Lecture 1

Vector Spaces and Linear Equations

What is linear algebra? There are at least two different answers, depending on one’s viewpoint:

1) Application-oriented/computational viewpoint:

Linear algebra is the study of linear equations and methods to solve them.

We will discuss some examples of linear equations in Section 1.1 below.

2) Abstract/conceptual viewpoint:

Linear algebra is the study of vector spaces and linear maps.

The term “vector space” will be defined in Section 1.2 below, the term “linear map” in Lecture 2.

The first answer represents a more “practical” viewpoint that focuses on applications and on computations. The second answer represents a more “abstract”—and arguably also more geometric—viewpoint that emphasizes the importance of linear algebra not only as a computational tool but also as a theory that develops concepts that help to gain intuition and insight. Both perspectives, the computational one and the conceptual one, are important, and we will strive to balance them.

1.1 Linear Equations

A linear equation is an equation in which each term is either a constant or a product of a constant and a variable.\(^1\) Here is a simple linear equation:

\[
y = 9.81x
\]  

\(^1\) Traditionally, variables are often denoted by letters from the end of the alphabet like \(x, y, z\), or by a letter with a subscript, like \(x_1, x_2, x_3, x_4, x_5, \ldots\) if we need more variables and run out of letters. Note that this is only a convention, and there is nothing wrong with using other symbols, we just have to specify clearly which symbols are considered constants and which variables.
This equation has many possible interpretations. For instance, it might express the force $y$ (measured in Newton) that acts on an object of mass $x$ (measured in kg) in the Earth’s gravitational field; or $y$ might be the distance (in meters) that an animal, say a rabbit, runs in time $x$ (measured in seconds).

More abstractly, we might want to consider a linear equation without specifying numeric constants, e.g., the following equation in two variables $x$ and $y$ and involving two arbitrary, non-specified constants $a$ and $b$:

$$y = ax + b$$  \hspace{1cm} (1.2)

For instance, this equation could be a mathematical model\(^2\) for the growth of, e.g., larvae, with $b$ being the initial amount of larvae (in grams, say), $a$ the growth rate, and $y$ denoting the amount of larvae after time $x$.

The advantage of writing an equation in this form (1.2) with abstract, unspecified constants, is that we can formulate statements that are true for all linear equations of this form. For instance, we can visualize Equation (1.2) as specifying a (nonvertical) line $\ell$ in the plane with coordinates $x$ and $y$, where $a$ denotes the slope\(^3\) if of the line and $b$ is the $y$-coordinate at which the line intersects the $y$-axis/

Example 1.1 (Balancing Chemical Reactions). Here is an example of how a system of several linear equations may arise in applications. Consider a chemical reaction equation, e.g.,

$$\text{NaOH} + \text{H}_2\text{SO}_4 \rightarrow \text{Na}_2\text{SO}_4 + \text{H}_2\text{O}. \hspace{1cm} (1.3)$$

One basic question that arises is: In what proportions should the reactants (on the left-hand side) be mixed, and in what proportions do we obtain the products (on the right hand side). By basic rules of chemistry, for each element that appears in the reaction

\(^2\)We stress that in many situations, a linear equation only gives an idealized or simplified model that only describes or even approximates the true behavior of the system in a particular range of parameters, e.g., the growth of larvae before they reach pupation.

\(^3\)The slope is defined as the change in the $y$-coordinate of a point on the line as the $x$-coordinate changes by 1.
equation, we need to have the same number of atoms before and after the equation. Let us use variables \( x, y, z, w \) to denote the respective amounts of the chemicals (e.g., number of molecules):

\[
\begin{align*}
    x &= \text{amount of NaOH} \\
    y &= \text{amount of } \text{H}_2\text{SO}_4 \\
    z &= \text{amount of Na}_2\text{SO}_4 \\
    w &= \text{amount of } \text{H}_2\text{O}
\end{align*}
\]

Then we get the following system of four equations (one for each element):

\[
\begin{align*}
    x &= 2z \quad (\text{Na}) \\
    x + 4y &= 4z + w \quad (\text{O}) \\
    y &= z \quad (\text{S}) \\
    x + 2y &= 2w \quad (\text{H})
\end{align*}
\]

This is still a fairly small system of equations that may be able to solve by trial and error (e.g., by the third equation \( y = z \), we can immediately eliminate one variable), but for more complicated reactions involving more chemicals, we might get more complicated systems with dozens of variables.

More generally, systems of linear equations that arise in application of linear algebra (in physics, engineering, computer science, chemistry, biology, statistics, . . . ) may be huge, involving many (hundred, thousands, maybe millions) of variables and equations. It is impossible to solve such huge equations by hand, but even if we want to use a computer, we have to understand the basic theory of linear equations.

**General Systems of Linear Equations.** Here is the general form of a system of \( m \) linear equations in \( n \) variables in standard form, i.e., with all terms involving variables on the left-hand side and all constant terms on the right-hand side.

\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\
    a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\
    \vdots & \quad \vdots & \quad \vdots \\
    a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m
\end{align*}
\]

(1.5)

Note that the coefficients \( a_{ij} \) on the left-hand side have a double subscript, with the first index \( i \), called the row index indicating the equation, and the second index \( j \), called the column index indicating the variable.

A system of equations of this form (1.5) is called homogeneous if all constant terms \( b_j \) on the right-hand side are zero.

Solving such a system involves two questions or tasks:

1) Is there any solution, i.e., are there numbers \( x_1, x_2, \ldots, x_n \) satisfy (1.5)?

2) Describe all possible solutions.

**Remark.** In this course, we will mostly focus on the situation where the coefficients \( a_{ij} \) and \( b_j \) of the system of linear equations are real numbers, and we seek solutions \( x_i \) that are real numbers as well.

Later, we will also consider settings where the coefficients and/or solutions are of a different kind, e.g., complex numbers.
1.2 Vector Spaces

In this section, we introduce the notion of a vector space. Apart from introducing one of the key notions of linear algebra, the goal is to illustrate a general process of how abstract concepts are formed in mathematics.

1) Start with a simple situation in which we can consider concrete examples or that we can visualize. Here, this will be vectors in the 2-dimensional coordinate plane $\mathbb{R}^2$ or in 3-dimensional coordinate space $\mathbb{R}^3$. In these settings, we can form a geometric intuition of “vectors”, “vector addition” and related concepts. We can also use these to describe physical quantities like forces, which hopefully adds even more intuition.

2) Next, we note that the basic definitions work just as well in general $n$-dimensional coordinate space $\mathbb{R}^n$. Here, we start viewing a sequence of real numbers $x_1, x_2, \ldots, x_n$, which may or may not constitute a solution to a system of linear equations as in (1.5), as a single object, which we call a vector. We may not be able to visualize $\mathbb{R}^4$ or $\mathbb{R}^17$ anymore or have a physical interpretation for it, but as we have seen above, linear systems of equations with a large number $n$ of variables arise in many applications. Thinking about a sequence of numbers as a single mathematical object is an important conceptual step. In particular, it allows us to transfer our geometric intuition, by looking for analogies with the low-dimensional situations that we can visualize.

3) A key step for transfer to work is that we formulate the concepts and methods that we use with great care, focusing on a few key aspects that capture the essence of our geometric intuition but are otherwise pretty general. In this way, we can use analogy and intuition gained from low dimensions as a “guiding light”, while only relying on a few easily verified abstract properties for formal and precise proofs. This also allows us to transfer our intuition and insights to even more general and sometimes surprising situations that we may not initially have imagined. It should be appreciated that this step of abstraction, generalization and transfer of intuition is subtle and takes some getting used to (in particular, it took mathematicians hundreds of years to explicitly formulate the notion of an abstract vector space).

Sometimes, this whole process fails and low-dimensional intuition goes completely wrong, but in linear algebra, it generally works amazingly well.

The Coordinate Plane $\mathbb{R}^2$

Abstractly, the coordinate plane $\mathbb{R}^2$ is defined as the set of all ordered pairs $(x_1, x_2)$ of real numbers, i.e.,

$$\mathbb{R}^2 = \{(x_1, x_2) \mid x_1, x_2 \in \mathbb{R}\}.$$

We can visualize $\mathbb{R}^2$ as a 2-dimensional plane in which we have chosen two coordinate axes:
Every element of $\mathbb{R}^2$ then has two geometric interpretations:

1) A pair $(x_1, x_2)$ can be interpreted as the point of the plane with coordinates $x_1$ and $x_2$. In particular, the pair $(0,0)$ is the origin of the coordinate system, where the coordinate axes meet.

2) A pair $(x_1, x_2)$ can also be interpreted as a vector or arrow starting at $(0,0)$ and ending at the point $(x_1, x_2)$. Intuitively, we may think of such a vector as having a physical interpretation, e.g., as a displacement, or a physical force acting on an object.

**Remark.** The words “coordinate axis”, “point”, “origin”, “vector” and 2-dimensional plane have an intuitive geometric meaning for us. However, from an abstract mathematical point of view, the way we use these words is simply a convention or a collection of definitions. We have simply decided to call the elements of $\mathbb{R}^2$ points or vectors, and we simply define the $x_1$-axis as the set of all pairs $(x_1,0)$ with $x_1 \in \mathbb{R}$ and the $x_2$-axis as the set of pairs $(0,x_2)$ with $x_2 \in \mathbb{R}$.

**The 3-Dimensional Coordinate Space $\mathbb{R}^3$**

Analogously, the 3-dimensional coordinate space $\mathbb{R}^3$ is defined as the set of all ordered triples $(x_1, x_2, x_3)$ of real numbers. We can interpret this as an idealized mathematical model of the ambient 3-dimensional space we live in, by choosing 3 coordinate axes. As before, we can view every element $(x_1, x_2, x_3)$ of $\mathbb{R}^3$ geometrically, either as a point with coordinates $x_1$, $x_2$, and $x_3$, or as the vector that ends at that point and starts at the origin $(0,0,0)$. Moreover, we can interpret vectors physically as describing, e.g., forces that act on an object in 3-space.

**The $n$-Dimensional Coordinate Space $\mathbb{R}^n$**

It should now be clear how to define $\mathbb{R}^4$ or $\mathbb{R}^{57}$ mathematically, even if we cannot immediately visualize these spaces anymore.
Definition 1.2. Fix a natural number \( n \geq 1 \). We define \( \mathbb{R}^n \) the set of all ordered \( n \)-tuples\(^4\) \((x_1, x_2, \ldots, x_n)\) of real numbers and call this set the standard \( n \)-dimensional vector space or \( n \)-dimensional coordinate plane. The elements of \( \mathbb{R}^n \) are called \( n \)-dimensional vectors or, more precisely, \( n \)-dimensional coordinate vectors, and the numbers \( x_i \) that appear in a vector \((x_1, x_2, \ldots, x_n)\) are called the components of the vector.

We will now start denoting \( n \)-dimensional vectors by a single symbol\(^5\) like \( x \) or \( y \) or \( u \) or \( v \) and write things like “Let \( v = (v_1, v_2, \ldots, v_n) \) be an \( n \)-dimensional vector.” In particular, we will use the same symbol 0 to denote the \( n \)-dimensional vector \((0, 0, \ldots, 0)\) with all components equal to zero, for every \( n \geq 1 \). Sometimes we will write \( 0 \in \mathbb{R}^n \) if \( n \) is not clear from the context.

Remark (Row Vectors and Column Vectors). Later on, when we discuss matrices and matrix multiplication, we will distinguish two ways of writing \( n \)-dimensional vectors, row vectors, which we write as a horizontal list or row of components

\[
x = (x_1, \ldots, x_n),
\]

and column vectors for which we write the components in a vertical column

\[
x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.
\]

For the time being, this distinction is not important and we continue with the less space-consuming row vector notation.

The physical interpretation of 2- or 3-dimensional vectors as forces or displacements that can be added motivates the following definition:

Definition 1.3 (Vector Addition and Scalar Multiplication in \( \mathbb{R}^n \)). We add \( n \)-dimensional vectors componentwise. In other words, if \( x = (x_1, x_2, \ldots, x_n) \) and \( y = (y_1, y_2, \ldots, y_n) \) are vectors in \( \mathbb{R}^n \) then their sum is defined as

\[
x + y := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).
\]

Here is an illustration in 2 dimensions:

\( ^4 \)By definition, an ordered \( n \)-tuple is a sequence of \( n \) entries in which entries may be repeated and the order in which the entries appear matters. In particular, \((x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_n)\) if and only if \( x_i = y_i \) for all \( 1 \leq i \leq n \).

\( ^5 \)Other commonly used notations for vectors are \( \vec{x} \) or \( \mathbf{x} \), but for simplicity, we will stick to simple, non-boldface symbols.
Similarly, we “stretch” vectors componentwise. More precisely, let $a$ be a real number, also called a scalar in this context. We define the product of the scalar $a$ with a vector $x = (x_1, x_2, \ldots, x_n)$ by multiplying each component of $x$ with $a$, i.e.,

$$ax := (ax_1, ax_2, \ldots, ax_n).$$

In particular, we introduce the notation

$$-x := (-1)x := (-x_1, -x_2, \ldots, -x_n),$$

which is “the vector $x$ reflected in the origin” (intuitively, we reverse the direction of the vector and translate it that it starts at the origin).

We summarize the most important properties of vector addition and scalar multiplication in $\mathbb{R}^n$.

**Proposition 1.4.** For all vectors $x, y, z$ in $\mathbb{R}^n$ and all scalars $a, b \in \mathbb{R}$, the following properties hold:

1) $x + y = y + x$

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

3) There exists a unique vector $0$ such that $0 + x = x + 0 = x$ for all vectors $x$.

4) For every vector $x$ there exists a unique vector $-x$ such that $x + (-x) = 0$

5) $a(bx) = (ab)x$

6) $1x = x$

7) $a(x + y) = ax + ay$

8) $(a + b)x = ax + bx$.

We will not give a detailed proof of this proposition here. Each property is fairly easy to verify, based on the fact that the operations are defined for each component separately and on the usual rules of addition and multiplication of real numbers, but doing all of this in all detail is quite tedious.
Here, we just remark that Property 2), sometimes called the associative law, says that we even though we have formally only defined the sum of two vectors, we can also form the sum \( x + y + z \) of three vectors by adding first forming the sum of two them and then adding the third one, and it does not matter in which order. Analogously, we can define the sum of four or more vectors by adding two of them at a time, in arbitrary order, e.g.,

\[
x + y + z + w = (x + y) + (z + w) = x + (y + (z + w)) = \ldots
\]

Abstract Vector Spaces

As indicated in the beginning of this section, we now abstraction one step further:

**Definition 1.5 (Abstract Vector Space).** An abstract vector space is a set \( V \) of elements, called **vectors**, on which two operations are defined, called **vector addition** and **scalar multiplication** that satisfy properties analogous to those listed in Proposition 1.4.

More precisely, we assume that for all vectors \( v, w \in V \) and for every scalar \( a \in \mathbb{R} \), there are uniquely defined vectors \( v + w \) and \( av \) such that the following conditions are satisfied for all scalars \( a, b \in \mathbb{R} \) and all vectors \( u, v, w \in V \):

1) \( v + w = w + v \)

2) \( u + (v + w) = (u + v) + w \)

3) There exists a unique vector \( 0 \in V \), called the zero vector, such that \( 0 + v = v + 0 = v \) for all vectors \( v \in V \).

4) For every vector \( v \in V \) there exists a unique vector \( -v \) such that \( v + (-v) = 0 \)

5) \( a(bv) = (ab)v \)

6) \( 1v = v \)

7) \( a(v + w) = av + aw \)

8) \( (a + b)x = ax + bx \).

**Remark.** We stress that such abstract vector spaces have, a priori, nothing to do with the standard coordinate spaces \( \mathbb{R}^n \). In particular, we do no longer assume that a vector is a sequence of numbers (its components).

Why do we bother with this abstraction (which, after all, requires some time and mental effort to get used to)? There are several reasons:

1) Formulating a few key properties and then basing our reasoning on these also focuses our attention and often (although not always) leads to simpler and more transparent arguments and proofs.

2) Furthermore, the abstraction forces us to no longer think of vectors as sequences of components but as geometric objects in themselves. Arguably, this actually strengthens the analogy with the low-dimensional situation in \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \), for which we have a geometric or physical intuition that makes it easy to think about a vectors independently of components and coordinates.
3) Even in situations where the vectors in a vector space can be described as sequences of components, such a description may not be obvious or unique. In particular, some descriptions may be “better” (easier to understand or to work with than others), and viewing vectors as abstract mathematical objects makes it easier to go back and forth between different descriptions (we will encounter this later under the keyword “change of coordinates”). It will also help us to clarify the notion of dimension.