Algebraic structure of $\mathbb{R}^n$ (8.1)

Now we’re starting multivariable advanced calculus. We define $\mathbb{R}^n$ to be the set of $n$-tuples of reals. This is a vector space, in the sense that you can add two elements of $\mathbb{R}^n$ in a sensible way to get an element of $\mathbb{R}^n$, and you can multiply elements of $\mathbb{R}^n$ by scalars in a sensible way. (By sensible, I mean that the axioms of vector spaces are satisfied, but this is something that you saw in your linear algebra course.) We will sometimes think of elements of $\mathbb{R}^n$ as points in $n$-space, and sometimes as “physics-type” vectors, i.e., quantities with magnitude and direction.

We define the standard dot product (scalar product, inner product) of $\vec{x}$ and $\vec{y}$ as $x_1y_1 + \cdots + x_ny_n$, where $\vec{x} = (x_1, \ldots, x_n)$.

Theorem 8.2 is various algebraic properties of vector addition, dot product, scalar multiplication, etc., things like $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$.

I won’t go through the proofs.

We need to generalize absolute value of $\mathbb{R}$ to something in $\mathbb{R}^n$. The best way of thinking of absolute value is as distance: $|x|$ is the distance from the number $x$ to 0, and $|x - y|$ is the distance between $x$ and $y$ on the number line. In $\mathbb{R}^n$, the Euclidean distance between $\vec{x}$ and $\vec{y}$ is $\sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$, and this will give us our generalization.

Definition: The Euclidean norm $\|\vec{x}\|$ is defined to be $\sqrt{\sum_{i=1}^{n} x_i^2}$.

The Euclidean distance between $\vec{x}$ and $\vec{y}$ is $\|\vec{x} - \vec{y}\|$. Note that $\vec{x} \cdot \vec{x} = \|\vec{x}\|^2$.

So, what is this word “norm”?

Definition: a norm on $\mathbb{R}^n$ is a function $N : \mathbb{R}^n \to \mathbb{R}$ satisfying:

1. $N(\vec{x}) \geq 0$ for all $\vec{x} \in \mathbb{R}^n$, and $N(\vec{x}) = 0$ iff $\vec{x} = \vec{0}$.
2. $N(\alpha \vec{x}) = |\alpha|N(\vec{x})$ for all $\alpha \in \mathbb{R}$, $\vec{x} \in \mathbb{R}^n$.
3. $N(\vec{x} + \vec{y}) \leq N(\vec{x}) + N(\vec{y})$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$.

These are the properties of absolute value that we’ll want to generalize. We need to show that I’m not lying when I say “Euclidean norm”, i.e., that $\|\vec{x}\|$ is a norm and has the above three properties. (Note: we can
have norms defined on other vector spaces besides \( \mathbb{R}^n \), but we won’t be worrying about that.)

Proof of 1) is pretty clear. For 2):

\[
\|\alpha \vec{x}\| = \sqrt{\sum_{k=1}^{n} \alpha^2 x_k^2} = |\alpha| \sqrt{\sum_{k=1}^{n} x_k^2}.
\]

Part 3 is harder. For this we will need the following inequality:

**Theorem:** (Cauchy-Schwartz inequality) For all \( \vec{x}, \vec{y} \in \mathbb{R}^n \),

\[
|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|.
\]

**Proof:** The result is obvious if \( \vec{y} = \vec{0} \), so let’s assume that \( \vec{y} \neq \vec{0} \). Define a function of \( t \) by

\[
P(t) = \|\vec{x} - t\vec{y}\|^2.
\]

Clearly \( P(t) \geq 0 \) for all \( t \). In fact, it’s just a quadratic in \( t \):

\[
P(t) = \|\vec{x} - t\vec{y}\|^2
= (\vec{x} - t\vec{y}) \cdot (\vec{x} - t\vec{y})
= \vec{x} \cdot \vec{x} - 2t \vec{x} \cdot \vec{y} + t^2 \vec{y} \cdot \vec{y}
= \|\vec{x}\|^2 - 2t \vec{x} \cdot \vec{y} + t^2 \|\vec{y}\|^2.
\]

Where does this have its minimum? \( P'(t) = -2\vec{x} \vec{y} + 2t \|\vec{y}\|^2 \), so the minimum occurs at \( t = \frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\|^2} \). Crucially, it’s still non-negative at that point:

\[
0 \leq P\left( \frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\|^2} \right) = \|\vec{x}\|^2 - 2 \left( \frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\|^2} \right)^2 + \frac{(\vec{x} \cdot \vec{y})^2}{\|\vec{y}\|^4} \|\vec{y}\|^2,
\]

hence

\[
0 \leq \|\vec{x}\|^2 \|\vec{y}\|^2 - (\vec{x} \cdot \vec{y})^2,
\]

giving

\[
|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|,
\]

as desired.

Now to prove the triangle inequality:

**Theorem:** \( \|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \) holds for all \( \vec{x}, \vec{y} \) in \( \mathbb{R}^n \).

**Proof:**

\[
\|\vec{x} + \vec{y}\|^2 = (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y})
= \|\vec{x}\|^2 + 2\vec{x} \cdot \vec{y} + \|\vec{y}\|^2
\leq \|\vec{x}\|^2 + 2\|\vec{x}\| \|\vec{y}\| + \|\vec{y}\|^2
\leq \|\vec{x}\|^2 + 2\|\vec{x}\| \|\vec{y}\| + \|\vec{y}\|^2 = (\|\vec{x}\| + \|\vec{y}\|)^2
\]

and take square roots to get the result.

©2012, T. Vogel
There are other norms one can define on \( \mathbb{R}^n \). Here are two:

- \( l_1 \) norm: \( \| \vec{x} \|_1 = \sum_{k=1}^{n} |x_k| \).
- \( l_\infty \) norm, or sup norm: \( \| \vec{x} \|_\infty = \max (|x_1|, \ldots, |x_n|) \).

The Euclidean norm can also be called the \( l_2 \) norm. Although the Euclidean norm is by far the most important, we will occasionally use the other two. We’ll need to see how they relate.

**Proposition:** For any \( \vec{x} \in \mathbb{R}^n \) and for any \( j \),

\[
|x_j| \leq \| \vec{x} \| \leq \sqrt{n} \| \vec{x} \|_\infty.
\]

**Proof:** The first inequality is clear:

\[
|x_j| = \sqrt{x_j^2} \leq \sqrt{\sum_{j=1}^{n} x_j^2} = \| \vec{x} \|.
\]

For the second, suppose that \( k \) is such that \( |x_k| = \max (|x_1|, \ldots, |x_n|) \). Then

\[
\sqrt{\sum_{j=1}^{n} x_j^2} \leq \sqrt{n x_k^2} = \sqrt{n} |x_k| = \sqrt{n} \| \vec{x} \|_\infty.
\]

**Proposition:** For any \( \vec{x} \in \mathbb{R}^n \), \( \| \vec{x} \| \leq \| \vec{x} \|_1 \).

**Proof:**

\[
\| \vec{x} \|_1^2 = (|x_1| + \cdots + |x_n|)^2 \\
= |x_1|^2 + 2|x_1||x_2| + \cdots + |x_n|^2 \\
\geq |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2,
\]

after throwing out a lot of positive cross terms. But the last line is \( \| \vec{x} \|_1^2 \).

Recall from calculus, that if \( \theta \) is the angle between two vectors \( \vec{a} \) and \( \vec{b} \) in \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \), then

\[
\cos \theta = \frac{\vec{a} \cdot \vec{b}}{\| \vec{a} \| \| \vec{b} \|}.
\]

The proof is generally skipped at that level, but is pretty easy. The angle \( \theta \) is one angle in a triangle with sides of length \( \| \vec{a} \| \), \( \| \vec{b} \| \), and \( \| \vec{b} - \vec{a} \| \). (picture). By the law of cosines,

\[
\| \vec{b} - \vec{a} \|^2 = \| \vec{a} \|^2 + \| \vec{b} \|^2 - 2\| \vec{a} \| \| \vec{b} \| \cos \theta.
\]

But

\[
\| \vec{b} - \vec{a} \|^2 = (\vec{b} - \vec{a}) \cdot (\vec{b} - \vec{a}) = \| \vec{a} \|^2 + \| \vec{b} \|^2 - 2\vec{a} \cdot \vec{b}
\]

by expanding out the dot product, so the result follows.