, \ldots, 0)$. Here is a simple visualization of a vector u in the vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$.





[1] A set V satisfying properties 1 to 5 is called "Abelian group".

- $V = M_{m\times n}(\mathbb{R})$, i.e., the set of real $m \times n$ matrices with the addition operation we defined in Lecture 4, is a vector space over $\mathbb{R}$. The neutral element with respect to the addition operation is

$$0_{M_{m\times n}} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \qquad \text{(zero matrix).} \tag{1}$$

- $V = M_{n\times m}(\mathbb{C})$ is a vector space over $\mathbb{C}$ and over $\mathbb{R}$.
- $V = \mathbb{P}_n(\mathbb{R})$, i.e., the space of polynomials of degree n with real coefficients, is a vector space over $\mathbb{R}$. An element of $\mathbb{P}_n(\mathbb{R})$ is

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n, \qquad a_j \in \mathbb{R}, \quad x \in \mathbb{R}. \tag{2}$$

The addition operation between two polynomials, say $p(x) = a_0 + a_1x + \cdots + a_nx^n$ and $q(x) = b_0 + b_1x + \cdots + b_nx^n$, is defined as

$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n, \qquad a_j, b_j \in \mathbb{R}, \quad x \in \mathbb{R}. \tag{3}$$

The neutral element with respect to the addition operation is the zero polynomial $p(x) = 0$.

- $V = C^{(1)}(\mathbb{R})$ (space of real-valued continuously differentiable functions defined on the real line) is a vector space over $\mathbb{R}$. An element of $C^{(1)}(\mathbb{R})$ is, e.g., $v(x) = e^{-x^2}\sin(x)$. The neutral element with respect to the addition operation is the zero function $v(x) = 0$.
- Then space of linear transformations between two vector spaces V and W is a vector space over $\mathbb{R}$. The elements of such vector space are linear maps $\mathcal{L} : V \to W$.

Vector subspace. Let V be a vector space over a field K. We say that $W \subseteq V$ is a vector subspace of V if

1. $0_V \in W$
2. $u, v \in W \quad \Rightarrow (u + v) \in W$
3. $cu \in W \quad \forall u \in W, \quad \forall c \in K$

Clearly, a vector subspace is itself a vector space. Note that the only condition we need for $W \subseteq V$ to be a vector subspace of V is that it is closed under addition and multiplication.

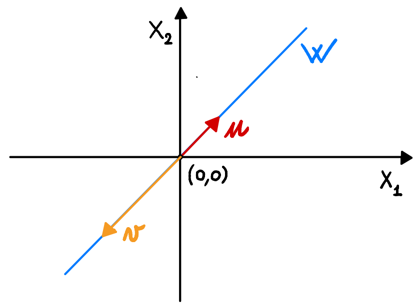
Example 1: A line passing through the origin of a Cartesian coordinate system is a vector subspace of $\mathbb{R}^2$. In fact, such line is defined by the set of points $(x_1, x_2) \in \mathbb{R}^2$ satisfying the equation $a_1x_1 + a_2x_2 = 0$ (for some $a_1, a_2 \in \mathbb{R}$). As we shall see hereafter, the set

$$W = \{(x_1, x_2) \in \mathbb{R}^2 \, : \, a_1x_1 + a_2x_2 = 0\}, \tag{4}$$

which represents the line, can be equivalently written as (assuming $a_2 \neq 0$)

$$W = \left\{u \in \mathbb{R}^2 \, : \, u = x(1, -a_1/a_2), \quad x \in \mathbb{R}\right\}. \tag{5}$$

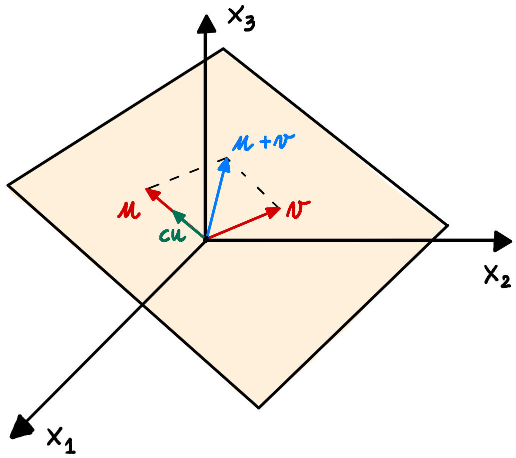
Clearly, W is a vector subspace of $\mathbb{R}^2$. In fact, 1) the zero of $\mathbb{R}^2$ is in W (the line passes through the origin); 2) a rescaling of a vector u on the line W is either zero or a vector that is still on the line; 3) the addition of two vectors u and v on the line is either zero or it is a vector that sits on the same line.



Example 2: A plane passing through the origin of a three-dimensional Cartesian coordinate system is a vector subspace of $\mathbb{R}^3$. Such plane can be defined as

$$W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \; : \; a_1x_1 + a_2x_2 + a_3x_3 = 0 \qquad a_1, a_2, a_3 \in \mathbb{R}\}. \tag{6}$$

Clearly, W is a vector subspace of $\mathbb{R}^3$. In fact, 1) the zero of $\mathbb{R}^3$ is in W (the plane passes through the origin); 2) a rescaling of a vector u on the plane is either zero or a vector that is still on the plane; 3) the addition of two vectors on the plane is either zero or a vector on the plane.



Example 3: The space of continuously differentiable functions is a vector subspace of the space of continuous functions. In fact: 1) the addition between two differentiable functions $f(x)$ and $g(x)$ is a differentiable function $f(x) + g(x)$; 2) multiplication of a differentiable function $f(x)$ by a scalar c is a differentiable function $cf(x)$.

Example 4: The space 3×3 symmetric matrices is a vector subspace of $M_{3\times3}(\mathbb{R})$. In fact, if A and B are symmetric then: 1) $A + B$ is symmetric, 2) the zero matrix $0_{M_{3\times3}}$ is symmetric, and 3) cA is symmetric for all $c \in \mathbb{R}$.

Example 5: The space of polynomials of degree at most 3, i.e., $\mathbb{P}_3(\mathbb{R})$, is a vector subspace of the space of polynomials of degree at most 8, i.e., $\mathbb{P}_8(\mathbb{R})$.

Linear combination. Let V be a vector space over K. A linear combination of $v_1, \ldots, v_n \in V$ is an expression of the form

$$x_1 v_1 + \cdots + x_n v_n. \tag{7}$$

We say that the set of vectors $v_1, \ldots, v_n \in V$ *generates* V if for every $v \in V$ there exist n numbers $x_1, \ldots, x_n \in K$ such that

$$v = x_1 v_1 + \cdots x_n v_n. \tag{8}$$

Example 1: The vectors

$$v_1 = (1, 0), \qquad v_1 = (1, 1), \tag{9}$$

generate $\mathbb{R}^2$.

Example 2: The matrices

$$v_1 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \tag{10}$$

generate the space of 2×2 matrices with real coefficients $M_{2\times 2}(\mathbb{R})$. Similarly, the matrices

$$v_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{11}$$

generate the space of 2×2 *symmetric* matrices.

Example 3: The polynomials

$$p_1(x) = 1, \qquad p_2(x) = x \qquad p_3(x) = x^2 \tag{12}$$

generate the vector space of polynomials of degree at most 2.

Definition. Let V be a vector space over K. The space generated by $v_1, \ldots, v_p \in V$ is called *span* of $v_1, \ldots, v_p$ and denoted by $\text{span}\{v_1, \ldots, v_p\}$.

Theorem 1. Let V be a vector space over K. The span of an arbitrary number of vectors $v_1, ..., v_p \in V$ is a vector subspace of V.

Proof. Let $v_1, \ldots, v_p$ be vectors in V. Consider the space generated by $v_1, \ldots, v_p$, i.e.,

$$W = \text{span}\{v_1, \ldots, v_p\} = \{v \in V : v = x_1 v_1 + \cdots + x_p v_p, \quad x_i \in K\}. \tag{13}$$

and pick two elements in W

$$u = x_1 v_1 + \cdots + x_p v_p, \qquad v = y_1 v_1 + \cdots + y_p v_p. \tag{14}$$

Clearly, $0_V \in Q$, $(u + v) \in W$, and $cu \in W$ (for all $c \in K$).

$\square$

By using the last theorem we immediately see why lines and planes are vector subspaces of $\mathbb{R}^3$. In fact, a line is a vector subspace generated by a nonzero vector $u \in \mathbb{R}^3$. Specifically, consider the line $(x_1, -3x_1, 2x_1)$ (for all $x_1 \in R$). This line is generated by the vector $u = (1, -3, 2)$. Similarly, the plane $x_1 + x_2 - 2x_3 = 0$ is generated, e.g., by the two vectors $v_1 = (1, 1, 1)$ and $v_2 = (2, 0, 1)$. In fact, any element on the plane can be expressed as a linear combination of v_1 and v_2.

Linear independence. Let V be a vector space over K, $v_1, \ldots, v_n \in V$. We say that n vectors $v_1, \ldots, v_n$ are linearly *independent* if

$$x_1 v_1 + \cdots + x_n v_n = 0_V \quad \Rightarrow \quad x_1, \ldots, x_n = 0 \tag{15}$$

Example 1: The following vectors of $\mathbb{R}^2$

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{16}$$

are linearly independent. In fact,

$$x_1 v_1 + x_2 v_2 = 0_{\mathbb{R}^2} \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases} \tag{17}$$

Example 2: The following two vectors of $\mathbb{R}^3$

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \tag{18}$$

are linearly independent. In fact,

$$x_1 v_1 + x_2 v_2 = 0_{\mathbb{R}^2} \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{19}$$

Let us compute the reduced row echelon form of the augmented matrix

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 0 \end{array}\right] \quad \Rightarrow \quad \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]. \tag{20}$$

Hence, the system is consistent (see the last row), i.e., it has a solution. Moreover, the solution is unique and given by $x_1 = x_2 = 0$.

Example 3: The following 2×2 matrices

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 2 \\ 3 & 4 \end{bmatrix} \tag{21}$$

are *linearly dependent.* In fact,

$$x_1 A + x_2 B + x_3 C = 0_{M_{2\times 2}} \quad \Leftrightarrow \quad \begin{bmatrix} x_1 + x_2 + 2x_3 & x_1 + x_2 + 2x_3 \\ 2x_1 + x_2 + 3x_3 & 3x_1 + x_2 + 4x_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{22}$$

which yields the system

$$\begin{cases} x_1 + x_2 + 2x_3 = 0 \\ 2x_1 + x_2 + 3x_3 = 0 \\ 3x_1 + x_2 + 4x_3 = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x_1 = -x_3 \\ x_2 = -x_3 \\ x_3 \quad \text{free} \end{cases} \tag{23}$$

Therefore the condition $x_1 A + x_2 B + x_3 C = 0_{M_{2\times 2}}$ implies that

$$x_3 A + x_3 B = x_3 C \qquad \forall x_3 \in \mathbb{R} \tag{24}$$

and therefore the matrices A, B and C are linearly dependent.

Basis of a vector space. Let V be a vector space over K. A basis of V is a set of linearly independent vectors in V that generate V.

Example: The vectors

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{25}$$

are a basis for the vector space $\mathbb{R}^2$. In fact, they are linearly independent and they generate $\mathbb{R}^2$. To show that they generate $\mathbb{R}^2$ we need to show that every vector $u \in \mathbb{R}^2$ can be represented as a linear combination of v_1 and v_2. In other words, given $u \in \mathbb{R}^2$ we need to show that there exist x_1 and x_2 such that

$$x_1 v_1 + x_2 v_2 = u. \tag{26}$$

This is equivalent to show that the following linear system of equations has a unique solution

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \tag{27}$$

which is obvious since the matrix of coefficients is invertible.

Definition (Coordinates relative to a basis)**.** Let V be a vector space over K, $v_1, \ldots, v_n \in V$ a basis for V, and $v \in V$. The numbers $x_1, \ldots, x_n$ such that $v = x_1 v_1 + \cdots + x_n v_n$ are called *coordinates* of v relative $v_1, \ldots, v_n$.

Theorem 2. The coordinates of an arbitrary vector v in a vector space V are uniquely determined by the basis.

Proof. Let $v_1, \ldots, v_n$ be a basis for V. Suppose that for some $v \in V$ there are two set of coordinates $\{x_i\}$ and $\{y_i\}$ such that

$$v = x_1 v_1 + \cdots + x_n v_n = y_1 v_1 + \cdots + y_n v_n \quad \Rightarrow \quad (x_1 - y_1) v_1 + \cdots + (x_n - y_n) v_n = 0_V. \tag{28}$$

This implies that $x_i = y_i$ since the vectors $v_1, \ldots, v_n$ are linearly independent.

□

Example: The coordinates of $v = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ relative to $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $v = \begin{bmatrix} -10 \\ 9 \end{bmatrix}$ can be computed by solving the linear system of equations

$$x_1 v_1 + x_2 v_2 = v \quad \Rightarrow \quad \begin{bmatrix} 1 & -10 \\ 1 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \tag{29}$$

Example: Find the coordinates of $p(x) = x^3 + x + 1$ relative to the following basis of $\mathbb{P}_3(\mathbb{R})$

$$p_0(x) = 5, \qquad p_1(x) = x, \qquad p_2(x) = x^2 + 1, \qquad p_3(x) = x^3 - x^2. \tag{30}$$

Let $y_0, \ldots, y_3$ be the coordinates of $p(x)$ relative to $\{p_0(x), \ldots, p_3(x)\}$. We have,

$$y_0 p_0(x) + \cdots + y_3 p_3(x) = x^3 + x + 1. \tag{31}$$

Developing the products we find

$$y_3 x^3 + (y_2 - y_3) x^2 + y_1 x + (5 y_0 + y_2) = x^3 + x + 1, \tag{32}$$

Which yields the linear system

$$\begin{cases} y_3 = 1 \\ y_2 - y_3 = 0 \\ y_1 = -1 \\ 5y_0 + y_2 = 1 \end{cases} \qquad \Rightarrow \qquad \begin{cases} y_3 = 1 \\ y_2 = 1 \\ y_1 = -1 \\ y_0 = 0 \end{cases} \tag{33}$$

Example: The coordinates of the symmetric matrix

$$v = \begin{bmatrix} -2 & 3 \\ 3 & 4 \end{bmatrix}, \tag{34}$$

relative to the basis v_1, v_2 and v_3 defined in Eqs. (11) are $x_1 = -2$, $x_2 = 4$ and $x_3 = 3$.

Dimension of a vector space. The dimension of a vector space V is the number of linearly independent vectors required to generate V, i.e., the number of elements in any basis of V. We denote the dimension of V as $\dim(V)$. We have, for example,

- $\dim(M_{2\times 2}(\mathbb{R})) = 4$,
- $\dim(\mathbb{R}^3) = 3$,
- $\dim(\mathbb{P}_4(\mathbb{R})) = 5$,
- $\dim\left(C^{(1)}(\mathbb{R})\right) = \infty$

It is easy to show that if $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_m$ are $m > n$ vectors of V then $w_1, \ldots, w_m$ are necessarily linear dependent. This means that number of vectors in every basis of V is that minimum one that is needed to generate V. To show this, we let us write each vector w_i in terms of the basis

$$\begin{cases} w_1 = x_{11}v_1 + \cdots + x_{1n}v_n \\ \vdots \\ w_m = x_{m1}v_1 + \cdots + x_{mn}v_n \end{cases} \tag{35}$$

Now, suppose that $w_1, \ldots, w_m$ are linearly independent, i.e.,

$$0_V = y_1w_1 + \cdots + y_mw_m \quad \Rightarrow \quad y_1, \ldots, y_m = 0. \tag{36}$$

By substituting (35) into (36) we obtain,

$$0_V = y_1w_1 + \cdots y_mw_m = (y_1x_{11} + \cdots + y_mx_{m1})v_1 + \cdots + (y_1x_{1n} + \cdots + y_mx_{mn})v_n, \tag{37}$$

which implies that

$$\begin{cases} y_1x_{11} + \cdots + y_mx_{m1} = 0 \\ \vdots \\ y_1x_{1n} + \cdots + y_mx_{mn} = 0 \end{cases} \tag{38}$$

This is a homogeneous linear system of $n < m$ equation in m unknowns $(y_1, \ldots, y_m)$. which always admits a nontrivial (i.e., nonzero) solution. Hence, $y_1, ..., y_m$ cannot be all zero, and therefore $w_1, ..., w_m$ are necessarily linearly dependent.

We conclude this section by emphasizing that a set of p linearly independent vectors in a vector space V of dimension $n > p$ can be always complemented with additional linearly independent vectors to become a basis of V.

The rank of a matrix. Consider the following $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}. \tag{39}$$

The columns of A generate a vector space called *column space of* A. Similarly, the rows of A generate a vector space called *row space of* A

$$\text{Column space of } A\text{:} \quad \text{span} = \left\{ \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} \right\}. \tag{40}$$

$$\text{Row space of } A: \quad \text{span} = \left\{ \begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix}, \begin{bmatrix} a_{21} & \cdots & a_{2n} \end{bmatrix}, \dots, \begin{bmatrix} a_{m1} & \cdots & a_{mn} \end{bmatrix} \right\}. \tag{41}$$

Note that the column space of A is a vector subspace of $\mathbb{R}^m$, while the row space of A is a vector subspace of $\mathbb{R}^n$.

The dimension of the column space is called *column rank*, while the dimension of the row space is called *row rank*. Both ranks can be computed by reducing the matrix to an echelon form using elementary row or column operations, i.e.,

1. Adding a scalar multiple of one row (column) to another row (column);
2. Interchange rows (columns),
3. Multiplying one row (column) by a non-zero number.

Theorem 3. Elementary row or column operations do not change the row rank nor the column rank of a matrix[2].

This statement follows immediately by noting that linear taking linear combinations of a fixed number of vectors does not change the dimension of the span of such vectors. Moreover, taking permutations of the entries of a set of vectors in the same way for all vectors does not alter linear independence.

By performing both rows and column operations it is possible to transform any $m \times n$ matrix into the following canonical form (block matrix)

$$A = \begin{bmatrix} I_r & 0_{M_{r\times(n-r))}} \\ 0_{M_{(m-r)\times r}} & 0_{M_{(m-r)\times(n-r)}} \end{bmatrix}, \tag{42}$$

where I_r is a $r \times r$ identity matrix, and all other matrices are zero matrices.

This means that the dimension of the row space of a matrix is always the same as the dimension of the column space. Phrasing this differently:

Theorem 4. The row rank of a matrix is always the same as the column rank.

Hence, we can omit "row" or "column" and just speak of the *rank of a matrix*. Clearly, for an $m \times n$ matrix the rank r is always smaller or equal than the minimum between the number of rows m and the number of columns n, i.e.,

$$r \leq \min\{m, n\}. \tag{43}$$

Example 1: By using elementary row and column operations reduce the matrix

$$\begin{bmatrix} 1 & 1 & 1 & 2 & -1 \\ 0 & 2 & 1 & 0 & 1 \\ 2 & 0 & 1 & 4 & -1 \end{bmatrix} \tag{44}$$

[2]Note that elementary column operations can change the solution to a linear system of equations. In fact, if we perform Gauss elimination along a row we are essentially eliminating the coefficient multiplying, say, x_k using the coefficient of the variable x_j. Clearly, this changes the solution of the linear system.

to the canonical form (42).

$$\begin{bmatrix}1 & 1 & 1 & 2 & -1\\ 0 & 2 & 1 & 0 & 1\\ 2 & 0 & 1 & 4 & -1\end{bmatrix} \xrightarrow{R_3:R_3-2R_1} \begin{bmatrix}1 & 1 & 1 & 2 & -1\\ 0 & 2 & 1 & 0 & 1\\ 0 & -2 & -1 & 0 & 1\end{bmatrix} \xrightarrow[C_4:C_4-2C_1]{C_2:C_2-C_1} \begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 0 & 2 & 1 & 0 & 1\\ 0 & -2 & -1 & 0 & 1\end{bmatrix} \tag{45}$$

$$\xrightarrow{R_3:R_3+R_2} \begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 0 & 2 & 1 & 0 & 1\\ 0 & 0 & 0 & 0 & 2\end{bmatrix} \xrightarrow[C_2:C_2-2C_1]{C_4:C_4-C_2} \begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 2\end{bmatrix} \xrightarrow[C_2\leftrightarrow C_3]{C_3\leftrightarrow C_5/2} \begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\end{bmatrix} \tag{46}$$

Hence, the rank of the matrix (44) is $r = 3$.

Example 2: Find the rank of the matrix

$$A = \begin{bmatrix}1 & 3 & 4\\ 1 & 1 & -1\\ 2 & 6 & 0\end{bmatrix}. \tag{47}$$

A is row equivalent to the following matrix[3]

$$A = \begin{bmatrix}1 & 3 & 4\\ 0 & -2 & -5\\ 0 & 0 & -8\end{bmatrix}. \tag{48}$$

Clearly, the columns of this matrix are linearly independent and therefore the rank is 3.

Example: The rank of the following matrices is equal to 2

$$A = \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}, \quad B = \begin{bmatrix}1 & 2 & 1\\ 0 & 0 & 1\\ -1 & -2 & 0\end{bmatrix}, \quad C = \begin{bmatrix}1 & 1 & 1 & 2 & -1\\ 0 & 2 & 1 & 0 & -1\\ 2 & 0 & 1 & 4 & -1\end{bmatrix}, \quad D = \begin{bmatrix}1 & 1\\ 0 & 2\\ 2 & 0\\ 1 & 1\end{bmatrix}. \tag{49}$$

[3]Recall that elementary row or column operations do not change the rank of a matrix.

Lecture 7: Linear transformations

Let V and W be two vector spaces over a field K. We say that a transformation

$$F : V \mapsto W \tag{1}$$

is *linear* if

1. $F(u+v) = F(u) + F(v) \qquad \forall u, v \in V,$
2. $F(cu) = cF(u) \qquad \forall u \in V, \quad \forall c \in K.$

Conditions 1. and 2. imply that

$$F(au+bv) = aF(u) + bF(v) \qquad \forall u, v \in V, \quad \forall a, b \in K. \tag{2}$$

Let us discuss a few examples of linear and nonlinear transformations.

- *Example 1:* The transformation

$$\begin{aligned} F : \mathbb{R} &\to \mathbb{R} \\ x &\to \sin(x) \end{aligned}$$

is nonlinear. In fact, $\sin(x+y) \neq \sin(x) + \sin(y)$ for arbitrary x and y in $\mathbb{R}$.

- *Example 2:* Let $V = C^{(1)}(\mathbb{R})$ (vector space of real-valued continuously differentiable functions), $W = C^{(0)}(\mathbb{R})$ (vector space of real-valued continuous functions), $K = \mathbb{R}$. The transformation

$$\begin{aligned} F : C^1(\mathbb{R}) &\to C^0(\mathbb{R}) \\ f(x) &\to \frac{df(x)}{dx} \end{aligned}$$

is linear. In fact, we have

$$\frac{d}{dx}(af(x) + bg(x)) = a\frac{df(x)}{dx} + b\frac{dg(x)}{dx} \qquad \forall f, g \in C^{(1)}(\mathbb{R}), \quad \forall a, b \in \mathbb{R}. \tag{3}$$

- *Example 3:* The transformation

$$\begin{aligned} F : \mathbb{R}^3 &\to \mathbb{R}^2 \\ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &\to \begin{bmatrix} x_1 - x_2 \\ 2x_1 + x_2 - x_3 \end{bmatrix} \end{aligned}$$

is linear. In fact, we have

$$F\left(a\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + b\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\right) = \begin{bmatrix} a(x_1 - x_2) + b(y_1 - y_2) \\ a(2x_1 + x_2 - x_3) + b(2y_1 + y_2 - y_3) \end{bmatrix} = aF\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) + bF\left(\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\right).$$

- *Example 4:* The transformation

$$\begin{aligned} F : \mathbb{R}^3 &\to \mathbb{R}^2 \\ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &\to \begin{bmatrix} x_1 + x_2 + 1 \\ x_3 + x_1 \end{bmatrix} \end{aligned} \tag{4}$$

is <u>*not*</u> linear. In fact,

$$F\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\right) \neq F\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) + F\left(\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\right).$$

Transformations of the form (4) are called *affine* transformations. Affine transformations are obtained by adding a constant vector to a linear transformation. For the transformation (4) we have

$$F\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) \to \underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\text{linear transformation}} + \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\text{constant vector}} . \tag{5}$$

- *Example 5:* The transformation[1]

$$\begin{aligned} \text{trace} : M_{n\times n}(\mathbb{R}) &\to \mathbb{R} \\ A &\to \sum_{k=1}^{n} a_{kk} \quad \text{(trace of the matrix } A) \end{aligned} \tag{6}$$

is linear. In fact,

$$\text{trace}(aA + bB) = a\,\text{trace}(A) + b\,\text{trace}(B). \tag{7}$$

Hereafter we show that the composition of two linear transformation is a linear transformation.

Theorem 1. Let U, V, and W be vector spaces. Consider the linear transformations $F : U \to V$ and $G : V \to W$. Then $G(F(u)) : U \to W$ is a linear transformation.

Proof. If F and G are linear transformations then

$$G(F(au + bv)) = G(aF(u) + bF(v)) = aG(F(u)) + bG(F(v)). \tag{8}$$

Hence, the composition of F and G is a linear transformation.

□

[1]The trace of a square matrix is defined to be the sum of all diagonal entries of A.

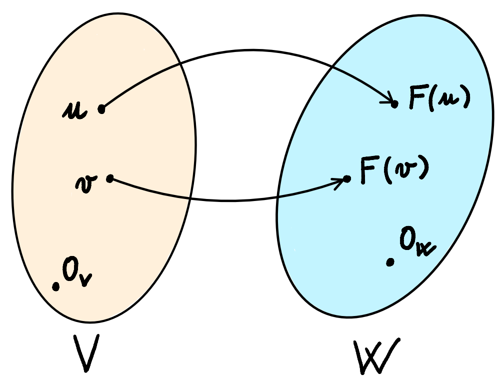
Injective, surjective and invertible transformations. Let V and W be two vector spaces. Consider the following transformation

$$F : V \to W \tag{9}$$

Here, F can be linear on nonlinear.

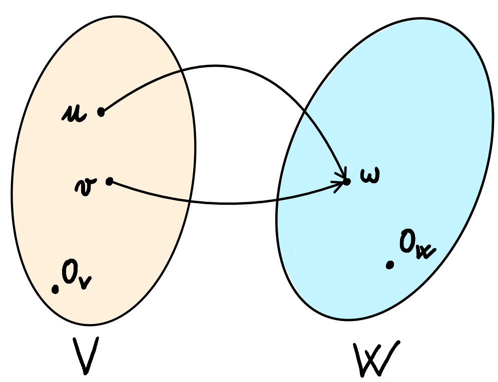
1. We say that F is *injective* or *one-to-one* if:

$$\text{for all} \quad u, v \in V \quad F(u) = F(v) \quad \Rightarrow \quad u = v \tag{10}$$



2. We say that F is *surjective* or *onto* if

$$\text{for all} \quad w \in W \quad \text{there exists (at least one)} \quad u \in V \quad \text{such that} \quad F(u) = w \tag{11}$$



Note that there may be more than one element in V that is mapped onto w. In the figure above, two elements u and v are mapped onto the same element w.

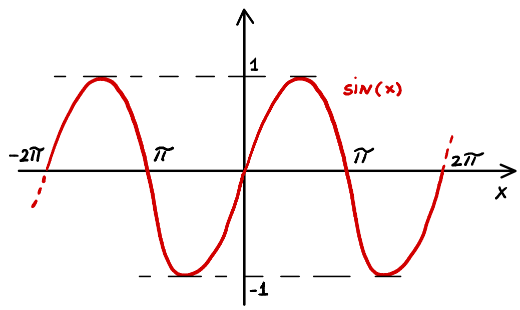
3. We say that F is *invertible*[2] if is is one-to-one and onto (injective and surjective).

[2] Invertible transformations are often called bijections or bijective transformations.

Example 6: The nonlinear transformation

$$\begin{aligned} F : \mathbb{R} &\to \mathbb{R} \\ x &\to \sin(x) \end{aligned}$$

is not injective nor surjective on the real line.



In fact, there are multiple points on the x axis with the same value of $\sin(x)$. For example,

$$\sin(1) = \sin(1 + 2k\pi) \quad k \in \mathbb{Z}. \tag{12}$$

Hence the function is not injective. The function $\sin(x)$ is also not surjective in $\mathbb{R}$, as there is no $x \in \mathbb{R}$ such that $\sin(x) = 2$. However, if we restrict the domain and range of F as follows

$$\begin{aligned} F : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] &\to [-1, 1] \\ x &\to \sin(x) \end{aligned}$$

then F is invertible, since it is injective and surjective. The inverse function is denoted by $\sin^{-1}(x)$ or $\arcsin(x)$

Example 7: The linear transformation

$$\begin{aligned} F : \mathbb{R}^2 &\to \mathbb{R}^2 \\ \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} &\to \underbrace{\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \begin{bmatrix} x_1 + 2x_2 \\ -x_1 + x_2 \end{bmatrix}, \end{aligned}$$

is one-to-one and onto. In fact it is easy to show $Ax = Ay$ implies $x = y$ (inejctivity), and that for each $y \in \mathbb{R}$ there exits $x \in \mathbb{R}^2$ such that $Ax = y$. Therefore the transformation F is invertible. The inverse transformation is defined by the inverse matrix A^{-1}

$$\begin{aligned} F^{-1} : \mathbb{R}^2 &\to \mathbb{R}^2 \\ \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} &\to \underbrace{\frac{1}{3}\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}}_{A^{-1}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \frac{1}{3}\begin{bmatrix} x_1 - 2x_2 \\ x_1 + x_2 \end{bmatrix}. \end{aligned}$$

Definition. Let V, W be vector spaces, $F : V \to W$ a linear transformation. If F is invertible then we say that F is an *isomorphism* between V and W. If there exists an isomorphism between the vector spaces V and W (i.e., an invertible linear trasformation) then we say that V and W are *isomorphic.*

Theorem 2. Let V be a vector space of dimension n over a field K. Then V is isomorphic to K^n.

Proof. Let $v_1, \dots, v_n \in V$ be a basis of F. Any vector $v \in V$ can be represented uniquely relative to the basis as

$$v = x_1 v_1 + \cdots + x_n v_n \qquad x_i \in K. \tag{13}$$

The transformation

$$F : V \to K^n \tag{14}$$

$$v \to \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{15}$$

is linear, one-to-one and onto. These properties follow immediately from the definition of basis (surjectivity), and from the fact that the coordinates of $v \in V$ relative to a basis are unique (injectivity). Hence, (15) defines a bijection between V and K^n. This means that V is isomorphic to K^n.

□

Example 8: The space of polynomials of degree at most 4 with real coefficients, i.e., $\mathbb{P}_4(\mathbb{R})$, is isomorphic to $\mathbb{R}^5$. In fact, if we set up a basis for $\mathbb{P}_4(\mathbb{R})$, i.e., a set of 5 linearly independent polynomials of degree at most 4, e.g.,

$$p_4(x) = x^4 - 3x, \quad p_3(x) = x^3, \quad p_2(x) = x^3 + x^2 + 1, \quad p_1(x) = x - x^3, \quad p_0(x) = x^2 + 1, \tag{16}$$

then we see that each polynomial in $p \in \mathbb{P}_4(\mathbb{R})$ is uniquely identified by 5 real coefficients $(x_0, \dots, x_4)$:

$$p(x) = x_4 p_4(x) + x_3 p_3(x) + x_2 p_2(x) + x_1 p_1(x) + x_0 p_0(x). \tag{17}$$

Hence, there exists a bijection between $\mathbb{R}^5$ and the space of polynomials $\mathbb{P}_4(\mathbb{R})$. In other words, $\mathbb{P}_4(\mathbb{R})$ and $\mathbb{R}^5$ are isomorphic.

Example 9: The vector space of 3×3 *symmetric* matrices with real coefficient is isomorphic to $\mathbb{R}^6$.

Since the inverse of an isomorphism is an isomorphism we have that all vector spaces of dimension n over some field K are isomorphic to one another. For example, the vector space of polynomials of degree at most 3 is isomorphic to the vector space of 2×2 matrices with real coefficients.

Theorem 3. The set of all linear mappings between two vector spaces V and W is a vector space. Such a space is denoted by $\mathcal{L}(V, W)$.

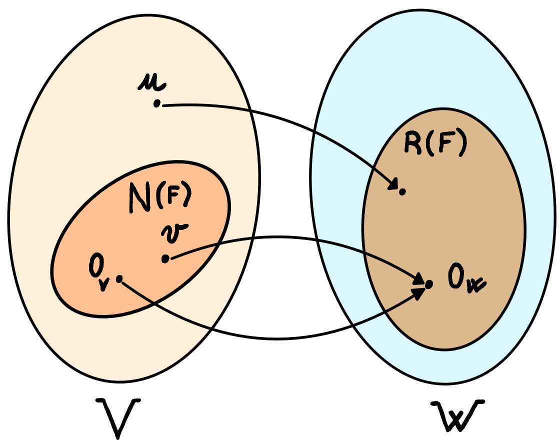
Nullspace and range of a linear transformation. Let V, W be vector spaces. Consider the linear transformation

$$F : V \to W. \tag{18}$$

- The *nullspace* (or kernel[3]) of F is the set vectors in V that are mapped into 0_W (zero vector of W), i.e.,

$$N(F) = \{v \in V \quad \text{such that} \quad F(v) = 0_W\} \qquad \text{(nullspace of } F\text{)}. \tag{19}$$

  Clearly, since F is linear we have that the element 0_V is always mapped onto 0_W. Therefore, 0_V is always in the nullspace of F.



- The *range* of F is the set of vectors w in W such that w is the image of some $v \in V$ under F, i.e., there exists $v \in V$ such that $F(v) = w$.

$$R(F) = \{F(v) \in W \quad \text{such that} \quad v \in V\} \tag{20}$$

  Note that the range of $R(F)$ has 0_W in it. In fact, since F is linear we have that $F(0_V) = 0_W$.

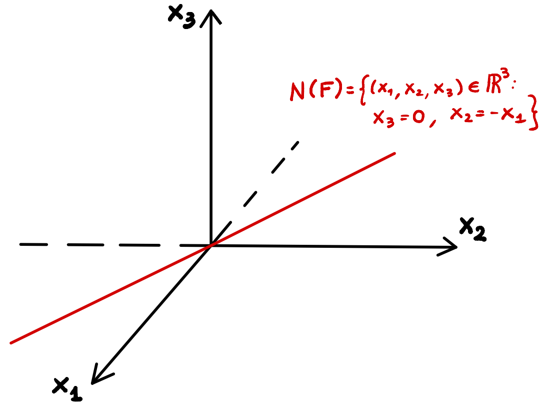
Let us determine the nullspace and the range of simple linear transformations.

Example 10: Consider the following linear transformation

$$\begin{aligned} &F : \mathbb{R}^3 \to \mathbb{R}^2 \\ &\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \to \begin{bmatrix} x_1 + x_2 + x_3 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \end{aligned} \tag{21}$$

The nullspace of F is the set of vectors in $\mathbb{R}^3$ that mapped onto the zero vector of $\mathbb{R}^2$. Hence, the nullspace of F is defined by the following homogeneous linear system of equations

$$\begin{cases} x_1 + x_2 + x_3 = 0 \\ x_3 = 0 \end{cases} \qquad \Rightarrow \qquad \begin{cases} x_1 = -x_2 \\ x_3 = 0 \end{cases} \tag{22}$$



Note that the nullspace of F is a vector subspace of $\mathbb{R}^3$ (line passing through the origin). The range of F can be constructed by taking an arbitrary element of $\mathbb{R}^3$ and mapping it via F. Such range coincides with *column space* of the matrix A, i.e., the span of the columns of A. In fact,

$$\begin{bmatrix} x_1 + x_2 + x_3 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{23}$$

Hence,

$$R(F) = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^2. \tag{24}$$

Theorem 4. Let V, W be vector spaces, $F : V \to W$ linear. Then

1. $N(F)$ is a vector subspace of V.
2. $R(F)$ is a vector subspace of W.

Proof. Let $u, v \in N(F)$. Clearly, $u + v$ is in $N(F)$. In fact, since F is linear we have $F(u + v) = F(u) + F(v) = 0_W$. Thus, $u + v$ is in $N(F)$. Moreover, $0_V \in N(F)$ and $cu \in N(F)$ for all $u \in N(F)$ and all $c \in K$. This implies that $N(F)$ is a vector subspace of V. To prove that $R(F)$ is a vector subspace of W, let $w, s \in R(F)$. This means that there exist $u, v \in V$ such that $F(u) = w$ and $F(v) = s$. Obviously, $(w + s) \in R(F)$. In fact, by using the linearity of F we have $F(u + v) = w + s$, and therefore $w + s \in R(F)$. Also, 0_W is $R(F)$ and $cu \in R(F)$ for all $u \in R(F)$. Thus, $R(F)$ is a vector subspace of W.

□

The nullspace and the range of linear transformation also characterize the injectivity and surjectivity of the transformation. In particular we have the following theorems.

Theorem 5. Let V, W be vector spaces, $F : V \to W$ a linear transformation. Then F is injective (one-to-one) if and only if $N(F) = \{0_V\}$, i.e., the if nullspace of F reduces the single element $\{0_V\}$.

[3]The nullspace/kernel of a linear transformation F is often denoted as $\ker(F)$.

Proof. To prove the theorem we need to prove two statements:

1. F is injective $\Rightarrow N(F) = \{0_V\}$.

   Suppose that F is one-to-one. We want to show that this implies $N(F) = \{0_V\}$. To this end, let $v \in N(F)$, i.e., $F(v) = 0_W$. Clearly $v = 0_V$ is mapped onto 0_W, i.e., $0_V \in N(F)$. The assumption that F is one-to-one rules out the existence of any other element in V mapped onto 0_W. In other words, 0_V is the only element of V mapped into 0_W. Hence, if F is one-to-one then $N(F) = \{0_V\}$.

2. $N(F) = \{0_V\} \Rightarrow F$ is injective.

   Conversely, let us assume that $N(F) = \{0_V\}$. We want to show that this implies that F is one-to-one. To this end, suppose there are two elements $u, v \in V$ such that $F(u) = F(v)$. By using the linearity of F we have $F(u - v) = 0_W$, i.e., $(u - v) \in N(F)$. Since, by assumption, the only element in the nullspace of F is 0_V we have that $u - v = 0_V$, i.e., $u = v$. In other words, $N(F) = \{0_V\}$ implies that F is one-to-one.

□

Theorem 6. Let V, W be vector spaces, $F : V \to W$ linear. Then F is surjective (onto) if and only if $\dim(R(F)) = \dim(W)$.

Proof. As before, to prove the theorem we need to prove two statements:

1. F is surjective $\Rightarrow \dim(R(F)) = \dim(W)$,
2. F is surjective $\Leftarrow \dim(R(F)) = \dim(W)$.

Let F be surjective (or onto), i.e., $\forall w \in W$ there exists at least one $v \in V$ such that $F(v) = w$. This means that $R(F) = W$ and therefore $\dim(R(F)) = \dim(W)$. Conversely, suppose that $\dim(R(F)) = \dim(W)$. We know that $R(F)$ is a vector subspace of W. Since the dimension of $R(F)$ and W are the same (by assumption) then $R(F) = W$, i.e., F is surjective (or onto).

□

Next we discuss a very important theorem for linear transformations between vector spaces.

Theorem 7. Let V and W be vector space and $F : V \to W$ be any linear transformation. Then

$$\dim(V) = \dim(N(F)) + \dim(R(F)). \tag{25}$$

Proof. If $R(F) = 0_W$ the statement is trivial since the entire V is mapped to the 0_W. This implies $N(F) = V$, and of course $\dim(N(F)) = \dim(V)$. Consider now $\dim(R(F)) = s > 0$ and let $\{w_1, \ldots, w_s\}$ be a basis of $R(F)$. Then there exist s elements $v_1, \ldots, v_s \in V$ such that $F(v_1) = w_1, \ldots, F(v_s) = w_s$. Suppose $\dim(N(F)) = q$ and let $\{u_1, \ldots, u_q\}$ be a basis for $N(F)$.

We would like to show that $\{u_1, \ldots, u_q, v_1, \ldots, v_s\}$ is a basis of V[4]. To this end, pick an arbitrary $v \in V$. Then, there exists $x_1, \ldots, x_s \in K$ such that $F(v) = x_1 w_1 + \ldots + x_s w_s$ (since $w_1, \ldots, w_s$ is a

[4]Note that if $\{u_1, \ldots, u_q, v_1, \ldots, v_s\}$ is a basis of V then $\dim(V) = q+s$, where $q = \dim(N(F)$ and $s = \dim(R(F))$.

basis for $R(F)$). Recalling that $F(v_1) = w_1$, ..., $F(v_s) = w_s$

$$\begin{aligned} F(v) &= x_1 F(v_1) + \ldots + x_s F(v_s) \\ &= F(x_1 v_1 + \ldots + x_s v_s). \end{aligned}$$

By using the linearity of F we obtain

$$F(v - x_1 v_1 - \cdots - x_s v_s) = 0_W \quad \Rightarrow \quad (v - x_1 v_1 - \cdots - x_s v_s) \in N(F).$$

At this point we represent $(v - x_1 v_1 - \cdots - x_s v_s)$ relative to the basis of $N(F)$

$$v - x_1 v_1 - \ldots - x_s v_s = y_1 u_1 + \ldots + y_q u_q,$$

i.e.,

$$v = x_1 v_1 + \ldots + x_s v_s + y_1 u_1 + \ldots + y_q u_q.$$

This shows that $V = \text{span}\{v_1, \ldots, v_s, u_1, \ldots, u_q\}$, i.e., that V is generated by $\{v_1, \ldots, v_s, u_1, \ldots, u_q\}$. To prove the theorem it remains to prove that the the vectors $\{v_1, \ldots, v_s, u_1, \ldots, u_q\}$ are linearly independent. In this way we can claim that $n = s + q$, i.e., $\dim(V) = \dim(N(F)) + \dim(R(F))$.

To this end, consider the linear combination

$$x_1 v_1 + \ldots + x_s v_s + y_1 u_1 + \ldots + y_q u_q = 0_V. \tag{26}$$

By applying F and recalling that $F(u_i) = 0_W$ ($u_i \in N(F)$) we obtain

$$x_1 w_1 + \ldots + x_s w_s = 0_W \quad \Rightarrow \quad x_1, \ldots, x_s = 0. \tag{27}$$

In fact $\{w_1, \ldots, w_s\}$ is a basis for $R(F)$ and therefore w_i are linearly independent. Substituting this result back into (26) yields

$$y_1 u_1 + \ldots + y_q u_q = 0_V \quad \Rightarrow \quad y_1, \ldots, y_q = 0 \tag{28}$$

since $\{u_1, \ldots, u_q\}$ is a basis for $N(F)$. Equations (27), (28) and (26) allow us to conclude that $\{v_1, \ldots, v_s, u_1, \ldots, u_q\}$ are linearly independent. Moreover the vectors $\{v_1, \ldots, v_s, u_1, \ldots, u_q\}$ generate V, and therefore they are a basis for V. This implies that

$$\dim(V) = s + q = \dim(N(F)) + \dim(R(F)). \tag{29}$$

$\square$

Matrix rank theorem. Theorem 7 can be applied to linear transformations defined by matrices. To this end, consider the transformation $F : \mathbb{R}^n \to \mathbb{R}^m$ from $\mathbb{R}^n$ into $\mathbb{R}^m$ defined as $F(x) = Ax$, where A is an $m \times n$ matrix:

$$\underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{x} \rightarrow \underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{x} \tag{30}$$

We know that range of F coincides with the column space of A. Also the dimension of the column space is the *rank* of the matrix A. Therefore from equation (25) it follows that

$$\boxed{n = \dim(N(A)) + \operatorname{rank}(A)}. \tag{31}$$

Matrix associated with a linear transformation Let V and W be finite-dimensional vector spaces, and let

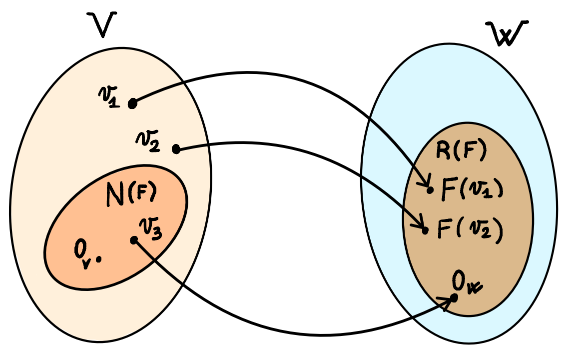
$$F : V \to W \tag{32}$$

an arbitrary linear transformation. In this section we show how to represent F in terms of a matrix. To this end, suppose that

$$\begin{aligned} \mathcal{B}_V &= \{v_1, \ldots, v_n\} &\to\quad &\text{basis of } V, &\dim(V) = n, \\ \mathcal{B}_W &= \{w_1, \ldots, w_m\} &\to\quad &\text{basis of } W, &\dim(W) = m. \end{aligned}$$

The transformation F is *uniquely determined* by the image of the basis $\mathcal{B}_V$ under F, i.e.,

$$\{v_1, \ldots, v_n\} \quad \to \quad \{F(v_1), \ldots, F(v_n)\}. \tag{33}$$



Clearly, for all $i = 1, \ldots, n$ we have that $F(v_i) \in R(F) \subseteq W$. Therefore, each $F(v_i)$ can be represented in terms of the basis $\mathcal{B}_W$ as

$$\begin{cases} F(v_1) = a_{11} w_1 + \cdots + a_{m1} w_m \\ \quad\vdots \\ F(v_n) = a_{1n} w_1 + \cdots + a_{mn} w_m \end{cases}. \tag{34}$$

Note that a_{ij} is the i-th component of $F(v_j)$ relative to the basis $\{w_1, \ldots, w_m\}$. The matrix associated with the linear transformation F depends bases $\mathcal{B}_V$ and $\mathcal{B}_W$ and it is defined as

$$A_{\mathcal{B}_V}^{\mathcal{B}_W}(F) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}. \tag{35}$$

Next, consider an arbitrary element $v \in V$, and represent it in terms of the basis $\mathcal{B}_V$

$$v = x_1 v_1 + \cdots + x_n v_n. \tag{36}$$

By applying F and taking (34) into account we obtain

$$\begin{aligned} F(v) =& x_1 F(v_1) + \cdots + x_n F(v_n) \\ =& x_1 \left(a_{11} w_1 + \cdots + a_{m1} w_m\right) + \cdots + x_n \left(a_{1n} w_1 + \cdots + a_{mn} w_m\right) \\ =& \underbrace{\left(a_{11} x_1 + \cdots + a_{1n} x_n\right)}_{y_1} w_1 + \cdots + \underbrace{\left(a_{m1} x_1 + \cdots + a_{mn} x_n\right)}_{y_m} w_m. \end{aligned} \tag{37}$$

At this point we define the following two column vectors

$$[v]_{\mathcal{B}_V} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad [F(v)]_{\mathcal{B}_W} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \tag{38}$$

representing the *coordinates* of v and $F(v)$ relative to the bases $\mathcal{B}_V$ and $\mathcal{B}_W$, respectively[5]. With this notation, we see from (37) and (35) that

$$[F(v)]_{\mathcal{B}_W} = A_{\mathcal{B}_V}^{\mathcal{B}_W}(F)[v]_{\mathcal{B}_V}. \tag{39}$$

Therefore, the coordinates of $F(v)$ relative to $\mathcal{B}_W$ are obtained by taking the matrix-vector product between the matrix $A_{\mathcal{B}_V}^{\mathcal{B}_W}(F)$ and the coordinates of v relative to $\mathcal{B}_V$.

Example 11: Let V and W be vector spaces of dimension $\dim(V) = 2$ and $\dim(W) = 3$, respectively. We consider the following bases in V and W:

$$\mathcal{B}_V = \{v_1, v_2\}, \qquad \mathcal{B}_W = \{w_1, w_2, w_3\}. \tag{40}$$

Relative to such bases, suppose that F is defined as

$$\begin{cases} F(v_1) = w_1 - 2w_2 - w_3 \\ F(v_2) = w_1 + w_2 + w_3 \end{cases}. \tag{41}$$

Then the matrix representing F is

$$A_{\mathcal{B}_V}^{\mathcal{B}_W}(F) = \begin{bmatrix} 1 & 1 \\ -2 & 1 \\ -1 & 1 \end{bmatrix}. \tag{42}$$

If $v = x_1 v_1 + x_2 v_2$ is an arbitrary vector in V then

$$\begin{aligned} F(v) =& x_1 F(v_1) + x_2 F(v_2) \\ =& x_1(w_1 - 2w_2 - w_3) + x_2(w_1 + w_2 + w_3) \\ =& \underbrace{(x_1 + x_2)}_{y_1} w_1 + \underbrace{(x_2 - 2x_1)}_{y_2} w_2 + \underbrace{(x_2 - x_1)}_{y_3} w_3. \end{aligned} \tag{43}$$

Note that the coordinates of $F(v)$ relative to the basis $\mathcal{B}_W$, i.e., $\{y_1, y_2, y_3\}$ are given by the standard matrix-vector product

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{44}$$

[5]We know from Lecture 6 that such coordinates are uniquely defined by the basis.

Change of basis transformation Consider the following two bases in the vector space V

$$\mathcal{B}_1 = \{u_1, \ldots, u_n\}$$
$$\mathcal{B}_2 = \{v_1, \ldots, v_n\}$$

Obviously, we can express any element in $\mathcal{B}_1$ as a linear combination of elements in $\mathcal{B}_2$ and vice versa. For example,

$$\begin{cases} v_1 = \alpha_{11}u_1 + \cdots + \alpha_{n1}u_m \\ \quad\vdots \\ v_n = \alpha_{1n}u_1 + \cdots + \alpha_{nn}u_n \end{cases} . \tag{45}$$

The matrix associated with the linear transformation "change of basis *from* $\mathcal{B}_2$ *to* $\mathcal{B}_1$" is

$$M_{\mathcal{B}_2}^{\mathcal{B}_1} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}. \tag{46}$$

Such a matrix is invertible and it allows us to transform the coordinates of any vector $v \in V$ from those relative to $\mathcal{B}_1$ to those relative to $\mathcal{B}_2$, i.e.,

$$[v]_{\mathcal{B}_2} = M_{\mathcal{B}_1}^{\mathcal{B}_2}[v]_{\mathcal{B}_1}. \tag{47}$$

Moreover, we have

$$[v]_{\mathcal{B}_1} = M_{\mathcal{B}_2}^{\mathcal{B}_1}[v]_{\mathcal{B}_2} = \left(M_{\mathcal{B}_1}^{\mathcal{B}_2}\right)^{-1}[v]_{\mathcal{B}_2} \quad \text{which implies} \quad M_{\mathcal{B}_2}^{\mathcal{B}_1} = \left(M_{\mathcal{B}_1}^{\mathcal{B}_2}\right)^{-1}. \tag{48}$$

The change of basis transformation can be also used to represent a linear transformation $F : V \to W$ relative to different bases in V and W. To show this, let

$$\begin{aligned} \mathcal{B}_1, \mathcal{B}_2 \quad &\to \quad \text{Bases of } V, \quad \dim(V) = n, \\ \mathcal{B}_3, \mathcal{B}_4 \quad &\to \quad \text{Bases of } W, \quad \dim(W) = m. \end{aligned}$$

We have,

$$[F(v)]_{\mathcal{B}_4} = M_{\mathcal{B}_3}^{\mathcal{B}_4}[F(v)]_{\mathcal{B}_3} = M_{\mathcal{B}_3}^{\mathcal{B}_4}A_{\mathcal{B}_2}^{\mathcal{B}_3}[v]_{\mathcal{B}_2} = \underbrace{M_{\mathcal{B}_3}^{\mathcal{B}_4}A_{\mathcal{B}_2}^{\mathcal{B}_3}M_{\mathcal{B}_1}^{\mathcal{B}_2}}_{A_{\mathcal{B}_1}^{\mathcal{B}_4}}[v]_{\mathcal{B}_1}, \tag{49}$$

i.e.,

$$A_{\mathcal{B}_1}^{\mathcal{B}_4} = M_{\mathcal{B}_3}^{\mathcal{B}_4}A_{\mathcal{B}_2}^{\mathcal{B}_3}M_{\mathcal{B}_1}^{\mathcal{B}_2}. \tag{50}$$

The matrix $A_{\mathcal{B}_1}^{\mathcal{B}_4}$ represents the linear transformation F relative to the bases $\mathcal{B}_1$ (basis of V) and $\mathcal{B}_4$ (basis of W). Similarly, $A_{\mathcal{B}_2}^{\mathcal{B}_3}$ represents the linear transformation F relative to the bases $\mathcal{B}_2$ (basis of V) and $\mathcal{B}_3$ (basis of W).

Example 12: (Change of basis in $\mathbb{R}^2$) Consider the following bases of $\mathbb{R}^2$

$$\mathcal{B}_1 = \{e_1, e_2\}, \qquad e_1 = \begin{bmatrix}1\\0\end{bmatrix}, \quad e_2 = \begin{bmatrix}0\\1\end{bmatrix} \quad \text{(canonical basis of } \mathbb{R}^2\text{)},$$

$$\mathcal{B}_2 = \{v_1, v_2\}, \qquad v_1 = \begin{bmatrix}1\\1\end{bmatrix}, \quad v_2 = \begin{bmatrix}1\\2\end{bmatrix}.$$

Define the change of basis transformation

$$F : \mathbb{R}^2 \to \mathbb{R}^2 \tag{51}$$

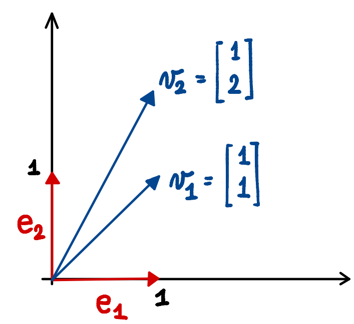
as

$$F\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\1\end{bmatrix}, \qquad F\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}1\\2\end{bmatrix}. \tag{52}$$

Clearly,

$$\begin{cases} v_1 = e_1 + e_2 \\ v_2 = e_1 + 2e_2 \end{cases}. \tag{53}$$

The following figure sketches $\{e_1, e_2\}$ and $\{v_1, v_2\}$ as vectors in the Cartesian plane.



Any vector $v \in \mathbb{R}^2$ can be expressed relatively to $\mathcal{B}_1$ or $\mathcal{B}_2$:

$$\begin{aligned} v =& x_1 v_1 + x_2 v_2 \\ =& x_1(e_1 + e_2) + x_2(e_1 + 2e_2) \\ =& (x_1 + x_2)e_1 + (x_1 + 2x_2)e_2. \end{aligned} \tag{54}$$

Denote by

$$[v]_{\mathcal{B}_1} = \begin{bmatrix}y_1\\y_2\end{bmatrix}, \qquad [v]_{\mathcal{B}_2} = \begin{bmatrix}x_1\\x_2\end{bmatrix}. \tag{55}$$

the coordinates of v relative to $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively. Then equation (54) implies that

$$[v]_{\mathcal{B}_1} = M_{\mathcal{B}_2}^{\mathcal{B}_1}[v]_{\mathcal{B}_2}, \qquad \text{where} \qquad M_{\mathcal{B}_2}^{\mathcal{B}_1} = \begin{bmatrix}1 & 1\\1 & 2\end{bmatrix}. \tag{56}$$

$M_{\mathcal{B}_2}^{\mathcal{B}_1}$ is the matrix associated with the change of basis transformation $\mathcal{B}_2 \to \mathcal{B}_1$. Clearly, $M_{\mathcal{B}_2}^{\mathcal{B}_1}$ is invertible with inverse

$$M_{\mathcal{B}_1}^{\mathcal{B}_2} = \left(M_{\mathcal{B}_2}^{\mathcal{B}_1}\right)^{-1} = \begin{bmatrix}2 & -1\\-1 & 1\end{bmatrix} \tag{57}$$

$M_{\mathcal{B}_1}^{\mathcal{B}_2}$ is the matrix associated with the change of basis transformation $\mathcal{B}_1 \to \mathcal{B}_2$. Let us see if this is true. To this end, we consider the vector $v = e_1$ and compute the coordinates of this vector relative to $\mathcal{B}_2$. We have

$$[v]_{\mathcal{B}_1} = \begin{bmatrix}1\\0\end{bmatrix} \quad \Rightarrow \quad [v]_{\mathcal{B}_2} = \underbrace{\begin{bmatrix}2 & -1\\-1 & 1\end{bmatrix}}_{M_{\mathcal{B}_1}^{\mathcal{B}_2}}\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}2\\-1\end{bmatrix} \tag{58}$$

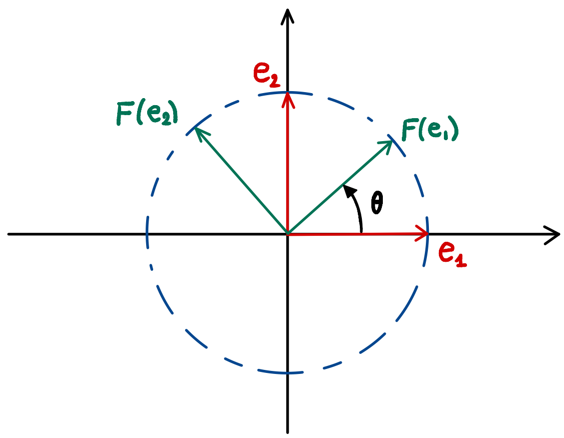
Example 13: (Rotations in $\mathbb{R}^2$) Consider the linear transformation $F : \mathbb{R}^2 \to \mathbb{R}^2$ defined as follows (counterclockwise rotation of the basis vectors by an angle θ)

$$\begin{cases} F(e_1) = \cos(\theta)e_1 + \sin(\theta)e_2 \\ F(e_2) = -\sin(\theta)e_1 + \cos(\theta)e_2 \end{cases}, \tag{59}$$

where

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{60}$$

is the canonical basis of $\mathbb{R}^2$.



The matrix associated with the transformation F relative to the basis $\mathcal{B}_V = \{e_1, e_2\}$ is

$$A^{\mathcal{B}_V}_{\mathcal{B}_V}(F) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \qquad \text{(2D rotation matrix).} \tag{61}$$

Any vector with components $[v]_{\mathcal{B}_V}$ is rotated to a vector $F(v)$ with components

$$[F(v)]_{\mathcal{B}_V} = A^{\mathcal{B}_V}_{\mathcal{B}_V}[v]_{\mathcal{B}_V}. \tag{62}$$

For example, the vector $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ has components $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ relative to the canonical basis $\mathcal{B}_V$, and it is transformed to a vector $F(v)$ with components

$$[F(v)]_{\mathcal{B}_V} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2\cos(\theta) - \sin(\theta) \\ 2\sin(\theta) + \cos(\theta) \end{bmatrix}. \tag{63}$$

In particular, if $\theta = \pi/2$ (90 degrees counterclockwise rotation) then

$$[F(v)]_{\mathcal{B}_V} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}. \tag{64}$$

The inverse transformation (inverse rotation) is obtained by replacing θ with $-\theta$ in (61), i.e.,

$$\left[A^{\mathcal{B}_V}_{\mathcal{B}_V}(F)\right]^{-1} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}. \tag{65}$$

It is straightforward to verify that for all $\theta \in [0, 2\pi]$ we have

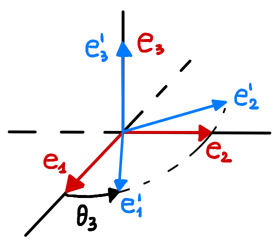
$$\left[A_{\mathcal{B}_V}^{\mathcal{B}_V}(F)\right]^{-1} A_{\mathcal{B}_V}^{\mathcal{B}_V}(F) = I_2. \tag{66}$$

In fact,

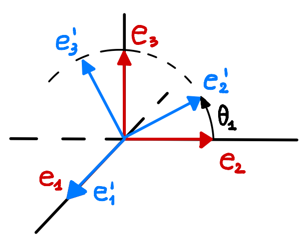
$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The rotation matrix is an *orthogonal matrix*[6].

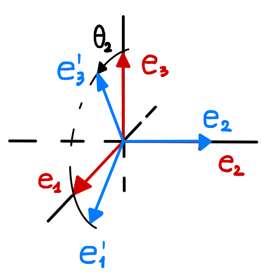
Example 14: (Rotations in $\mathbb{R}^3$) We can define rotations along each of the three axes of a 3D Cartesian coordinate system, i.e.,



$$R_3 = \begin{bmatrix} \cos\theta_3 & -\sin\theta_3 & 0 \\ \sin\theta_3 & \cos\theta_3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



$$R_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & -\sin\theta_1 \\ 0 & \sin\theta_1 & \cos\theta_1 \end{bmatrix}$$



$$R_2 = \begin{bmatrix} \cos\theta_2 & 0 & -\sin\theta_2 \\ 0 & 1 & 0 \\ \sin\theta_2 & 0 & \cos\theta_2 \end{bmatrix}$$

Note that the composition of two rotations in $\mathbb{R}^3$ does not commute. For example,

$$R_1 R_3 \neq R_3 R_1.$$

[6] In general, we say that $A \in M_{n\times n}(\mathbb{R})$ is orthogonal if

$$A^T = A^{-1}. \tag{67}$$

This is equivalent to the statement that orthogonal matrices satisfy

$$AA^T = I_n. \tag{68}$$

Example 15: (Orthogonal projection) Consider

$$F : \mathbb{R}^3 \to \mathbb{R}^3$$

and the canonical bases of $\mathbb{R}^3$

$$\mathcal{B}_3 = \{e_1, e_2, e_3\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

We define F by mapping the basis $\mathcal{B}_3$ as follows

$$F(e_1) = e_1, \qquad F(e_2) = e_2, \qquad F(e_3) = 0_{\mathbb{R}^3}.$$

The associated matrix defines an orthogonal projection onto the (x_1, x_2)

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{69}$$

Note that $P^2 = P$. The orthogonal projection transformation basically project any vector $v \in \mathbb{R}^3$ onto the plane spanned by e_1 and e_2. If we are interested in a projection onto different plane, we can use e.g., the 3D rotation matrices R_i and rotate the plane before applying the projection. Note that with just R_1 and R_3 we can orient the plane (x_1, x_2) in all possible directions. We maintain that

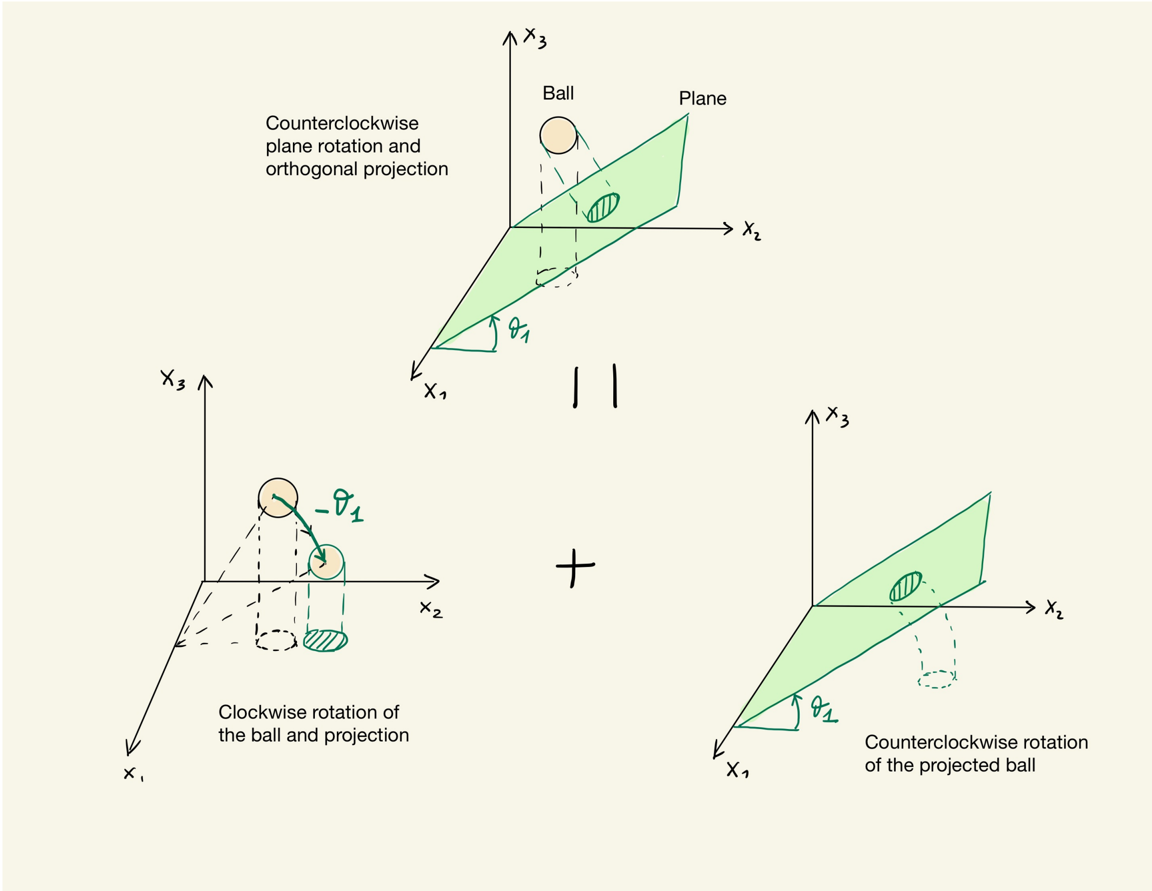
$$P(\theta_1, \theta_3) = R_1(\theta_1)R_3(\theta_3)PR_3^T(\theta_3)R_1^T(\theta_1) \tag{70}$$

is an orthogonal projection onto a tilted plane identified by the angles (θ_1, θ_3). To explain this formula suppose for simplicity that we just rotate the plane (x_1, x_2) counterclockwise of an angle θ_1 around the x_1 axis. The projection of any object onto such plane is obtained by rotating the object clockwise of an angle θ_1 around x_1 (matrix $R_1^T(\theta_3)$ projecting onto the (x_1, x_2) plane and then rotating the result back (matrix $R_1(\theta_3)$). Clearly, (70) satisfies the condition for orthogonal projections,

$$P^2(\theta_1, \theta_3) = P(\theta_1, \theta_3). \tag{71}$$

Example 16: (Oblique projection) Let $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ be a vector of $\mathbb{R}^3$ representing the direction of a light beam. A light beam passing through an arbitrary point $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ has the form

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = c \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \text{where } c \in \mathbb{R} \tag{72}$$



If we set $y_3 = 0$ we obtain $c = -x_3/v_3$. With such a value for c, the light beam passing through the point x intersects the horizontal plane. The linear transformation defined by

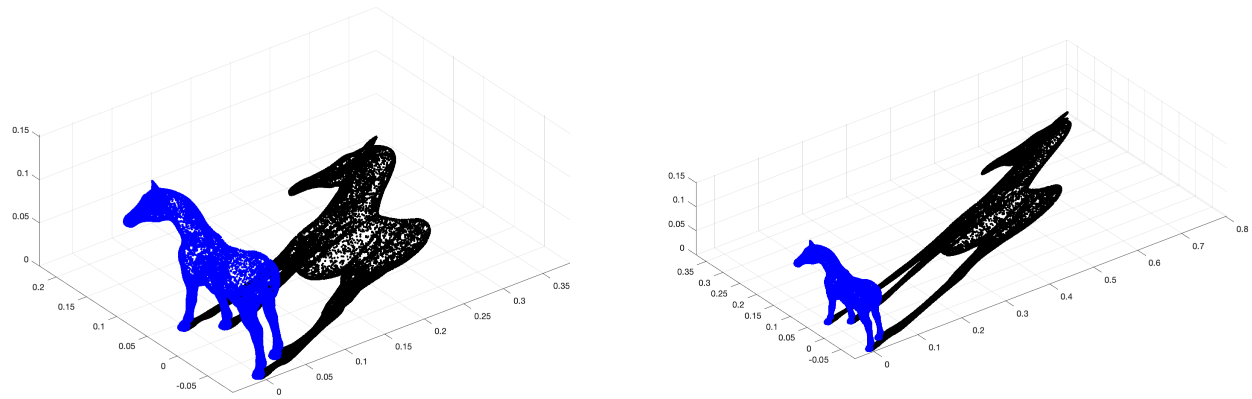
$$\begin{cases} y_1 = -\dfrac{v_1}{v_3}x_3 + x_1 \\ y_2 = -\dfrac{v_2}{v_3}x_3 + x_2 \\ y_3 = 0 \end{cases} \tag{73}$$

defines an oblique projection onto the horizontal plane. The matrix associated with such oblique projection transformation (relative to the canonical basis of $\mathbb{R}^3$) is

$$P = \begin{bmatrix} 1 & 0 & -v_1/v_3 \\ 0 & 1 & -v_2/v_3 \\ 0 & 0 & 0 \end{bmatrix} \tag{74}$$

The oblique projection can be used to compute the *shadow* of any object in 3D. The following figure shows the shadow projected by a horse for various angles of the light beam.

Note that for $v_1 = v_2 = 0$ the oblique projection reduces to the projection we studied in the previous example.



Example 17: Let $\mathbb{P}_4 = \text{span}\{1, x, x^2, x^3, x^4\}$ be the space of polynomials of degree at most 4. Define the linear transformation

$$\begin{aligned} F: \quad & \mathbb{P}_4 \to \mathbb{P}_3 \\ & p(x) \to \frac{dp(x)}{dx}. \end{aligned}$$

The canonical bases of $\mathbb{P}_4$ and $\mathbb{P}_3$ are

$$\begin{aligned} \mathcal{B}_4 &= \{1, x, x^2, x^3, x^4\}, \\ \mathcal{B}_3 &= \{1, x, x^2, x^3\}. \end{aligned}$$

We define the derivative transformation by mapping each element of $\mathbb{P}_4$ and representing the result in terms of $\mathbb{P}_3$. This yields

$$F(1) = 0, \qquad F(x) = 1, \qquad F(x^2) = 2x, \qquad F(x^3) = 3x^2, \qquad F(x^4) = 4x^3.$$

The matrix associated with F (derivative operator) relative to the bases $\mathcal{B}_4$ and $\mathcal{B}_3$ is

$$A_{\mathcal{B}_4}^{\mathcal{B}_3} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}.$$

For example, let us compute the derivative of the polynomial

$$p(x) = 1 - 3x + 6x^3. \tag{75}$$

The coordinates of $p(x)$ relative to $\mathcal{B}_4$ are

$$[p(x)]_{\mathcal{B}_4} = \begin{bmatrix} 1 & -3 & 0 & 6 & 0 \end{bmatrix}^T.$$

This implies that

$$\Rightarrow \left[\frac{dp(x)}{dx}\right]_{\mathcal{B}_3} = A_{\mathcal{B}_4}^{\mathcal{B}_3}[p(x)]_{\mathcal{B}_4} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 0 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 18 \\ 0 \end{bmatrix}.$$

Therefore we obtained

$$\frac{dp(x)}{dx} = -3 + 0x + 18x^2 + 0x^3 = -3 + 18x^2, \tag{76}$$

which is indeed the derivative of the polynomial (75).

Lecture 8: Scalar products, norms and orthogonality

Let U, V and W be three real vector spaces. We say that the transformation[1]

$$G : U \times V \mapsto W \tag{2}$$

is *bilinear* if for all $u_1, u_2 \in U$, all $v_1, v_2 \in V$, and all $c \in \mathbb{R}$

1. $G(u_1 + u_2, v_1) = G(u_1, v_1) + G(u_2, v_1)$,
2. $G(u_1, v_1 + v_2) = G(u_1, v_1) + G(u_1, v_2)$,
3. $G(cu_1, v_1) = G(u_1, cv_1) = cG(u_1, v_1)$.

If the bilinear transformation G is real-valued, i.e.,

$$G : U \times V \mapsto \mathbb{R}, \tag{3}$$

then we say that F is a *bilinear form.*

Example 1: Consider $U = \mathbb{R}^m$, $V = \mathbb{R}^n$ and a matrix $A \in M_{m\times n}(\mathbb{R})$. Then

$$\begin{aligned} G : \mathbb{R}^m \times \mathbb{R}^n &\to \mathbb{R} \\ (u, v) &\to \sum_{i=1}^{m}\sum_{j=1}^{n} u_i A_{ij} v_j \end{aligned}$$

is a bilinear form.

If $U = V$ then we say that $G : V \times V \mapsto \mathbb{R}$ is a bilinear form on V. Moreover, if for all $v_1, v_2 \in V$ we have that

$$G(v_1, v_2) = G(v_2, v_1) \tag{4}$$

then we say that the bilinear form on V is *symmetric.*

Matrix associated to a bilinear form. Similarly to linear transformations, it is possible to define the matrix associated to a bilinear form on V. To this end, suppose that V is n-dimensional and consider the basis $\mathcal{B}_V = \{v_1, \dots, v_n\}$. Let $u, v \in V$ such that

$$u = x_1 v_1 + \cdots + x_n v_n, \qquad v = y_1 v_1 + \cdots + y_n v_n. \tag{5}$$

The coordinates of u and v relative to $\mathcal{B}_V$ are

$$[u]_{B_V} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad [v]_{B_V} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}. \tag{6}$$

[1] The multiplication symbol $\times$ in (2) means "Cartesian product" of two sets. The elements of the set $U \times V$ are of pairs of vectors (u, v) where $u \in V$ and $v \in V$. We have already seen an example of a vector space constructed using multiple Cartesian products, i.e.,

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ times}}. \tag{1}$$

A substitution of u and v into $G(u,v)$ yields

$$G(u,v) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i G(v_i, v_j) y_j. \tag{7}$$

Define the matrix $A_{\mathcal{B}_V}$ associated with $G(u,v)$ relative to B_V as

$$A_{\mathcal{B}_V} = \begin{bmatrix} G(v_1,v_1) & \cdots & G(v_1,v_n) \\ \vdots & \ddots & \vdots \\ G(v_n,v_1) & \cdots & G(v_n,v_n) \end{bmatrix} \tag{8}$$

This allows us to write (7) as

$$G(u,v) = [u]_{B_V}^T A_{\mathcal{B}_V} [v]_{B_V}. \tag{9}$$

If G is symmetric then $A_{\mathcal{B}_V}$ is a symmetric matrix.

Scalar products. A scalar product on a real vector space V is a symmetric bilinear form on V. The scalar product between two vectors in V is a real number[2]. We denote such scalar product as

$$\langle u, v\rangle = G(u,v) \qquad \forall u, v \in V. \tag{10}$$

A scalar product in V is also called "inner product" in V.

Examples of scalar products:

1. $V = \mathbb{R}^n$: $\langle u, v\rangle = \sum_{i=1}^{n} u_i v_i$ (scalar product on $\mathbb{R}^n$). This scalar product is often called "dot-product" and denoted as $u \cdot v$. The matrix associated with the dot product is the identity matrix.

2. $V = \mathbb{R}^n$: $\langle u, v\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} u_i g_{ij} v_j$ where g_{ij} is a $n \times n$ symmetric matrix. In differential geometry and in the theory of general relativity g_{ij} is called *metric tensor* and it represents the metric properties of the space-time (the curvature of the space-time is a nonlinear function of g_{ij}).

3. $V = M_{n\times n}(\mathbb{R})$: $\langle A, B\rangle = \mathrm{Tr}(AB^T) = \sum_{i,j=1}^{n} A_{ij}B_{ij}$ for all $A, B \in M_{n\times n}(\mathbb{R})$ (scalar product between two matrices). Here $\mathrm{Tr}(AB^T)$ denotes the trace of the matrix AB^T, which is clearly a symmetric bilinear form[3].

[2] It is possible to define scalar products on complex vector spaces. In this setting, the symmetric bilinear form returns a complex number and it is called Hermitian bilinear form.

[3] In fact,

$$\mathrm{Tr}(AB^T) = \mathrm{Tr}(BA^T), \qquad \mathrm{Tr}((A+C)B^T) = \mathrm{Tr}(AB^T) + \mathrm{Tr}(CB^T) = \mathrm{Tr}(B(A+C)^T) \tag{11}$$

4. $V = C^0([0,1])$ (space of continuous functions defined on $[0,1]$). The integral

$$\langle u, v\rangle = \int_0^1 u(x)v(x)dx$$

is a symmetric bilinear form on $C^0([0,1])$ which defines a scalar product. In particular, if u and v are two polynomials of degree at most n defined on $[0,1]$ then $\langle u, v\rangle$ is a scalar product on $\mathbb{P}_n$.

A scalar product on $\langle \cdot, \cdot\rangle : V \times V \to \mathbb{R}$ is said to be *non-degenerate* if

$$\langle v, w\rangle = 0 \qquad \text{for all } w \quad \Rightarrow \quad v = 0_V. \tag{12}$$

Theorem 1. Let V be a real vector space of dimension n. A scalar product

$$\langle \cdot, \cdot\rangle : V \times V \to \mathbb{R}$$

is non-degenerate if and only if the matrix associated with $\langle \cdot, \cdot\rangle$ relative to any basis $\mathcal{B}_V = \{v_1, \dots, v_n\}$ is invertible.

Proof. We know that for all $u, v \in V$

$$\langle u, v\rangle = [u]^T_{\mathcal{B}_V} A_{\mathcal{B}_V} [v]_{\mathcal{B}_V}. \tag{13}$$

If the inner product is non-degenerate then, by definition

$$[u]^T_{\mathcal{B}_V} A_{\mathcal{B}_V} [v]_{\mathcal{B}_V} = 0 \qquad \text{for all} \quad [u]_{\mathcal{B}_V} \quad \Rightarrow \quad [v]_{\mathcal{B}_V} = 0_{\mathbb{R}^n}. \tag{14}$$

This implies that the nullspace of $A_{\mathcal{B}_V}$ reduces to the singleton $\{0_{\mathbb{R}^n}\}$. In fact if there exists another nonzero vector $[w]$ in the nullspace of $A_{\mathcal{B}_V}$ then clearly $A_{\mathcal{B}_V}[w] = 0_{\mathbb{R}^n}$ and the implication in (14) is not true. Conversely, suppose that $A_{\mathcal{B}_V}$ is invertible. Consider the column vector $A_{\mathcal{B}_V}[v]_{\mathcal{B}_V}$ and take all "dot products" with the elements the canonical basis of $\mathbb{R}^n$. This yields the system

$$[e_k]^T A_{\mathcal{B}_V} [v]_{\mathcal{B}_V} = 0 \qquad k = 1, \dots, n, \tag{15}$$

which is a homogeneous linear system of equations in n unknowns $[v]_{\mathcal{B}_V}$ that can be written as

$$A_{\mathcal{B}_V} [v]_{\mathcal{B}_V} = 0_{\mathbb{R}^n}. \tag{16}$$

The solution to this system is clearly $[v]_{\mathcal{B}_V} = 0_{\mathbb{R}^n}$, since $A_{\mathcal{B}_V}$ is invertible by assumption.

$\square$

Examples of non-degenerate scalar products: The scalar products 1., 2. (with g_{ij} invertible) and 3. at page 2, and 4. at page 3 are all non-degenerate scalar products. Let us show that 3. is indeed a non-degenerate scalar product on $M_{n\times n}(\mathbb{R})$. We need to show that

$$\mathrm{Tr}(AB^T) = 0 \quad \text{for all} \quad B \in M_{n\times n}(\mathbb{R}) \quad \text{implies} \quad A = 0_{M_{n\times n}}. \tag{17}$$

The trace of the matrix product can be expressed as

$$\mathrm{Tr}(AB^T) = \sum_{i,j=1}^{n} A_{ij} B_{ij}. \tag{18}$$

Since B is arbitrary it easily follows form $\mathrm{Tr}(AB^T) = 0$ that $A = 0_{M_{n\times n}}$. In fact, evaluate the equation above using B equal to each element of the canonical basis of $M_{n\times n}(\mathbb{R})$. This yields $A_{ij} = 0$.

Examples of degenerate scalar products: Hereafter we provide a few examples of degenerate scalar products.

1. Let $V = \mathbb{R}^2$. Consider two vectors

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{19}$$

the scalar product

$$\langle x, y \rangle = x_1 y_1 \tag{20}$$

is degenerate. In fact, the condition

$$\langle x, y \rangle = 0 \quad \text{for all} \quad y \in \mathbb{R}^2 \quad \text{does not imply} \quad x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{21}$$

To see this simply consider the vector $x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Alternatively, we observe that the scalar product (20) can be written as

$$\langle x, y \rangle = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{A_{\mathcal{B}_V}} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{22}$$

Recalling Theorem 2, we see that the matrix $A_{\mathcal{B}_V}$ associated with the scalar product relative to the canonical basis of $\mathbb{R}^2$ is not invertible and therefore the scalar product is degenerate.

Positive definite scalar products. Let V be a real vector space. A scalar product on V is said to be *positive definite* if

$$\langle v, v \rangle > 0 \quad \text{for all nonzero} \quad v \in V. \tag{23}$$

Clearly, if $v = 0_V$ then $\langle v, v \rangle = 0$.

Examples of positive definite scalar products: The scalar products 1., 3. and 4. defined at page 2 are all non-degenerate and positive definite. The scalar product 2. at page 2 is non-degenerate and positive definite if and only if the matrix g_{ij} is positive definite, i.e., if

$$\sum_{i,j=1}^{n} g_{ij} x_i x_j > 0 \quad \text{for all nozero vectors} \quad x \in \mathbb{R}^n. \tag{24}$$

Positive definite matrices are necessarily invertible. In fact, from condition (24) it follows that for any nonzero vector x, gx is nonzero. Therefore g is full rank ($\Rightarrow$ nullspace reduces to the $0_{\mathbb{R}^n}$) and therefore invertible.

Norms. A norm on a vector space V is is a function $\|\cdot\| : V \to \mathbb{R}$ with the following properties

1. $\|av\| = |a| \|v\|$ for all $v \in V$ and for all $a \in \mathbb{R}$ (or $\mathbb{C}$).
2. $\|u + v\| \leq \|u\| + \|v\|$
3. $\|u\| = 0 \quad \Leftrightarrow \quad u = 0_V$
4. $\|u\| > 0 \quad$ for all nonzero $u \in V$.

The norm defines the *length* of vectors in a vector space. We have already seen a norm when we studied complex numbers, i.e., the modulus of a complex number. Let us provide a few examples of norms.

Examples:

- Let $V = \mathbb{R}^n$. For every $v \in \mathbb{R}^n$ we define

$$\|v\|_1 = \sum_{i=1}^{n} |v_i| \qquad \text{(1-norm)} \tag{25}$$

$$\|v\|_2 = \left(\sum_{i=1}^{n} |v_i|^2\right)^{1/2} \qquad \text{(2-norm)} \tag{26}$$

$$\|v\|_p = \left(\sum_{i=1}^{n} |v_i|^p\right)^{1/p} \qquad \text{(p-norm, $p \geq 1$ real number)} \tag{27}$$

$$\|v\|_\infty = \max_{i=1,\dots,n} |v_i| \qquad \text{(infinity norm)} \tag{28}$$

  All these norms satisfy properties 1.-4. above. Moreover, it can be shown that

$$\|v\|_\infty = \lim_{p\to\infty} \|v\|_p . \tag{29}$$

- Let $V = C^0([0,1])$ (space of continuous functions in $[0,1]$). We define

$$\|u\|_\infty = \max_{x\in[0,1]} |u(x)| \qquad \text{(uniform norm)}, \tag{30}$$

$$\|u\|_2 = \int_0^1 |u(x)|^2 \, dx \qquad \text{($L^2([0,1])$-norm)}. \tag{31}$$

- Let $V = M_{n\times n}(\mathbb{R})$ (space of $n \times n$ matrices with real coefficients). Let us define the following matrix norm

$$\|A\| = \max_{v\neq 0_{\mathbb{R}^n}} \frac{\|Av\|_p}{\|v\|_p} = \max_{\|v\|_p=1} \|Av\|_p \qquad p \geq 1. \tag{32}$$

It is straightforward to show that

$$\|A\|_\infty = \max_{i=1,..,n}\left(\sum_{j=1}^{n}|A_{ij}|\right), \tag{33}$$

$$\|A\|_1 = \max_{j=1,..,n}\left(\sum_{i=1}^{n}|A_{ij}|\right). \tag{34}$$

For example,

$$\|Av\|_\infty = \max_{i=1,...,n}\left|\sum_{j=1}^{n}A_{ij}v_j\right| \le \max_{i=1,...,n}\left(\sum_{j=1}^{n}|A_{ij}|\,|v_j|\right) \le \|v\|_\infty \max_{i=1,...,n}\left(\sum_{j=1}^{n}|A_{ij}|\right) \tag{35}$$

which implies that

$$\frac{\|Av\|_\infty}{\|v\|_\infty} \le \max_{i=1,...,n}\left(\sum_{j=1}^{n}|A_{ij}|\right) \qquad \text{for all } v \neq 0_{\mathbb{R}^n}, \tag{36}$$

i.e.,

$$\max_{v\neq 0_{\mathbb{R}^n}} \frac{\|Av\|_\infty}{\|v\|_\infty} = \max_{i=1,...,n}\left(\sum_{j=1}^{n}|A_{ij}|\right) = \|A\|_\infty. \tag{37}$$

The matrix norms (33) and (34) are said to be *compatible* with associated vector norms (or *induced* by the vector norms) since they verify the inequalities

$$\|Av\|_p \le \|A\|_p \|v\|_p \qquad p = 1, \infty. \tag{38}$$

Norms induced by scalar products. Any non-degenerate positive-definite scalar product on V induces a norm[4]

$$\|v\| = \sqrt{\langle v, v\rangle}. \tag{39}$$

In particular, the standard dot product in $\mathbb{R}^n$

$$\langle v, v\rangle = \sum_{i=1}^{n} v_i^2$$

induces the 2-norm defined in (26). Similarly, the scalar product between two matrices $A, B \in M_{n\times n}(\mathbb{R})$

$$\langle A, B\rangle = \mathrm{Tr}\left(AB^T\right)$$

induces the following norm in the space of matrices $M_{n\times n}(\mathbb{R})$

$$\|A\|_F = \sqrt{\sum_{i,j=1}^{n} A_{ij}^2} \qquad \text{(Frobenious norm)}. \tag{40}$$

The Frobenious norm is compatible with the vector norm in the sense that $\|Ax\|_2 \le \|A\|_F \|x\|_2$.

[4]It is straightforward to show that properties 1.-4. at page 5 are all satisfied by the norm (39).

Theorem 2 (Cauchy-Schwarz inequality)**.** Let V be a real vector space. Then for all $u, v \in V$ we have

$$|\langle u, v\rangle| \leq \|u\| \, \|v\| \,, \tag{41}$$

where $\|u\| = \sqrt{\langle u, u\rangle}$ and $\|v\| = \sqrt{\langle v, v\rangle}$.

Proof. For $v = 0_V$ the inequality reduces to $0 = 0$. Let $u, v \in V$ be nonzero. The vector

$$u - \frac{\langle u, v\rangle}{\langle v, v\rangle} v \tag{42}$$

is orthogonal[5] to v since

$$\left\langle v, u - \frac{\langle u, v\rangle}{\langle v, v\rangle} v \right\rangle = \langle v, u\rangle - \frac{\langle u, v\rangle}{\langle v, v\rangle} \langle v, v\rangle = 0. \tag{43}$$

Next, consider the identity

$$u = u - \frac{\langle u, v\rangle}{\langle v, v\rangle} v + \frac{\langle u, v\rangle}{\langle v, v\rangle} v. \tag{44}$$

We have

$$\begin{aligned}
\|u\|^2 &= \langle u, u\rangle \\
&= \left\langle u - \frac{\langle u, v\rangle}{\langle v, v\rangle} v, u - \frac{\langle u, v\rangle}{\langle v, v\rangle} v \right\rangle + \frac{|\langle u, v\rangle|^2}{|\langle v, v\rangle|^2} \langle v, v\rangle \\
&= \left\| u - \frac{\langle u, v\rangle}{\langle v, v\rangle} v \right\|^2 + \frac{|\langle u, v\rangle|^2}{\|v\|^2} \\
&\geq \frac{|\langle u, v\rangle|^2}{\|v\|^2}.
\end{aligned} \tag{45}$$

From the last inequality we have $|\langle u, v\rangle|^2 \leq \|u\|^2 \|v\|^2$. Taking the square root yields equation (41).

□

Cosine similarity. The Cauchy-Schwartz inequality (41) implies that

$$-1 \leq \frac{\langle u, v\rangle}{\|u\| \, \|v\|} \leq 1. \tag{46}$$

The quantity

$$\cos(\vartheta) = \frac{\langle u, v\rangle}{\|u\| \, \|v\|} \tag{47}$$

is known as *cosine similarity* between the vectors u and v. In the case where u and v are vectors of $\mathbb{R}^2$, $\mathbb{R}^3$ or $\mathbb{R}^n$, the cosine similarity coincides with cosine of the angle between the two vectors. Such an angle is measures on the two-dimensional plane spanned by the two vectors. From (47) it follows that

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2 \|u\| \, \|v\| \cos(\vartheta). \tag{48}$$

which is the well-known law of cosines for triangles. The cosine similarity is practically utilized in many different fields, e.g., natural language processing (similarity between texts) and econometrics (analysis of time series).

[5] We say that to vectors $u, w \in V$ are orthogonal with respect to the scalar product $\langle \cdot, \cdot \rangle$ if $\langle u, w\rangle = 0$.

Orthogonality. Consider a vector space V and a non-degenerate positive definite scalar product $\langle \cdot, \cdot \rangle$ on V. Two vectors $u, v \in V$ are said to be *orthogonal* relative to $\langle \cdot, \cdot \rangle$ if

$$\langle u, v \rangle = 0. \tag{49}$$

Examples:

- $V = \mathbb{R}^2$. The following vectors

$$u = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad v = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \tag{50}$$

are orthogonal in $\mathbb{R}^2$ relative to the standard inner product

$$\langle u, v \rangle = \sum_{i=1}^{2} u_i v_i = -1 + 1 = 0. \tag{51}$$

- $V = M_{2\times 2}(\mathbb{R})$. The following matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{52}$$

are orthogonal in $M_{2\times 2}(\mathbb{R})$ relative to the inner product

$$\langle A, B \rangle = \mathrm{Tr}\left(AB^T\right). \tag{53}$$

In fact,

$$\mathrm{Tr}\left(AB^T\right) = \mathrm{Tr}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = \mathrm{Tr}\left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right) = 0. \tag{54}$$

- $V = \mathbb{P}_2([-1, 1])$ (vector space of polynomials of degree at most two). The polynomials

$$p_1(x) = x \qquad \text{and} \qquad p_2(x) = \frac{3}{2}x^2 - \frac{1}{2} \tag{55}$$

are orthogonal with respect to the scalar product

$$\langle p_1, p_2 \rangle = \int_{-1}^{1} p_1(x) p_2(x) dx. \tag{56}$$

In fact,

$$\langle p_1, p_2 \rangle = \int_{-1}^{1} p_1(x) p_2(x) dx = \int_{-1}^{1} \left(\frac{3}{2}x^3 - \frac{1}{2}x\right) dx = \left[\frac{3}{8}x^4 - \frac{1}{4}x^2\right]_{-1}^{1} = 0. \tag{57}$$

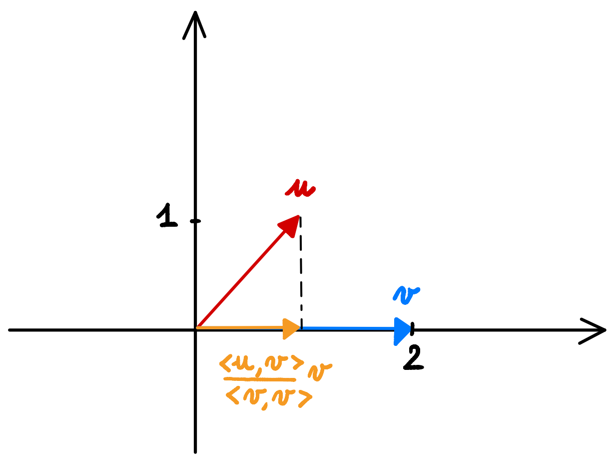
Orthogonal projections. Consider two vectors u and v in a vector space V. The orthogonal projection of u onto v is defined as

$$P_v u = \frac{\langle u, v \rangle}{\langle v, v \rangle} v = \left\langle u, \frac{v}{\|v\|} \right\rangle \frac{v}{\|v\|}, \tag{58}$$

where $\|\cdot\|$ is the norm induced by the scalar product. Clearly $v/\|v\|$ is a vector with norm equal to one, i.e., a *unit vector*.

Examples:

- Let $V = \mathbb{R}^2$ and consider the following vectors
$$u = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad v = \begin{bmatrix} 2 \\ 0 \end{bmatrix}. \tag{59}$$



The projection of u onto v is
$$P_v u = \frac{\langle u, v\rangle}{\langle v, v\rangle} v = \frac{2}{4}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{60}$$
Note that if we subtract $\dfrac{\langle u, v\rangle}{\langle v, v\rangle} v$ from u we obtain a vector that is orthogonal to v.
$$u - \frac{\langle u, v\rangle}{\langle v, v\rangle} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{61}$$

- $V = \mathbb{R}^3$. Given three vectors $v_1, v_2, v_3 \in \mathbb{R}^3$ we can compute the orthogonal projection of any vector onto any other vector, e.g., the orthogonal projection of v_2 onto v_1
$$P_{v_1} v_2 = \frac{\langle v_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1. \tag{62}$$
We can also construct an orthogonal set of vector by transforming the given set of linearly independent vectors $\{v_1, v_2, v_3\}$ as follows
$$\begin{aligned} u_1 &= v_1, \\ u_2 &= v_2 - \frac{\langle v_2, u_1\rangle}{\langle u_1, u_1\rangle} u_1, \\ u_3 &= v_3 - \frac{\langle v_3, u_1\rangle}{\langle u_1, u_1\rangle} u_1 - \frac{\langle v_3, u_2\rangle}{\langle u_2, u_2\rangle} u_2. \end{aligned}$$

This procedure is known as *Gram-Schmidt orthogonalization*, and it allows us to transform any set of linearly independent vectors into an orthogonal one. Such set of ortogonal vectors can be then normalized.

Gram-Schmidt orthogonalization. The previous example suggests that we can transform any basis $\{v_1, \dots, v_n\}$ of a n-dimensional vector space V into an orthonormal basis[6] by using the Gram-Schmidt procedure. In fact, we can first compute

$$
\begin{aligned}
u_1 =& v_1,\\
u_2 =& v_2 - \frac{\langle v_2, u_1\rangle}{\langle u_1, u_1\rangle} u_1,\\
u_3 =& v_3 - \frac{\langle v_3, u_1\rangle}{\langle u_1, u_1\rangle} u_1 - \frac{\langle v_3, u_2\rangle}{\langle u_2, u_2\rangle} u_2,\\
&\vdots\\
u_n =& v_n - \frac{\langle v_n, u_1\rangle}{\langle u_1, u_1\rangle} u_1 - \cdots \frac{\langle v_n, u_{n-1}\rangle}{\langle u_{n-1}, u_{n-1}\rangle} u_{n-1}.
\end{aligned}
$$

and then normalize the vectors $\{u_1, \dots, u_n\}$ to obtain the orthonormal basis

$$
\left\{ \frac{u_1}{\|u_1\|}, \dots, \frac{u_n}{\|u_n\|} \right\}. \tag{63}
$$

Alternatively, we can normalize each vector u_i right after we compute it. This reduces the number of calculations in the Gram-Schmidt procedure as we can write

$$
\begin{aligned}
u_1 &= v_1, & \widehat{u}_1 &= u_1/\|u_1\|,\\
u_2 &= v_2 - \langle v_2, \widehat{u}_1\rangle \widehat{u}_1 & \widehat{u}_2 &= u_2/\|u_2\|,\\
u_3 &= v_3 - \langle v_3, \widehat{u}_1\rangle \widehat{u}_1 - \langle v_3, \widehat{u}_2\rangle \widehat{u}_2 & \widehat{u}_3 &= u_3/\|u_3\|,\\
&\cdots
\end{aligned}
$$

It is straightforward to show that

$$
\langle u_i, u_j\rangle = \delta_{ij} \|u_j\|^2, \tag{64}
$$

where δ_{ij} is the Kronecker delta function[7]. For example,

$$
\langle u_1, u_2\rangle = \left\langle v_1, v_2 - \frac{\langle v_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 \right\rangle = \langle v_1, v_2\rangle - \frac{\langle v_2, v_1\rangle}{\langle v_1, v_1\rangle} \langle v_1, v_1\rangle = 0. \tag{66}
$$

Example: Let us use the Gram-Schmidt procedure to orthogonalize the following vectors in $\mathbb{R}^2$

$$
v_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.
$$

[6]Note that the orthogonal basis we obtain from the Gram-Schmidt procedure is not unique. In fact a reordering of the vectors $\{v_1, \dots, v_n\}$ yields a different orthogonal basis at the end of the procedure.

[7]The Kronecher delta is defined as:

$$
\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{65}
$$

We have

$$u_1 = v_1, \qquad u_2 = v_2 - \frac{\langle v_2, u_1\rangle}{||u_1||^2} u_1.$$

The norm of $u_1 = v_1$ is

$$\|u_1\|^2 = \|v_1\|^2 = \langle v_1, v_1\rangle = 2^2 + 1^2 = 5. \tag{67}$$

This implies that

$$u_2 = \begin{bmatrix}1\\2\end{bmatrix} - \frac{4}{5}\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}1-8/5\\2-4/5\end{bmatrix} = \begin{bmatrix}-3/5\\6/5\end{bmatrix}$$

Note that u_1 and u_2 are orthogonal. In fact,

$$\langle u_1, u_2\rangle = 2 \times \left(-\frac{3}{5}\right) + \frac{6}{5} = 0. \tag{68}$$

The norm of u_2 is

$$\|u_2\| = \sqrt{\langle u_2, u_2\rangle} = \sqrt{\frac{9}{25} + \frac{36}{25}} = \frac{3\sqrt{5}}{5}. \tag{69}$$

This means that

$$\left\{\frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|}\right\} = \left\{\frac{\sqrt{5}}{5}\begin{bmatrix}2\\1\end{bmatrix}, \frac{\sqrt{5}}{5}\begin{bmatrix}-1\\2\end{bmatrix}\right\} \tag{70}$$

is an orthonormal basis of $\mathbb{R}^2$. Note that $u_2/\|u_2\|$ can be obtained by rotating $u_1/\|u_1\|$ by 90 degrees counterclockwise.

Representation of vectors relative to orthonormal bases. Let $\mathcal{B}_V = \{\widehat{u}_1, \dots, \widehat{u}_n\}$ be an orthonormal basis of a n-dimensional vector space V. Any vector $v \in V$ can be represented relative to the basis $\mathcal{B}_V$ as

$$v = x_1\widehat{u}_1 + \cdots + x_n\widehat{u}_n. \tag{71}$$

by projecting the vector v onto $\widehat{u}_i$ and taking into account the orthonormality conditions $\langle u_i, u_j\rangle = \delta_{ij}$ yields

$$\begin{aligned}\langle v, \widehat{u}_j\rangle =& \langle x_1\widehat{u}_1 + \cdots + x_n\widehat{u}_n, \widehat{u}_j\rangle\\ =& x_1\langle \widehat{u}_1, \widehat{u}_j\rangle + \cdots + x_n\langle \widehat{u}_n, \widehat{u}_j\rangle\\ =& x_j\langle \widehat{u}_j, \widehat{u}_j\rangle\\ =& x_j\end{aligned} \tag{72}$$

i.e., the j-the coordinate of v relative to $\mathcal{B}_V$ coincides with the projection of v onto $\widehat{u}_j$. On the other hand, if we consider an orthogonal basis $\{u_1, \dots, u_n\}$ we obtain

$$v = y_1u_1 + \cdots + y_nu_n \quad \Rightarrow \quad x_j = \frac{\langle v, u_j\rangle}{\langle u_j, u_j\rangle}. \tag{73}$$

Theorem 3. Let $\mathcal{B}_V = \{\widehat{u}_1, \dots, \widehat{u}_n\}$ be an orthonormal basis of a n-dimensional vector space V. Then for any vector $v \in V$ we have

$$v = x_1\widehat{u}_1 + \cdots + x_n\widehat{u}_n \quad \text{and} \quad \|v\|^2 = \sum_{k=1}^{n} x_k^2. \tag{74}$$

Orthogonal complement. Let S be a subspace of V. The orthogonal complement of S in V is defined as

$$S^\perp = \{v \in V : \langle v, w\rangle = 0 \quad \text{for all} \quad w \in S\}$$

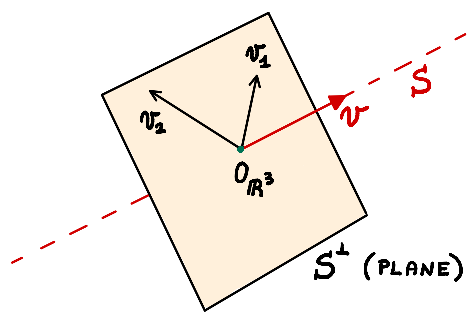
It can be shown that $S^\perp$ is a vector subspace of V. Any vector $v \in V$ can be expressed as a sum of two vectors $w_1 \in S$ and $w_2 \in S^\perp$, i.e.,

$$v = w_1 + w_2 \tag{75}$$

Equivalently, we say that V is the direct sum of S and $S^\perp$, and write

$$V = S \oplus S^\perp. \tag{76}$$

For example, any vector $v \in V = \mathbb{R}^3$ defines a one-dimensional vector subspace S. The orthogonal complement of S in $\mathbb{R}^3$ is a plane orthogonal S. Such plane is denoted by $S^\perp$



The plane, i.e., the vector space $S^\perp$, is identified mathematically by the condition

$$\langle x, v\rangle = 0 \qquad (v \text{ is given, } x \in \mathbb{R}^3 \text{ is arbitrary}) \tag{77}$$

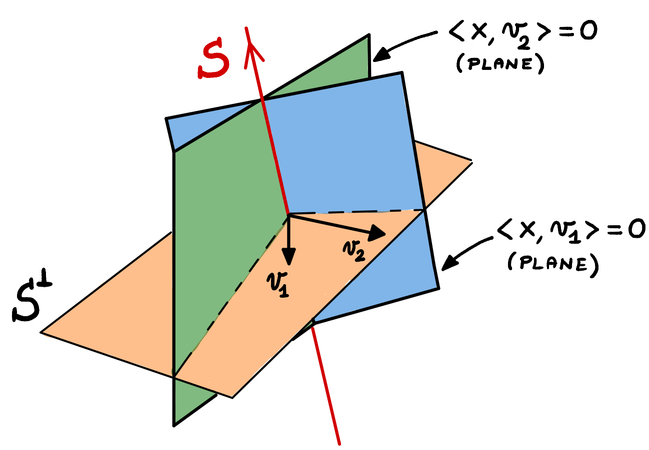
i.e.,

$$v_1x_1 + v_2x_2 + v_3x_3 = 0 \tag{78}$$

We know this expression very well, but now we learned something new, i.e., that the coefficients of v_1, v_2 and v_3 are the components of a vector that is orthogonal to the plane. Similarly, given two linearly independent vectors vectors v_1 and v_2 in $\mathbb{R}^3$, it is possible to determine the space that is orthogonal to the span of v_1 and v_2 by solving the system of equations

$$\langle x, v_1\rangle = 0, \qquad \langle x, v_2\rangle = 0 \qquad x \in \mathbb{R}^3. \tag{79}$$

This system represents the intersection of two planes orthogonal to v_1 and v_2.



Orthogonal complement of the range and the nullspace of a matrix. Next consider a $m \times n$ matrix A and let $\{v_1, \ldots, v_n\}$ be the columns of A. Denote by

$$R(A) = \text{span}\{v_1, \ldots, v_n\} \tag{80}$$

the column space of A, i.e., the range of the matrix A. We known that such space is a vector subspace of $\mathbb{R}^m$. The orthogonal complement of $R(A)$ is

$$[R(A)]^{\perp} = \{v \in \mathbb{R}^m : \langle v, w \rangle = 0 \quad \text{for all} \quad w \in R(A)\}. \tag{81}$$

Let us write the condition $\langle v, w \rangle = 0$ a more explicitly.

To this end, we notice that $R(A)$ can be characterized as the set of vectors $w \in \mathbb{R}^m$ such that such that $w = Ax$. Hence,

$$\begin{aligned} v \in [R(A)]^{\perp} \quad &\Leftrightarrow \quad \langle v, Ax \rangle = 0 \quad \text{for all } x \in \mathbb{R}^n \\ &\Leftrightarrow \quad \left\langle A^T v, x \right\rangle = 0 \quad \text{for all } x \in \mathbb{R}^n \\ &\Leftrightarrow \quad A^T v = 0_{\mathbb{R}^n} \\ &\Leftrightarrow \quad v \in N(A^T). \end{aligned}$$

This means that

$$[R(A)]^{\perp} = N(A^T). \tag{82}$$

In other words, the orthogonal complement of the column space of a matrix coincides with the nullspace of the matrix transpose. We can also prove the equality the other way around, i.e.,

$$\begin{aligned} v \in N(A^T) \quad &\Leftrightarrow \quad A^T v = 0 \qquad v \in \mathbb{R}^m \\ &\Leftrightarrow \quad \left\langle w, A^T v \right\rangle = 0 \quad \text{for all } w \in \mathbb{R}^n \\ &\Leftrightarrow \quad \langle Aw, v \rangle = 0 \quad \text{for all } w \in \mathbb{R}^n \\ &\Leftrightarrow \quad v \in [R(A)]^{\perp}. \end{aligned}$$

Repeating this simple proof for $N(A)$ yields

$$\left[R(A^T)\right]^{\perp} = N(A). \tag{83}$$

i.e., the orthogonal complement of the *row space* of A (i.e., the column space of A^T) coincides with the nullspace of A. Similarly, it can be shown that

$$[N(A)]^{\perp} = R(A^T) \qquad \text{and} \qquad \left[N(A^T)\right]^{\perp} = R(A). \tag{84}$$

Lecture 9: Determinants

Let $A \in M_{n\times n}$ be a square matrix with real or complex entries

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \tag{1}$$

The determinant of A is the real or complex number

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A^{ij}) \qquad \text{(for every fixed } i\text{)}, \tag{2}$$

where a_{ij} are the entries of A, and A^{ij} is a matrix obtained from A by crossing out the i-th row and the j-th column[1].

The expression (2) is called *Laplace expansion* of the determinant along the i-th row. As we will see hereafter $\det(A) = \det(A^T)$ and therefore there exists an equivalent *Laplace expansion* along the j-th column, which is

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A^{ij}) \qquad \text{(for every fixed } j\text{)}. \tag{4}$$

Remark: The fact that we can arbitrarily choose the row or the column along which develop the determinant and always obtain the same result suggests that the determinant is a rather special function. From a technical viewpoint it can be shown that (2) is the a unique alternating multilinear function[2]

$$\Lambda(\cdot) : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ times}} \to \mathbb{R}$$
$$(a_1, \dots, a_n) \to \Lambda(a_1, \dots, a_n) \tag{6}$$

satisfying

$$\Lambda(e_1, \dots, e_n) = 1, \tag{7}$$

where $\{e_1, \dots, e_n\}$ is the canonical basis of $\mathbb{R}^n$. In other words, we have

$$\det(A) = \Lambda(a_1, \dots, a_n), \tag{8}$$

where a_j is the j-th column of A.

[1] The number

$$c^{ij} = (-1)^{i+j} \det(A^{ij}) \tag{3}$$

is often called co-factor of a_{ij} in the determinant expansion.

[2] An alternating multilinear function is a function $\Lambda(a_1, \dots, a_n)$ that is linear in each argument a_j, e.g.,

$$\Lambda(a_1, a_2 + b_2, a_3) = \Lambda(a_1, a_2, a_3) + \Lambda(a_1, b_2, a_3),$$

and changes sign if we interchange a_j with a_i. For instance,

$$\Lambda(a_1, a_2, a_3) = -\Lambda(a_2, a_1, a_3) = \Lambda(a_3, a_1, a_2). \tag{5}$$

If $A = a$ is a number then we set $\det(A) = a$. Note that the determinant of a matrix is a *nonlinear* function of the matrix entries which is defined recursively in terms of determinants of the matrices A^{ij}, which have smaller dimension.

Examples: Let us provide a few examples of calculation of determinants

1. $A = a$ (real number). In this case we have $\det(A) = a$.

2. $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. We develop the determinant along the first row, i.e., set $i = 1$ in (2).

   This yields
   $$\det(A) = (-1)^{1+1} a_{11} \det(A^{11}) + (-1)^{1+2} a_{12} \det(A^{12}), \tag{9}$$
   where
   $$A^{11} = a_{22}, \qquad A^{12} = a_{21}. \tag{10}$$
   Therefore we obtain
   $$\det(A) = a_{11}a_{22} - a_{12}a_{21}. \tag{11}$$
   Note that we obtain exactly the same formula if we develop the determinant along the second row, the first column or the second column.

3. $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. We develop the determinant along the first row, i.e., set $i = 1$ in (2).

   This yields
   $$\det(A) = (-1)^{1+1} a_{11} \det(A^{11}) + (-1)^{1+2} a_{12} \det(A^{12}) + (-1)^{1+3} a_{13} \det(A^{13}) \tag{12}$$
   where
   $$A^{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \qquad A^{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} \qquad A^{11} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}. \tag{13}$$
   Computing the determinants of A^{11}, A^{12} and A^{13} yields the formula
   $$\det(A) = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}). \tag{14}$$
   Note that we obtain exactly the same formula if we develop the determinant along any other row or column.

Since we can equivalently expand the determinant along arbitrary rows orcolumns of A it is convenient to choose the row or the column with the largest number of zeros. This minimizes the number of calculations when computing the determinant using (2) or (4). For example, it is clear that it is convenient to compute the determinant of the following matrix along the second column:
$$\det\left(\begin{bmatrix} 3 & 1 & 4 \\ -1 & 0 & 2 \\ 1 & 0 & 5 \end{bmatrix}\right) = -1\det\left(\begin{bmatrix} -1 & 2 \\ 1 & 5 \end{bmatrix}\right) = -1(-5-2) = 7. \tag{15}$$

Properties of the determinant. The determinant of any $n \times n$ matrix A satisfies the following important properties:

1. $\det(A) = \det\left(A^T\right)$
2. $\det(A)$ is a *linear function* of the columns (or the rows) of the matrix A. In other words, if we denote by a_i the i-th column of A and B a column vector of the same length of a_i then
   (a) $\det\left(\begin{bmatrix}a_1 \cdots (a_i + b) \cdots a_n\end{bmatrix}\right) = \det\left(\begin{bmatrix}a_1 \cdots a_i \cdots a_n\end{bmatrix}\right) + \det\left(\begin{bmatrix}a_1 \cdots b \cdots a_n\end{bmatrix}\right)$,
   (b) $\det\left(\begin{bmatrix}a_1 \cdots ca_i \cdots a_n\end{bmatrix}\right) = c\det\left(\begin{bmatrix}a_1 \cdots a_i \cdots a_n\end{bmatrix}\right)$,

   where a_i is the i-th column of A.
3. If the columns or the rows of A are linearly dependent then $\det(A) = 0$. If the columns or the rows or A are linearly independent (i.e., A is full rank) then $\det(A) \neq 0$.
4. If a multiple of one row (or one column) is added or subtracted to another row (or column) then the determinant does not change (this follows from property 2 by setting with $B = cA_j$, and property 3).
5. If two rows (or two columns) are interchanged then the determinant changes sign.

These properties can be easily verified for 2×2 and 3×3 matrices. The proof of these properties for general $n \times n$ matrices can be found in the book.

Note that from property 2(b) it follows that for any number c and any $n \times n$ matrix A:

$$\det(cA) = c^n \det(A). \tag{16}$$

In fact, the matrix cA has all columns (n in total) multiplied by c.

Example: Let us show properties 1. to 5. for the simple matrix

$$A = \begin{bmatrix}1 & 2\\ 1 & 3\end{bmatrix}, \qquad \det(A) = 1. \tag{17}$$

We have:

1. $\det\left(A^T\right) = \det\left(\begin{bmatrix}1 & 1\\ 2 & 3\end{bmatrix}\right) = 1 = \det(A)$.
2. (a) $\det\left(\begin{bmatrix}1 & 1\\ 2 & 3\end{bmatrix}\right) = \det\left(\begin{bmatrix}1 & 0\\ 2 & 3\end{bmatrix}\right) + \det\left(\begin{bmatrix}1 & 1\\ 2 & 0\end{bmatrix}\right) = 3 - 2 = 1$.

   (b) Multiply the second column of A by 3. This yields $\det\left(\begin{bmatrix}1 & 3\\ 2 & 9\end{bmatrix}\right) = 3 = 3\underbrace{\det(A)}_{=1}$.
3. A has rank 2 and therefore its columns are linearly independent. However, if we consider the rank 1 matrix $\begin{bmatrix}1 & 3\\ 1 & 3\end{bmatrix}$ then

$$\det\left(\begin{bmatrix}1 & 3\\ 1 & 3\end{bmatrix}\right) = 3 - 3 = 0. \tag{18}$$

4. Multiply the first row of A by 2 and add it to the second row to obtain

$$\det\left(\begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}\right) = 7 - 6 = 1 = \det(A). \tag{19}$$

5. Interchange the first and the second row of A to obtain

$$\det\left(\begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix}\right) = 2 - 3 = -1 = -\det(A). \tag{20}$$

From Property 3. at page 3 it follows that:

Theorem 1. Let A be a $n \times n$ matrix, b a $n \times 1$ column vector. Then

1. A is invertible $\Leftrightarrow \det(A) \neq 0$.
2. A has linearly independent rows/columns $\Leftrightarrow \det(A) \neq 0$.
3. The linear system of equations $Ax = b$ has a unique solution $\Leftrightarrow \det(A) \neq 0$.

Determinant of matrix products and matrix inverses. By using the definition of determinant (2) it can be shown that for every $A, B \in M_{n\times n}$ we have

$$\det(AB) = \det(A)\det(B). \tag{21}$$

Clearly, this implies that

$$\det(AB) = \det(BA) \quad \text{and} \quad \det(A^p) = \det(A)^p \quad \text{for all } p \in \mathbb{N}. \tag{22}$$

By using these identities it is straightforward to show, e.g.,

$$\det(AB^TACAB) = \det(A)^3 \det(B)^2 \det(C), \tag{23}$$

where A, B and C are three $n \times n$ matrices. Moreover, if A is an invertible matrix then

$$1 = \det(AA^{-1}) = \det(A)\det(A^{-1}) \quad \Rightarrow \quad \det(A^{-1}) = \frac{1}{\det(A)}, \tag{24}$$

i.e. the determinant of the inverse matrix is the inverse of the determinant.

Computing the determinant of a matrix efficiently. How do we actually compute the determinant of a matrix? We have seen that one possibility is to use the definition (2), i.e., the Laplace rule. However, this not really computationally efficient if the dimension of the matrix is even moderately high, e.g., larger than 10 or 20. In fact, it can be shown that the number of operations to compute (2) is exactly

$$p = \lfloor n!e \rfloor - 2. \tag{25}$$

In this formula, $e = 2.7183...$ is the Napier number and the symbol $\lfloor n!e \rfloor$ denotes the nearest integer number smaller or equal than $n!e$, where $n!$ is the factorial of n. For instance, if $n = 2$ we have

$$p = \lfloor 2!e \rfloor - 2 = \lfloor 5.4366 \rfloor - 2 = 5 - 2 = 3.$$

In fact, as we see from equation (11), to compute the determinant we need two multiplications and one subtraction. Similarly, for 3×3 matrices ($n = 3$) we need

$$p = \lfloor 3!e \rfloor - 2 = \lfloor 16.3097 \rfloor - 2 = 16 - 2 = 14$$

operations. In fact, as we see from equation (12), to compute the determinant we need 9 multiplications and 5 subtractions. The number of operations increases exponentially fast as we increase the dimension of the matrix. For example, for a 20×20 matrix the Laplace rule (2) requires

$$p = \lfloor 20!e \rfloor - 2 \simeq 6.61 \times 10^{18} \qquad \text{operations.}$$

The 2022 Apple M2 Max processor is capable of 13.6 Teraflops in single precision (32 bits), i.e., 13.6×10^{12} single-precision floating point operations per second. Hence, to compute the determinant of a 20×20 matrix by using the Laplace rule on the latest MacBook Pro with M2 Max processor/GPU we need to let our laptop run for approximately

$$\frac{6.61 \times 10^{18} \ \text{operations}}{13.6 \times 10^{12} \ \text{flops}} = 4.86 \times 10^5 \ \text{seconds} \simeq 5.62 \ \text{days} \tag{26}$$

to complete the calculation. Repeating a similar calculation for a 22×22 would require

$$\frac{30.554 \times 10^{20} \ \text{operations}}{13.6 \times 10^{12} \ \text{flops}} = 2.2632 \times 10^8 \ \text{seconds} \simeq 7.18 \ \text{years.} \tag{27}$$

Fortunately, there is a more efficient algorithm to compute the determinant of a matrix. In fact, by using elementary row operations we known that we can reduce the matrix A to the following matrix in row-echelon form

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{12} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix} \tag{28}$$

The matrix U has the same determinant of A, up to a sign (i.e., $+$ or $-$) determined by how many times we interchange rows in the Gauss elimination with pivoting-by-row process. If we denote by s the number of row permutations we take in the Gauss elimination process we have

$$\det(A) = (-1)^s \prod_{k=1}^{n} u_{kk}. \tag{29}$$

In fact, the determinant of an upper-triangular (or a lower-triangular) matrix is simply the product of the diagonal elements. The total number of operations to transform an $n \times n$ matrix A into the upper triangular form U is

$$\frac{2}{3}n^3 - \frac{n^2}{2} - \frac{n}{6}. \tag{30}$$

The number of products in (29) is n, while taking the exponential is one operation. Hence, the total number of operations to compute the determinant with Gauss elimination is

$$\frac{2}{3}n^3 - \frac{n^2}{2} - \frac{n}{6} + n + 1 = \frac{2}{3}n^3 - \frac{1}{2}n^2 + \frac{5}{6}n + 1 \tag{31}$$

For a 22×22 matrix we get 6876 operations. If we use a 2022 Apple M2 Max processor, this requires

$$\frac{6876 \text{ operations}}{13.6 \times 10^{12} \text{ flops}} = 5.06 \times 10^{-10} \text{ seconds.} \tag{32}$$

Example: Consider the matrix

$$A = \begin{bmatrix} 1 & -1 & 3 \\ 1 & -2 & 1 \\ -1 & 1 & 3 \end{bmatrix} \quad \rightarrow \quad \text{Gauss elimination} \quad \rightarrow \quad U = \begin{bmatrix} 1 & -1 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & 6 \end{bmatrix} \tag{33}$$

Since we did not perform any permutation we have $s = 0$ in (29) and therefore

$$\det(A) = \det(U) = -6.$$

.

Cramer's rule. It is possible to express the solution to a linear system of equations in terms of determinants. Specifically, let

$$Ax = b \tag{34}$$

be a system of n linear equations in n unknowns. Suppose that the system has a unique solution (i.e., $\det(A) \neq 0$). Then

$$x_i = \frac{1}{\det(A)} \det([a_1 \cdots b \cdots a_n]), \tag{35}$$

where $[a_1 \cdots b \cdots a_n]$ is a matrix obtained by replacing the i-th column of A (denoted by a_i) with the column vector b.

Example: Compute the solution to the following system of equations using Cramer's rule:

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{b}. \tag{36}$$

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

$$x_1 = \det\left(\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\right) = -1, \qquad x_2 = \det\left(\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\right) = 1. \tag{37}$$

An explicit formula for A^{-1} . Let A be an invertible matrix. By definition, the inverse of A is a square matrix (denoted as A^{-1}) with the following properties

$$AA^{-1} = I_n \qquad A^{-1}A = I_n, \tag{38}$$

where I_n is the $n \times n$ identity matrix. Let h_i be the columns of the matrix A^{-1}, i.e.,

$$A^{-1} = [h_1 \quad h_2 \quad \cdots \quad h_n] \qquad h_i \in M_{n\times 1} \quad i = 1, \ldots, n. \tag{39}$$

By definition of matrix-vector product we have

$$AA^{-1} = [Ah_1 \quad Ah_2 \quad \cdots \quad Ah_n]. \tag{40}$$

At this point, define the following column vectors $e_i \in M_{n\times 1}$ $(i = 1, ..., n)$

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \qquad \cdots, \qquad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{41}$$

Note that e_i is the i-th column of the identity matrix I_n. With this notation we can write the matrix equation $AA^{-1} = I_n$ as

$$[Ah_1 \quad Ah_2 \quad \cdots \quad Ah_n] = [e_1 \quad e_2 \quad \cdots \quad e_n]. \tag{42}$$

Hence, the n columns of the inverse matrix A^{-1}, i.e., h_1, ..., h_n are solutions to n linear systems

$$Ah_1 = e_1, \qquad Ah_2 = e_2, \qquad \ldots, \qquad Ah_n = e_n. \tag{43}$$

By using Cramer's rule we obtain that the i-th component of the column vector h_j is

$$h_{ji} = \frac{1}{\det(A)} \det\left([a_1 \cdots e_j \cdots a_n]\right) \tag{44}$$

where $[a_1 \cdots e_j \cdots a_n]$ is a matrix in which we replaced the i-th column a_i with e_j. By using the Laplace rule along the i-th column of $[a_1 \cdots e_j \cdots a_n]$ we obtain

$$\det\left([a_1 \cdots e_j \cdots a_n]\right) = (-1)^{i+j} \det(A^{ji}) = C^{ji} \qquad (j,i)\text{-cofactor}. \tag{45}$$

This yields the following expression

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} C^{11} & \cdots & C^{1n} \\ \vdots & \ddots & \vdots \\ C^{n1} & \cdots & C^{nn} \end{bmatrix}^T. \tag{46}$$

Example: Compute the inverse of the following matrix

$$A = \begin{bmatrix} -2 & 4 \\ 4 & 3 \end{bmatrix}. \tag{47}$$

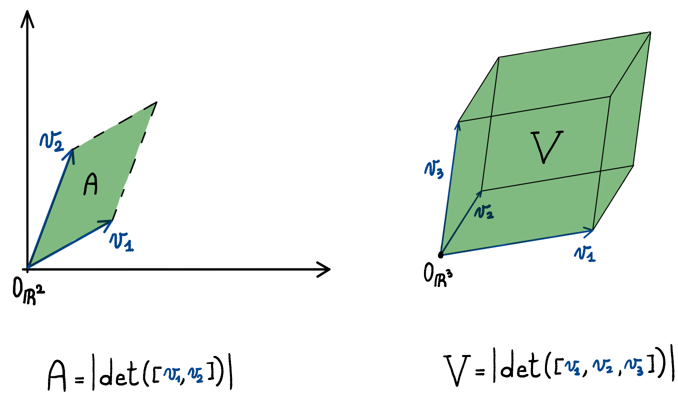
We have $\det(A) = -22$, and

$$\begin{aligned}
C^{11} &= \det\left(\begin{bmatrix} e_1 & a_2 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 1 & 4 \\ 0 & 3 \end{bmatrix}\right) = 3, \\
C^{12} &= \det\left(\begin{bmatrix} a_1 & e_1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} -2 & 1 \\ 4 & 0 \end{bmatrix}\right) = -4, \\
C^{21} &= \det\left(\begin{bmatrix} e_2 & a_2 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 0 & 4 \\ 1 & 3 \end{bmatrix}\right) = -4, \\
C^{22} &= \det\left(\begin{bmatrix} a_1 & e_2 \end{bmatrix}\right) = \det\left(\begin{bmatrix} -2 & 0 \\ 4 & 1 \end{bmatrix}\right) = -2.
\end{aligned}$$

Therefore,

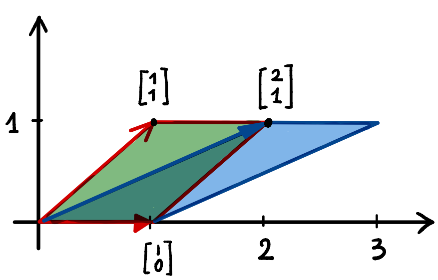
$$A^{-1} = \frac{1}{22} \begin{bmatrix} -3 & 4 \\ 4 & 2 \end{bmatrix}. \tag{48}$$

Volumes of parallelograms. The determinant of a matrix represents the volume enclosed by the vectors defined by the columns (or the rows) or the matrix.



At this point we notice that there are quite a lot of properties of A and V following from the properties of determinant. For example, if we add a scalar multiple of v_1 to v_2, then the area of the parallelogram defined by the two vectors does not change. This follows from the fact that the determinant is a linear function of the columns, and that the determinant of a matrix with linearly dependent columns is equal to zero. For example,

$$A = \left|\det\left(\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}\right)\right| = \left|\det\left(\begin{bmatrix}1 & 2\\0 & 1\end{bmatrix}\right)\right|. \tag{49}$$



Of course, the green and blue areas are the same. Other properties of the area of a parallelogram can be derived from properties of the determinant. Next, consider an invertible transformation $F : \mathbb{R}^n$ $\mathbb{R}^n$, represented by $n \times n$ invertible matrix L. We know that if $\{v_1, \dots, v_n\}$ is a basis of $\mathbb{R}^n$ then

$$\begin{bmatrix}u_1 & \dots & u_n\end{bmatrix} = \begin{bmatrix}Lv_1 & \dots & Lv_n\end{bmatrix} = L\begin{bmatrix}v_1 & \dots & v_n\end{bmatrix} \tag{50}$$

is also a basis of $\mathbb{R}^n$. The volume of the parallelograms enclosed by $\{v_1, \dots, v_n\}$ and $\{u_1, \dots, u_n\}$ are

$$V_0 = \left|\det\left(\begin{bmatrix}v_1 & \dots & v_n\end{bmatrix}\right)\right|, \qquad V_1 = \left|\det\left(\begin{bmatrix}u_1 & \dots & u_n\end{bmatrix}\right)\right|. \tag{51}$$

By applying the determinant operator to equation (50), and using the fact that the determinant of the matrix product is the product of the matrix determinants we see that

$$V_1 = |\det(L)|\, V_0. \tag{52}$$

This formula is very important in a variety of fields ranging from multi-dimensional integration theory to continuum mechanics.

Lecture 11: Eigenvalues and Eigenvectors

Consider a vector space V and the linear transformation transformation

$$F : V \mapsto V. \tag{1}$$

We say that $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) is an *eigenvalue* of F if there exists a *nonzero* vector $v \in V$ such that

$$F(v) = \lambda v. \tag{2}$$

We call v *eigenvector* of F corresponding to the eigenvalue λ.

Note that, by definition, we are not allowing eigenvectors to be zero, i.e., $v = 0_V$ is not an eigenvector. If we allow $v = 0_V$ to be an eigenvector, then any number λ would be an eigenvalue of F. However, we can have eigenvectors corresponding to zero eigenvalues. In this case the eigenvector belongs to the nullspace of F, since $\lambda = 0 \Rightarrow F(v) = 0_V$.

Next, suppose that V is n-dimensional and let $\mathcal{B}_V = \{u_1, \ldots, u_n\}$ be a basis of V. Denote by $A_{\mathcal{B}_V}$ be the matrix associated with the linear transformation F relative to the basis $\mathcal{B}_V$. We have seen that the coordinates of $F(v)$ relative to the basis $\mathcal{B}_V$ can be expressed as

$$[F(v)]_{\mathcal{B}_V} = A_{\mathcal{B}_V}[v]_{\mathcal{B}_V} \tag{3}$$

where

$$[v]_{\mathcal{B}_V} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{4}$$

are the coordinates of v relative to $\mathcal{B}_V$. We have

$$F(v) = \lambda v \quad \Leftrightarrow \quad A_{\mathcal{B}_V}[v]_{\mathcal{B}_V} = \lambda [v]_{\mathcal{B}_V}. \tag{5}$$

Hence, computing eigenvalues and eigenvectors of matrices is equivalent to compute eigenvalues and eigenvectors of linear transformations between finite-dimensional vector spaces.

Remark: Eigenvalues and eigenvectors can be defined also for linear transformations between infinite-dimensional vector spaces. For example, consider the derivative operator

$$\begin{aligned} F : C^{(\infty)}(\mathbb{R}) &\mapsto C^{(\infty)}(\mathbb{R}), \\ f(x) &\mapsto \frac{df(x)}{dx}. \end{aligned}$$

We have seen that d/dx defines a linear transformation (linear operator) between infinite-dimensional vector spaces. an eigenvector of d/dx corrersponding to an eigenvalue λ has the form $\psi(x) = e^{\lambda x}$. In fact

$$\frac{de^{\lambda x}}{dx} = \lambda e^{\lambda x} \qquad \Rightarrow \qquad \frac{d\psi(x)}{dx} = \lambda \psi(x). \tag{6}$$

Eigenvectors belonging to function spaces are often called *eigenfunctions*.

Eigenvalues of a matrix. Consider a $n \times n$ matrix A with real or complex coefficients. If $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) and $v \in \mathbb{R}^n$ (or $\mathbb{C}^n$) are, respectively, and eigenvalue of A and an eigenvector of A corresponding to λ then

$$Av = \lambda v. \tag{7}$$

Equation (7) is also called *eigenvalue problem* for the matrix A. We have

$$Av = \lambda v \quad \Leftrightarrow \quad (A - \lambda I)v = 0_{\mathbb{R}^n}, \tag{8}$$

Hence, the eigenvector v (which is non-zero by definition) is in the nullspace of the matrix $(A-\lambda I)$. This implies that the matrix $A - \lambda I$ is not injective and therefore not invertible. Equivalently, by using the matrix rank theorem we have that

$$\text{rank}(A - \lambda I) = n - \underbrace{\dim(N(A - \lambda I))}_{\geq 1} < n. \tag{9}$$

This shows that the matrix $(A - \lambda I)$ is not full rank and therefore it is not invertible. A necessary and sufficient condition for $(A - \lambda I)$ to be not invertible is

$$p(\lambda) = \det(A - \lambda I) = 0 \qquad \text{(characterististic equation)}, \tag{10}$$

The polynomial

$$p(\lambda) = \det(A - \lambda I) \tag{11}$$

is known as *characteristic polynomial* associated with the matrix A. The characteristic equation (10) implies that the eigenvalues of a matrix A are roots of the characteristic polynomial $p(\lambda)$.

How many eigenvalues do we have for a given $n \times n$ matrix A? The characteristic polynomial $p(\lambda)$ associated with a $n \times n$ matrix A is a polynomial of degree n with real or complex coefficients (complex coefficients if the matrix A has complex entries). By using the fundamental theorem of algebra (see Lecture 3) we conclude that every $n \times n$ matrix has exactly n complex eigenvalues. Some of such eigenvalues may be repeated, in which case we say that they have "algebraic multiplicity" greater than one. In other words, the multiplicity of an eigenvalue as a root of the characteristic polynomial is called *algebraic multiplicity* the eigenvalue.

If the matrix A is real then the characteristic polynomial $p(\lambda)$ has real coefficients and therefore the roots of $p(\lambda)$ are either real or complex conjugates.

Example 1: Compute the eigenvalues of the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}. \tag{12}$$

The characteristic polynomial is

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} = -(2-\lambda)(6+\lambda) - 9, \tag{13}$$

i.e.,

$$p(\lambda) = \lambda^2 + 4\lambda - 21. \tag{14}$$

The eigenvalues of A are roots of $p(\lambda)$. Setting $p(\lambda) = 0$ yields

$$\lambda_{1,2} = -2 \pm \sqrt{4+21} = -2 \pm 5 \qquad \Rightarrow \qquad \lambda_1 = 3, \quad \lambda_2 = -7. \tag{15}$$

In this case, both eigenvalues have algebraic multiplicity one, i.e., they are simple roots of $p(\lambda)$. The characteristic polynomial can be factored as

$$p(\lambda) = (\lambda - 3)(\lambda + 7), \tag{16}$$

suggesting once again that $\lambda = 3$ and $\lambda = -7$ are simple roots.

Example 2: Compute the eigenvalues of the matrix

$$A = \begin{bmatrix} 2 & 5 & 1 & -5 \\ 0 & 4 & 3 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{17}$$

In this case we have

$$A - \lambda I = \begin{bmatrix} 2-\lambda & 5 & 1 & -5 \\ 0 & 4-\lambda & 3 & 0 \\ 0 & 0 & 2-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix} \tag{18}$$

and

$$p(\lambda) = \det(A - \lambda I) = (2-\lambda)^2(4-\lambda)(1-\lambda). \tag{19}$$

Hence, the matrix A has three eigenvalues:

$$\begin{aligned} \lambda_1 &= 2 \qquad \text{with algebraic multiplicity 2,} \\ \lambda_2 &= 4 \qquad \text{with algebraic multiplicity 1,} \\ \lambda_3 &= 1 \qquad \text{with algebraic multiplicity 1.} \end{aligned}$$

Note that the eigenvalues coincides with the diagonal entries of the matrix A. This is a general fact about upper or or lower triangular matrices, i.e., the eigenvalues of such matrices coincides with the diagonal entries of the matrix. For example, the following matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{20}$$

has two eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 0$, both with algebraic multiplicity 2.

Example 3: Compute the eigenvalues of the following matrix

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}. \tag{21}$$

The characteristic polynomial is

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 2 \\ -1 & 1-\lambda \end{bmatrix} = -(1-\lambda)^2 + 2, \tag{22}$$

i.e.,

$$p(\lambda) = \lambda^2 - 2\lambda + 3. \tag{23}$$

Hence, the eigenvalues are

$$\lambda_1 = 1 + i\sqrt{2} \qquad \lambda_2 = 1 - i\sqrt{2} \tag{24}$$

Note that λ_1 and λ_2 are complex conjugates eigenvalues. Clearly, for 2×2 matrices with real entries the fundamental theorem of algebra tells us that the eigenvalues are either both real or complex conjugates.

Eigenvectors and eigenspaces. By definition, an eigenvector of a $n \times n$ matrix A is a nonzero vector $v \in \mathbb{R}^n$ such that

$$Av = \lambda v. \tag{25}$$

This means that v is an element of the nullspace of $(A - \lambda I)$ since v is mapped onto the zero of $\mathbb{R}^n$ by $(A - \lambda I)$. We know that such a nullspace is a vector subspace of $\mathbb{R}^n$.

In the context of eigenvalue problems, we call $N(A - \lambda I)$ the *eigenspace* of A corresponding to the eigenvalue λ. The dimension of the eigenspace $N(A - \lambda I)$ is called *geometric multiplicity* of the eigenvalue λ. By definition, an eigenvector cannot be zero and therefore the eigenspace corresponding to each eigenvalue has dimension at least equal to one. The dimension of the eigenspace corresponding to a certain eigenvalue can be computed by using the matrix rank theorem.

Example 4: Compute the eigenspaces of the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} \tag{26}$$

We have seen in a previous example that the eigenvalues of A are $\lambda_1 = 3$ and $\lambda_2 = -7$. Let us compute the eigenspace corresponding to λ_1. To this end, we first compute the dimension of such eigenspace by using the matrix rank theorem

$$\dim(N(A - \lambda_1 I)) = 2 - \operatorname{rank}(A - \lambda_1 I) = 2 - \operatorname{rank}\left(\begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}\right) = 2 - 1 = 1 \tag{27}$$

Hence, the eigenspace corresponding to λ_1 has dimension one. Any vector of such an eigenspace is an eigenvector of A corresponding to λ_1. To compute a basis for the eigenspace $N(A - \lambda_1 I)$ consider

$$(A - \lambda_1 I)v = 0_{\mathbb{R}^2} \quad \Leftrightarrow \quad \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad -v_1 + 3v_2 = 0 \tag{28}$$

Hence,

$$v = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \tag{29}$$

is a basis for $N(A - \lambda_1 I)$, and an eigenvector of A corresponding to λ_1. All eigenvectors of A corresponding to λ_1 are in the form

$$c \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{with } c \neq 0. \tag{30}$$

Similarly, the eigenspace corresponding to λ_2 has dimension 1 and can be determined by solving the linear system

$$(A - \lambda_2 I)v = 0_{\mathbb{R}^2} \quad \Leftrightarrow \quad \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad 3v_1 + v_2 = 0. \tag{31}$$

Hence,

$$v = \begin{bmatrix} 1 \\ -3 \end{bmatrix} \tag{32}$$

is a basis for $N(A - \lambda_2 I)$ and an eigenvector of A corresponding to λ_2. In summary, λ_1 and λ_2 are eigenvalues with algebraic multiplicity one and geometric multiplicity one. Geometric multiplicity one means that the eigenspaces $N(A - \lambda_1 I)$ and $N(A - \lambda_2 I)$ are both one-dimensional. A basis for $N(A - \lambda_1 I)$ and $N(A - \lambda_2 I)$ is given by (29) and (32), respectively.

The following theorem establishes a relationship between the algebraic multiplicity and the geometric multiplicity of an eigenvalue λ.

Theorem 1. Let λ be an eigenvalue of a $n \times n$ matrix A. Denote by s the algebraic multiplicity of λ. Then

$$\dim(N(A - \lambda I)) \leq s. \tag{33}$$

In other words the geometric multiplicity of λ (i.e., the dimension of the associated eigenspace) is always smaller or equal than the algebraic multiplicity).

Of course, if λ is a simple eigenvalue ($s = 1$) then $\dim(N(A - \lambda I)) = 1$, i.e., the eigenspace corresponding to simple eigenvalues is always one-dimensional. If λ has algebraic multiplicity 2, i.e., it is a repeated eigenvalue, then it is possible to have geometric multiplicity equal to one or equal to two. In the latter case the eigenspace is two-dimensional and any vector in such eigenspace (including linear combinations of multiple eigenvectors) is an eigenvector. Let us provide a simple example of a 2×2 matrix with one eigenvalue of algebraic multiplicity two and geometric multiplicity one

Example 5: Consider the following matrix

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}. \tag{34}$$

We know that $\lambda = 2$ is the only eigenvalue and it has algebraic multiplicity two. In fact, the characteristic polynomial is $p(\lambda) = (2 - \lambda)^2$. The geometric multiplicity of $\lambda = 2$ can be calculated by using the matrix rank theorem

$$\dim(N(A - \lambda I)) = 2 - \text{rank}(A - \lambda I) = 2 - \underbrace{\text{rank}\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right)}_{=1} = 2 - 1 = 1. \tag{35}$$

Hence, the eigenspace associated with $\lambda = 2$ is one-dimensional. A basis for such an eigenspace is obtained as follows:

$$(A - \lambda I)v = 0_{\mathbb{R}^2} \quad \Leftrightarrow \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad v_2 = 0. \tag{36}$$

We can choose as basis

$$v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{37}$$

Example 6: Compute the eigenvalues and the eigenvectors of the following matrix

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 2 \end{bmatrix}. \tag{38}$$

This is an upper triangular matrix and therefore the eigenvalues coincide with the diagonal entries. Hence we have $\lambda_1 = 2$ with algebraic multiplicity two and $\lambda_2 = 1$ with algebraic multiplicity one.

$$A - \lambda_1 I = \begin{bmatrix} 0 & 1 & 3 \\ 0 & -1 & 5 \\ 0 & 0 & 0 \end{bmatrix} \qquad \Leftrightarrow \qquad \dim(N(A - \lambda_1 I)) = 3 - \underbrace{\text{rank}(A - \lambda_1 I)}_{=2} = 1 \tag{39}$$

$$A - \lambda_2 I = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 0 & 5 \\ 0 & 0 & 1 \end{bmatrix} \qquad \Leftrightarrow \qquad \dim(N(A - \lambda_2 I)) = 3 - \underbrace{\text{rank}(A - \lambda_2 I)}_{=2} = 1 \tag{40}$$

Therefore, the dimension of the eigenspaces associated with λ_1 and λ_2 is one. Let us find a basis for such eigenspaces.

$$(A - \lambda_1 I)v = 0_{\mathbb{R}^3} \quad \Rightarrow \quad \begin{bmatrix} 0 & 1 & 3 \\ 0 & -1 & 5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} v_1 \text{ arbitrary} \\ v_2 + 3v_3 = 0 \\ -v_2 + 5v_3 = 0 \end{cases} \tag{41}$$

Hence, an eigenvector that spans $N(A - \lambda_1 I)$ is

$$v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \tag{42}$$

Similarly,

$$(A - \lambda_2 I)v = 0_{\mathbb{R}^3} \quad \Rightarrow \quad \begin{bmatrix} 1 & 1 & 3 \\ 0 & 0 & 5 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} v_1 + v_2 + 3v_3 = 0 \\ v_3 = 0 \end{cases} \tag{43}$$

Hence, an eigenvector that spans $N(A - \lambda_2 I)$ is

$$v = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}. \tag{44}$$

Theorem 2. Eigenvectors corresponding to different eigenvalues are linearly independent.

Proof. Let v_1 and v_2 be two eigenvectors of a matrix $A \in M_{n\times n}$ corresponding to two distinct eigenvalues λ_1 and λ_2. We want to show that

$$x_1 v_1 + x_2 v_2 = 0_{\mathbb{R}^n} \quad \Rightarrow \quad x_1 = x_2 = 0. \tag{45}$$

To this end we first multiply the equation above by λ_2 to obtain

$$x_1 \lambda_2 v_1 + x_2 \lambda_2 v_2 = 0_{\mathbb{R}^n} \tag{46}$$

Then we apply the matrix A to $x_1 v_1 + x_2 v_2 = 0_{\mathbb{R}^n}$ to obtain

$$x_1 A v_1 + x_2 A v_2 = x_1 \lambda_1 v_1 + x_2 \lambda_2 v_2 = 0_{\mathbb{R}^n} \tag{47}$$

Subtracting equation (46) from equation (47) yields

$$x_1 \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} \underbrace{v_1}_{\neq 0_{\mathbb{R}^n}} = 0_{\mathbb{R}^n} \qquad \Rightarrow \qquad x_1 = 0. \tag{48}$$

Substituting $x_1 = 0$ into $x_1 v_1 + x_2 v_2 = 0_{\mathbb{R}^n}$ yields $x_2 = 0$. Hence v_1 and v_2 are linearly independent.

□

Similarity transformations. Let $A, B \in M_{n\times n}$. We say that A is *similar* to B is there exists an invertible matrix $P \in M_{n\times n}$ such that

$$AP = PB \qquad \Leftrightarrow \qquad A = PBP^{-1} \tag{49}$$

The transformation $B \to PBP^{-1}$ is called *similarity transformation*. An example of similarity transformation is the change of basis transformation.

Theorem 3. Similar matrices have the same eigenvalues.

Proof. Let $A, B \in M_{n\times n}$ be two similar matrices, i.e., $P \in M_{n\times n}$ such that

$$A = PBP^{-1}. \tag{50}$$

Then

$$\det(A - \lambda I) = \det(PBP^{-1} - \lambda PP^{-1}) = \det(P)\det(B - \lambda I)\det(P^{-1}) = \det(B - \lambda I) \tag{51}$$

□

This theorem implies that the eigenvalues of a linear transformation $F : V \mapsto V$ ($\dim(V) = n$) do not depend on the basis we choose to represent F in V. In fact the matrices associated to F relative to different bases of V are related by a similarity transformation.

Diagonalization. Consider a $n \times n$ matrix A. We have seen in Theorem 2 that eigenvectors corresponding to different eigenvalues are linearly independent. Hence, if the algebraic multiplicity of each eigenvalue is equal to the geometric multiplicity then it is possible to construct a basis for $\mathbb{R}^n$ made of eigenvectors of A. Let us organize such n eigenvectors as columns of a matrix P

$$P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}. \tag{52}$$

Clearly,

$$AP = \begin{bmatrix} Av_1 & \cdots & Av_n \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \underbrace{\begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}}_{\Lambda} = P\Lambda, \tag{53}$$

where Λ is a diagonal matrix having the eigenvalues of A (counted with their multiplicity) along the diagonal. Equation (53) shows that if A has n linearly independent eigenvectors then A is similar to a diagonal matrix[1] Λ. The similarity transformation is defined by the matrix P in (52), i.e., the matrix that has the eigenvectors of A as columns.

A corollary of this statement is that matrices with simple eigenvalues are always diagonalizable, since they have n linearly independent eigenvectors. The following theorem summarizes what we just said.

Theorem 4. Let A be a $n \times n$ matrix with eigenvalues $\{\lambda_1, \dots, \lambda_p\}$ with algebraic multiplicities $\{s_1, \dots, s_p\}$, respectively. Then A is diagonalizable if and only if

$$\dim(N(A - \lambda_i I)) = s_i \quad \text{for all} \quad i = 1, \dots, p. \tag{54}$$

Example 7: The matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} \tag{55}$$

is diagonalizable. In fact we have seen that the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = -7$ (simple eigenvalues). This implies that the dimension of the associated eigenspace is one for both eigenvalues. The eigenvectors of A are

$$v_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \qquad \text{and} \qquad v_2 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}. \tag{56}$$

Define

$$P = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}. \tag{57}$$

It is straightforward to verify that

$$P^{-1} = \frac{1}{10} \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix} \tag{58}$$

and

$$A = P\Lambda P^{-1} \qquad \text{or} \qquad \Lambda = P^{-1}AP. \tag{59}$$

[1] In general, we say that a matrix A is *diagonalizable* if there exists an invertible matrix P such that A is similar to a diagonal matrix.

Example 8: The matrix

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \tag{60}$$

is not diagonalizable. In fact the algebraic multiplicity of the eigenvalue $\lambda = 2$ is two, while its geometric multiplicity is one. It is possible to show that there exists a basis made of "generalized eigenvectors" that makes A similar to a matrix J called *Jordan form* of A. In this particular example, the Jordan form of A coincides with A, i.e., A is already in a Jordan form (see the Remark at page 10).

Example 9: Verify that the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \tag{61}$$

is diagonalizable. The matrix is lower-triangular with eigenvalues $\lambda_1 = 1$ (algebraic multiplicity two) and $\lambda_2 = 2$ (algebraic multiplicity one). To verify that A is diagonalizable we just need to check that the geometric multiplicity of $\lambda_1 = 1$ is equal to two. To this end, we use the matrix rank theorem:

$$\dim(N(A - \lambda_1 I)) = 3 - \mathrm{rank}(A - \lambda_1 I) = 3 - \mathrm{rank}\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}\right) = 3 - 1 = 2 \tag{62}$$

This shows that the dimension of the nullspace of $N(A - \lambda_1 I)$, i.e., the dimension of the eigenspace associated with $\lambda_1 = 1$ is two. Let us compute a basis for such an eigenspace. To this end,

$$(A - \lambda_1 I)v = 0_{\mathbb{R}^3} \quad \Rightarrow \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} v_1 \text{ arbitrary} \\ v_2 \text{ arbitrary} \\ v_3 = -v_2 \end{cases} \tag{63}$$

Hence, a basis for the eigenspace corresponding to λ_1 is

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}. \tag{64}$$

On the other hand, the eigenspace $N(A - \lambda_2 I)$ is spanned by a vector that can be computed as

$$(A - \lambda_2 I)v = 0_{\mathbb{R}^3} \quad \Rightarrow \quad \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} v_1 = 0 \\ v_2 = 0 \\ v_3 \text{ arbitrary} \end{cases} \tag{65}$$

Therefore a matrix P that diagonalizes A is

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}. \tag{66}$$

Indeed, it can be verified by a direct calculation that

$$\underbrace{\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 2\end{bmatrix}}_{\Lambda} = \underbrace{\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 1 & 1\end{bmatrix}}_{P^{-1}}\underbrace{\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 1 & 2\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & -1 & 1\end{bmatrix}}_{P}. \tag{67}$$

Remark: It can be shown that the set of eigenvectors of any $n \times n$ matrix A can be complemented to a basis of $\mathbb{R}^n$ (or $\mathbb{C}^n$) by adding a certain number of *generalized eigenvectors*, as many as $s_i - \dim(N(A - \lambda_i I))$ in case the eigenspace $N(A - \lambda_i I)$ has dimension smaller than the algebraic multiplicity of λ_i. For instance, consider the matrix

$$A = \begin{bmatrix}2 & 1\\0 & 2\end{bmatrix} \tag{68}$$

We know that the eigenspace corresponding to $\lambda = 2$ is one-dimensional with basis

$$v = \begin{bmatrix}1\\0\end{bmatrix}. \tag{69}$$

To complement v to a basis of $\mathbb{R}^2$ we can construct another vector w as follows

$$(A - \lambda I)w = v. \tag{70}$$

Clearly, w is in the nullspace of the matrix $(A - \lambda I)^2$. It can be shown that w and v are linearly independent. We obtain.

$$(A - \lambda I)w = v \quad \Leftrightarrow \quad \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\begin{bmatrix}w_1\\w_2\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} w_1 \text{ arbitrary}\\ w_2 = 1\end{cases}. \tag{71}$$

At this point we can define

$$P = \begin{bmatrix}v & w\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \qquad \text{(matrix of generalized eigenvectors)}, \tag{72}$$

and apply A to P to obtain

$$AP = \begin{bmatrix}Av & Aw\end{bmatrix} = \begin{bmatrix}v & w\end{bmatrix}\underbrace{\begin{bmatrix}\lambda & 1\\0 & \lambda\end{bmatrix}}_{J} = PJ \tag{73}$$

Hence, A is similar to a matrix J in a particular form (not diagonal but almost diagonal), known as *Jordan canonical form*. In this particular example, A is already in a Jordan form so the similarity transformation defined by P turns out to be the identity transformation.

We conclude this section with an important theorem characterizing the *spectral properties* (i.e., eigenvalues and eigenvectors) of real symmetric matrices.

Theorem 5 (Spectral theorem for symmetric matrices)**.** If $A \in M_{n\times n}(\mathbb{R})$ is symmetric (i.e., $A = A^T$) then all eigenvalues are real and there exists an orthonormal basis of $\mathbb{R}^n$ made of eigenvectors of A.

Proof. To prove that the eigenvalues are real let us consider the following scalar product in $\mathbb{C}^n$:

$$\langle u, v\rangle = \sum_{i=1}^{n} u_i v_i^* \tag{74}$$

Suppose that u is an eigenvector of A. Then

$$\langle Au, u\rangle = \langle \lambda u, u\rangle = \lambda \langle u, u\rangle \tag{75}$$

On the other hand,

$$\langle u, Au\rangle = \langle u, \lambda u\rangle = \lambda^* \langle u, u\rangle . \tag{76}$$

The matrix A is symmetric. This implies that

$$\langle Au, u\rangle = \langle u, Au\rangle \qquad \Leftrightarrow \qquad \lambda^* = \lambda, \tag{77}$$

i.e., λ is real.

Let us now prove that eigenvectors of A corresponding to different eigenvalues are necessarily orthogonal. To this end, suppose that u_1 and u_2 are eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2. Then

$$\lambda_1 \langle u_1, u_2\rangle = \langle Au_1, u_2\rangle = \langle u_1, Au_2\rangle = \lambda_2 \langle u_1, u_2\rangle . \tag{78}$$

Since $\lambda_1 \neq \lambda_2$ we have that the previous equality is possible if and only if $\langle u_1, u_2\rangle = 0$. This means that u_1 and u_2 are orthogonal. Lastly, we need to prove that any symmetric matrix is diagonalizable. This is a little bit technical so we skip this proof.

□

Note that, in general, the eigenvectors of a matrix A are not orthogonal relative to the standard scalar product in $\mathbb{R}^n$ (or $\mathbb{C}^n$). However, if the matrix is symmetric then the eigenvectors are necessarily orthogonal[2], and they can be normalized, if needed. This yields a matrix of eigenvectors

$$P = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \quad \text{satisfying} \quad PP^T = I_n.$$

The condition $PP^T = I_n$ follows directly from $\langle u_i, u_j\rangle = \delta_{ij}$ (orthonormal eigenvectors). Hence the matrix P that contains the eigenvectors of a symmetric matrix is an orthogonal matrix.

Theorem 6. Let A be any $n \times n$ matrix. Then,

1. $\det(A) = \lambda_1\lambda_2\cdots\lambda_n$,
2. $\text{trace}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$,

where $\{\lambda_1, \ldots, \lambda_n\}$ are the eigenvalues of A counted with their multiplicity.

[2]Eigenvectors corresponding to different eigenvalues are necessarily orthogonal, while eigenvectors corresponding to the same eigenvalue with geometric multiplicity larger than one can be orthogonalized, e.g., by using Gram-Schmidt procedure.

Proof. To prove these identities, let us assume that A is diagonalizable[3]. In this case, we know that there exists a matrix P that has the eigenvectors of A as columns such that

$$A = P\Lambda P^{-1}, \qquad \text{where} \qquad \Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \tag{79}$$

is the diagonal matrix of eigenvalues. To prove 1. we simply notice that

$$\det(A) = \det(P\Lambda P^{-1}) = \det(P)\det(P^{-1})\det(\Lambda) = \det(\Lambda) = \lambda_1\lambda_2\cdots\lambda_n. \tag{80}$$

To prove 2. we notice that[4]

$$\text{trace}(A) = \text{trace}(P\Lambda P^{-1}) = \text{Tr}(PP^{-1}\Lambda) = \text{trace}(\Lambda) = \lambda_1 + \cdots + \lambda_n. \tag{81}$$

□

[3]The proof for the non-diagonalizable case is very much the same. The only difference is that we use the Jordan canonical form of A instead of the diagonal matrix of eigenvalues Λ.

[4]Recall that if A and B are two square matrices of the same size we have

$$\text{trace}(AB) = \text{trace}(BA).$$