Synopsis. This first lecture is just a bit of Linear Algebra backstory: As an introduction to the course, I thought to play with the structure of Euclidean space and linear algebra just to establish notation and begin the conversation. I also used a bit of Mathematica for visualization. It is listed on the course site.


**Real Euclidean Space** $\mathbb{R}^n$. The real plane is often described as the set of all ordered pairs of real numbers. We can write this as

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

The way the plane $\mathbb{R}^2$ is built out of two copies of the real line $\mathbb{R}$ is an example of a *Cartesian product*, a way of building a new set (called a product set) out of two sets, whose elements are pairs of elements of the two component sets, called factors, both $\mathbb{R}$ in this case. The set $\mathbb{R}^2$ is useful when studying functional relationships between sets because we can study the pairing given by the function as a subset living inside $\mathbb{R}^2$ by assigning the values of the input variable to the function $x$ to one of the ordered pairs, and the output variable $y = f(x)$ to the other (See Figure 1.1). This gives us a visual depiction of the functional relationship which facilitates the study of its properties.

We can construct a form of addition in the set $\mathbb{R}^2$ by using the notion of addition in $\mathbb{R}$ and forming an addition in $\mathbb{R}^2$ component-wise:

$$(a, b) + (c, d) = (a + c, b + d).$$

With this addition (and the identity element $(0, 0)$ and an inverse $(-a, -b)$ for every set element $(a, b)$), we can turn $\mathbb{R}^2$ into a group. Here we would call $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ the *direct product* of the two groups $\mathbb{R}$. (A direct product is a Cartesian product on the underlying sets with whatever added structure the individual sets have and give to the product.) We can also multiply elements of $\mathbb{R}^2$ by real numbers (scalars multiplication), where

$$c \cdot (a, b) = (ca, cb)$$

and these two notions behave well together.

Now $\mathbb{R}$ is also a field, but $\mathbb{R}^2$ is not: One cannot construct a good notion of multiplication in $\mathbb{R}^2$ that satisfies all of the field axioms. However, with the notion of addition of ordered pairs, along with scalar multiplication, we can give $\mathbb{R}^2$ the structure of a *vector space* over $\mathbb{R}$.
Definition 1.1. A linear or vector space over a field is a set \( V \) of objects together with two operations which can be added together and multiplied by field elements in a “compatible” way.

It is common, in a linear space, to call the individual set elements “vectors”. We also say that \( \mathbb{R}^2 \) is a vector space over \( \mathbb{R} \). But it will be a good idea to make a very important distinction:

Using Figure ?? as a guide, we will distinguish between points in \( \mathbb{R}^2 \), given by all 2-tuples of numbers written as

\[
\mathbb{R}^2 = \{ p = (x, y) \mid x, y \in \mathbb{R} \},
\]

and vectors in \( \mathbb{R}^2 \), denoted as the set of all possible \( 2 \times 1 \)-matrices, or 2-vectors

\[
\mathbb{R}^2 = \{ \mathbf{p} = \begin{bmatrix} x \\ y \end{bmatrix} \mid x, y \in \mathbb{R} \}.
\]

Some notes:

- Technically speaking, these two descriptions of the plane are quite different, even as there are equivalent. Note that this is a mathematical term that does need defining. For now we will leave it as is.
- In time, we will need to be able to define vectors based at arbitrary points in \( \mathbb{R}^2 \). Noticing a difference between points and vectors (with the same entries) as descriptions of the elements of the plane will help greatly later on when we define and understand vector fields.
- We can add still more structure to \( \mathbb{R}^2 \); a notion of a scalar product, sometimes called a dot product or an inner product on vectors (equivalently points):

\[
\begin{bmatrix} a \\ b \end{bmatrix} \cdot \begin{bmatrix} c \\ d \end{bmatrix} = ac + bd \in \mathbb{R}.
\]

With this new structure, the plane becomes an example of an inner product space. This is very useful for vector spaces, since with this new structure, we can define notions of a distance between vectors, a vector’s size, the angle between vectors, etc. And with these notions of measurement, the plane \( \mathbb{R}^2 \), as an inner product space, becomes a Euclidean Space (a space where one can do Euclidean geometry).
Absolutely all of this still works with \( n \)-tuples of numbers: Define, for \( n \in \mathbb{N} \),
\[
\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R} = \left\{ x = (x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R}, \text{ for } i = 0, 1, \ldots, n \right\}
\]
\[
= \left\{ x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mid x_i \in \mathbb{R}, \text{ for } i = 0, 1, \ldots, n \right\}.
\]

Now, a set of \( k \) \( n \)-vectors \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n \) are called \textit{linearly independent} if for real scalars \( c_i, i = 1, \ldots, k \),
\[
(1.1) \quad c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \ldots + c_k \mathbf{v}_k = 0
\]
is only solved by \( c_1 = c_2 = \ldots = c_k = 0 \). If this is true, then none of the vectors can be written as a linear combination of the others.

\textbf{Example 1.2.} \( \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \), \( \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \), and \( \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \) are \textit{linearly dependent} since \( 3\mathbf{v}_1 - \mathbf{v}_2 - \mathbf{v}_3 = 0 \). Thus, for instance, one can write \( \mathbf{v}_3 \) as a linear combination of the others;
\[
3\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{v}_3.
\]

If one can find \( n \) vectors that are linearly independent in \( \mathbb{R}^n \), then this set of \( n \) vectors can act as a \textit{basis}, in that any vector in \( \mathbb{R}^n \) can then be written as a linear combination of these. So if \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n \) are linearly independent (that is, if they form a basis), then
\[
\mathbb{R}^n = \text{span} \{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \} \\
= \left\{ \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mid \mathbf{x} = c_1 \mathbf{v}_1 + \ldots + c_n \mathbf{v}_n, \ c_i \in \mathbb{R} \right\}.
\]

Here, the term \textit{span} \{ \cdot \} is just the set of all linear combinations of...

An interesting side note: Using \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n \) as a basis for \( \mathbb{R}^n \), the lines through the origin formed by taking the set of all multiples of each vector \( \mathbf{v}_i \) can serve as axes for a coordinate system on \( \mathbb{R}^n \). Indeed, if each \( \mathbf{v}_i \) serves as a unit of measurement (a measuring stick) on the line that it helps to create, then the \( c_i \)'s in any linear combination of basis vectors are the coordinates in that coordinate system, and different from what would be considered the standard one.

\textbf{Example 1.3.} Construct the vectors
\[
\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.
\]

These vectors form a basis of \( \mathbb{R}^n \), since \( \mathbb{R}^n = \text{span} \{ \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \} \). This is called the \textit{standard basis} for \( \mathbb{R}^n \). See Figure ?? below.
Figure 1.3. The standard bases in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Note that these standard bases are used to define the equivalence between the notion of $\mathbb{R}^n$ defined as points and the notion of $\mathbb{R}^n$ defined as $n$-vectors.

**Definition 1.4.** A linear or vector subspace $W$ of a vector space $V$ is a subset of the elements of $V$ that satisfy

1. $0 \in W \subset V$,
2. If $w_1, w_2 \in W$, then $w_1 + w_2 \in W$, and
3. if $w \in W$, then for all $c \in \mathbb{R}$, $cw \in W$.

It is good to note here that ALL vector subspaces pass through the origin (contain the zero-vector).

And going back to Equation ??, note that for any $k \in \mathbb{N}$, the set of $n$-vectors span $\{v_1, v_2, \ldots, v_k\}$ is ALWAYS a linear subspace of $\mathbb{R}^n$. How big it is as a subspace depends on the number of $v_i$s are linearly independent.

**Example 1.5.** The set $\text{span} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$ is commonly referred to as the $xy$-plane in $\mathbb{R}^3$, thinking of the standard coordinates in $\mathbb{R}^3$. The span of these three vectors only makes a plane in three space since the third vector is simply twice the first plus 3/2 times the second. A basis for the span of these three 3-vectors can readily be the first two vectors in the standard basis of $\mathbb{R}^3$. Note that one can also call this linear subspace the $(z = 0)$-plane. In this way, the $xy$-plane is a version of $\mathbb{R}^2$ sitting inside $\mathbb{R}^3$ as a subspace of all vectors with 0 in the last component. See Figure ??.

**Example 1.6.** $\text{span} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$ is a line passing through the origin in $\mathbb{R}^3$.

**Example 1.7.** Let $V = \text{span} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$. Then $V \subset \mathbb{R}^3$ is a 2-dimensional subspace, since $a$ and $b$ are linearly independent (recall that the dimension of a (finite-dimensional) vector space is the number of elements in any basis), and $V \subset \mathbb{R}^3$ will look like
a plane passing through the origin (See Figure 1.5, with \( \mathbf{a} \) and \( \mathbf{b} \) in red). The two 3-vectors

\[
\mathbf{c} = \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}
\]

differ in that \( \mathbf{c} \) (shown in blue in Figure 1.5) is IN the plane \( V \), while \( \mathbf{d} \) (shown in green in the figure) is not. Indeed, \( \mathbf{c} = -\mathbf{a} + \mathbf{b} \), but there are not constants \( c_a, c_b \in \mathbb{R} \), where \( c_a \mathbf{a} + c_b \mathbf{b} = \mathbf{d} \). We would say that \( \mathbf{d} \) is linearly independent from \( V \).

Further, by Example 1.9, we can view the lines passing through \( \mathbf{a} \) and \( \mathbf{b} \) as coordinate axes for \( V \), and on each axis, use the length of the contained vector as the unit length along that axis, marking, for example, 0 at the origin and 1 at the head of \( \mathbf{a} \). This provides a coordinate system directly on \( V \), using the ordered pair \((c_a, c_b)\) as the coordinates in \( V \). Thus the vector \( \mathbf{c} \in V \subset \mathbb{R}^3 \) corresponds to the point \((3, 1, 2)\) \( \in \mathbb{R}^3 \), but in the coordinates defined directly on \( V \) by the basis \( \{\mathbf{a}, \mathbf{b}\} \), \( \mathbf{c} \in V \) corresponds to the point \((−1, 1)\) in the parameterization of \( V \) given by the basis. The idea of placing coordinates directly on a subspace instead of using the ambient coordinates of the larger space is an important one. We will spend much time on this.

One way to describe a subspace like \( V \in \mathbb{R}^3 \) is through another form of multiplication of vectors, this one where the product of two 3-vectors is again a 3-vector. (Note that this is extremely rare and for now is limited to \( \mathbb{R}^3 \).) The cross product of two vectors \( \mathbf{a} \times \mathbf{b} = \mathbf{n} \) is a vector normal (as in zero dot product) to both \( \mathbf{a} \) and \( \mathbf{b} \). Hence, for any vector \( \mathbf{n} \), the set of all vectors normal to \( \mathbf{n} \) is a two dimensional subspace \( V \in \mathbb{R}^3 \). And, if \( \mathbf{n} \) is given as the cross product of two linearly independent vectors \( \mathbf{a} \) and \( \mathbf{b} \), then \( \mathbf{a} \) and \( \mathbf{b} \) serve as a basis for \( V \). Indeed, endow \( \mathbb{R}^3 \) with the coordinates \( x, y, \) and \( z \). Then the equation

\[
\mathbf{n} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 = n_x x + n_y y + n_z z,
\]

defines a plane passing through the origin in \( \mathbb{R}^3 \). In Example 1.9, we have

\[
\mathbf{n} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \times \begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2(3) - 3(2) \\ -1(3) + 3(2) \\ 1(-2) - 2(2) \end{bmatrix} = \begin{bmatrix} 12 \\ 3 \\ -6 \end{bmatrix}.
\]

Thus the vector (sub)space \( V \) is defined

\[
V = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \bigg| 12x + 3y - 6z = 0 \right\}.
\]

Check for yourself that, for the vectors Example 1.9, \( \mathbf{a}, \mathbf{b}, \mathbf{c} \in V \), but \( \mathbf{d} \notin V \).
One conclusion that can be drawn from this is that one can define a plane in $\mathbb{R}^3$ via a single equation. But then, what is the equation of a line in $\mathbb{R}^3$? Here is an example:

**Example 1.8.** Consider the solution set for the set of equations:

\[
\begin{align*}
    x + 2y + 3z &= 4 \quad &\text{(eq1)} \\
    2x - 2y + 3z &= 1 \quad &\text{(eq2)} \\
\end{align*}
\]

2 equations in 3 unknowns.

So what does this solution set in $\mathbb{R}^3$ look like? To see, solve as best as one can:

\[
\begin{align*}
    (\text{eq1}) + (\text{eq2}) : & 3x + 6z = 5 \\
    2(\text{eq1}) - (\text{eq2}) : & 6y + 3z = 7
\end{align*}
\]

Then

\[
x = \frac{5 - 6z}{3}, \quad y = \frac{7 - 3z}{6}, \quad z \text{ is free.}
\]

Better yet, we can place a single parameter $t$ directly on this set by setting $z = t$, so that $x = \frac{5 - 6t}{3}$ and $y = \frac{7 - 3t}{6}$, along with $z = t$ makes a parameterized curve (a line) in $\mathbb{R}^3$. One could also write this as a function (using vector notation):

\[
c : \mathbb{R} \to \mathbb{R}^3, \quad c(t) = \begin{bmatrix} 5 - 6t \\ 7 - 3t \\ 6 \\ 3 \end{bmatrix}.
\]

Note that, in this parameterization, we still have 3 equations in 4 unknowns. Do you notice a pattern between the number of equations, the number of unknowns and the “size” of the space of solutions?

So, roughly speaking, a space $V$ is called linear if any linear combination of two elements in $V$ is still in $V$. So what, then, is a linear function?

**Definition 1.9.** A function $f : \mathbb{R} \to \mathbb{R}$ is called linear if

\[
f(c_1 x_1 + c_2 x_2) = c_1 f(x_1) + c_2 f(x_2), \quad \forall x_1, x_2 \in \mathbb{R}, \quad c_1, c_2 \in \mathbb{R}/
\]

Notes:

1. With appropriate changes, this works equally well for $f : \mathbb{R}^n \to \mathbb{R}^m$.
2. Using this definition, then, the function $f(x) = 3x$ is linear, but the function $g(x) = 3x + 1$ is NOT! To see this,

\[
g(2 + 3) = g(5) = 3(5) + 1 = 16 \\
\neq g(2) + g(3) = (3(2) + 1) + (3(3) + 1) = 17.
\]

The issue here is that for a function to be linear, the origin of the domain (the input space) must be mapped to the origin of the output space, so that $f(0) = 0$. But here $g(0) = 1$. And thus, $g(x)$ is not linear. It is an example of an affine function, one that can be seen as a composition of a linear function and a translation.

3. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be linear. Then, given a basis $\{v_1, \ldots, v_n\}$ for the domain $\mathbb{R}^n$, we can write any $x \in \mathbb{R}^n$ as

\[
x = c_1 v_1 + \ldots + c_n v_n.
\]
Then, since $f$ is linear, we have

$$f(x) = f(c_1v_1 + \ldots + c_nv_n) = c_1f(v_1) + \ldots + c_nf(v_n)$$

$$= m\left\{ \begin{bmatrix} f(v_1) & \ldots & f(v_n) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = A_{m \times n}x \right\}.$$ 

Hence, any linear map between vector spaces can always be represented by a matrix.