Introduction to Normed Vector Spaces

Audrey Terras

March 29, 2009

1 Why worry about infinite dimensional normed vector spaces?

We want to understand the integral from Lang's perspective rather than that of your calculus book. Secondly we want to understand convergence of series of functions - something that proved problematic for Cauchy in the 1800s. These things are important for many applications in physics, engineering, statistics. We will be able to study vibrating things such as violin strings, drums, buildings, bridges, spheres, planets, stock values. Quantum physics, for example, involves Hilbert space, which is a type of normed vector space with a scalar product where all Cauchy sequences of vectors converge.

The theory of such normed vector spaces was created at the same time as quantum mechanics - the 1920s and 1930s. So with this chapter of Lang you are moving ahead hundreds of years from Newton and Leibnitz, perhaps 70 years from Riemann.

Fourier series involve orthogonal sets of vectors in an infinite dimensional normed vector space:

$$C[a,b] = \{f : [a,b] \to \mathbb{R} \,|\, f \;\; \text{continuous}\}.$$

The L^2-norm of a continuous function f in $C[a,b]$ is

$$\|f\|_2 = \left(\int_a^b |f(x)|^2 \, dx \right)^{1/2}.$$

This is an analog of the usual idea of length of a vector $f = (f(1), ..., (f(n)) \in \mathbb{R}^n$:

$$\|f\|_2 = \left(\sum_{j=1}^{n} |f(j)|^2 \right)^{1/2}.$$

There are other natural norms for $f \in C[a,b]$ such as:

$$\|f\|_1 = \int_a^b |f(x)| \, dx.$$

$$\|f\|_\infty = \max_{a \le x \le b} |f(x)|.$$

On **finite** dimensional vector spaces such as $\mathbb{R}^n$ it does not matter what norm you use when you are trying to figure out whether a sequence of vectors has a limit. However, in infinite dimensional normed vector spaces convergence can disappear if a different norm is used. Not all norms are equivalent in infinite dimensions. See Homework 9, problem 6.

Note that $C[a,b]$ is infinite dimensional since the set $\{1, x, x^2, x^3, ..., x^n, ...\}$ is an infinite set of linearly independent vectors. Prove this as follows. Suppose that we have a linear dependence relation $\sum_{j=0}^{n} c_j x^j = 0$, for all x in $[a,b]$. This implies all the constants $c_j = 0$. **Why? Extra Credit. Prove this.**

Infinite dimensional vector spaces are thus more interesting than finite dimensional ones. Each (inequivalent) norm leads to a different notion of convergence of sequences of vectors.

2 What is a Normed Vector Space?

In what follows we define normed vector space by 5 axioms. We will not put arrows on our vectors. We will try to keep vectors and scalars apart by using Greek letters for scalars. Our scalars will be real. Maybe by the end of next quarter we may allow complex scalars. It simplifies Fourier series.

Definition 1 *A **vector space** V is a set of vectors $v \in V$ which is closed under addition and closed under multiplication by scalars $\alpha \in \mathbb{R}$. This means $u + v \in V$ and $\alpha v \in V$ and the following 5 axioms must hold for all $u, v, w \in V$ and all $\alpha, \beta \in \mathbb{R}$:*

VS1. $u + (v + w) = (u + v) + w$
VS2. $\exists\ 0 \in V$ *s.t.* $0 + v = v$
VS3. $\forall\ v \in V\ \exists\ v' \in V$ *s.t.* $v + v' = 0.$
VS4. $v + u = u + v$
VS5. $1v = v,\ \ \alpha(\beta v) = (\alpha\beta)v,\ (\alpha + \beta)v = \alpha v + \beta v,\ \ \alpha(u + v) = \alpha u + \alpha v$

You may say we cheated by putting 4 axioms into VS5.

Definition 2 *A vector space V is a **normed vector space** if there is a norm function mapping V to the non-negative real numbers, written $\|v\|$, for $v \in V$, and satisfying the following 3 axioms:*

$$N1.\ \ \|v\| \geq 0\ \forall v \in V\ \ \textit{and}\ \ \|v\| = 0\ \ \text{if and only if}\ \ v = 0.$$
$$N2.\ \ \|\alpha v\| = |\alpha|\,\|v\|,\ \ \forall v \in V\ \ \text{and} \forall \alpha \in \mathbb{R}.\ \text{Here}\ |\alpha| = \text{absolute value of}\ \alpha.$$
$$N3.\ \ \|u + v\| \leq \|u\| + \|v\|,\ \forall u, v \in V.\quad \text{Triangle Inequality.}$$

Definition 3 *The **distance between 2 vectors** u, v in a normed vector space V is defined by*
$d(u, v) = \|u - v\|$.

Example 1. 3-Space.

$$\mathbb{R}^3 = \left\{ \left(\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right) \middle| x_1, x_2, x_3 \in \mathbb{R} \right\}.$$

Define addition and multiplication by scalars as usual:

$$\left(\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right) + \left(\begin{array}{c} y_1 \\ y_2 \\ y_3 \end{array} \right) = \left(\begin{array}{c} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{array} \right).$$
$$\alpha \left(\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right) = \left(\begin{array}{c} \alpha x_1 \\ \alpha x_2 \\ \alpha x_3 \end{array} \right),\ \ \forall\ \alpha \in \mathbb{R}.$$

The usual norm is

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + x_3^2}\ \ if x = \left(\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right).$$

Other norms are possible:

$$\|x\|_2 = |x_1| + |x_2| + |x_3| \qquad \text{or} \qquad \|x\|_\infty = \max\{|x_1|, |x_3|, |x_3|\}.$$

The proof that these definitions make $\mathbb{R}^3$ a normed vector space is tedious. No doubt we will make it a homework problem.

Example 2. The space of continuous functions on an interval.

$$C[a,b] = \{f : [a,b] \to \mathbb{R} \,|\, f \;\; \text{continuous}\}\,.$$

For $f, g \in C[a,b]$, define $(f+g)(x) = f(x) + g(x)$ for all $x \in [a,b]$ and define for $\alpha \in \mathbb{R}$ $(\alpha f)(x) = \alpha f(x)$ for all $x \in [a,b]$. We leave it as an Exercise (see Homework 9) to check the axioms for a vector space. The most non-trivial one is the one that says $f+g$ and αf are both continuous functions on $[a,b]$.

Again there are many possible norms. We will look at 3:

$$\|f\|_2 = \left(\int_a^b |f(x)|^2 \, dx \right)^{1/2}.$$

$$\|f\|_1 = \int_a^b |f(x)| \, dx.$$

$$\|f\|_\infty = \max_{a \le x \le b} |f(x)|\,.$$

Most of the axioms for norms are easy to check. Let's do it for the $\|f\|_1$ norm.
N1. $\|v\| \ge 0 \;\; \forall \, v \in V$ and $\|v\| = 0$ if and only if $v = 0$.
N2. $\|\alpha v\| = |\alpha| \, \|v\|\,, \;\; \forall \, v \in V$ and $\forall \alpha \in \mathbb{R}$.
N3. Triangle Inequality. $\|u + v\| \le \|u\| + \|v\|\,, \;\; \forall \, u, v \in V$.

Proof of N1.

Since $|f(x)| \ge 0$ for all x we know that the integral is ≥ 0, because the integral preserves inequalities (Lang, Thm. 2.1, p. 104 or earlier in these notes). Suppose that $\int_a^b |f(x)| \, dx = 0$. Since f is continuous by Theorem 2.4 of Lang, p. 104 or these notes near Figure 1, this implies $f(x) = 0$ for all $x \in [a,b]$.

Proof of N2.

Also for any $\alpha \in \mathbb{R}$ and $f \in C[a,b]$, we have $\|\alpha f\|_1 = \int_a^b |\alpha f(x)| \, dx = \int_a^b |\alpha| \, |f(x)| \, dx = |\alpha| \int_a^b |f(x)| \, dx = |\alpha| \, \|f\|_1$. This proves N2 for norms. Here we used the multiplicative property of absolute value as well as the linearity of the integral (i.e., scalars come out of the integral lecture notes p. 80).

Proof of N3.

Using the definition of the 1-norm, and the triangle inequality for real numbers as well as the fact that the integral preserves $\le$, we see that

$$\|f + g\|_1 = \int_a^b |f(x) + g(x)| \, dx \le \int_a^b (|f(x)| + |g(x)|) \, dx.$$

To finish the proof, use the linearity of the integral to see that

$$\int_a^b (|f(x)| + |g(x)|) \, dx = \int_a^b |f(x)| \, dx + \int_a^b |g(x)| \, dx = \|f\|_1 + \|g\|_1\,.$$

Putting it all together gives $\|f + g\|_1 \le \|f\|_1 + \|g\|_1$.

3 Scalar Products.

You have seen the dot (or scalar or inner) product in $\mathbb{R}^3$. It is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \bullet \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = x_1y_1 + x_2y_2 + x_3y_3.$$

It turns out there is a similar thing for $C[a, b]$. First let's define the scalar product on a vector space and see how to get a norm if in addition the scalar product is positive definite.

Definition 4 *A (positive definite)* ***scalar product*** $< v, w >$ *for vectors* v, w *in a vector space* V *is a real number* $< v, w >$ *such that the following axioms hold:*

$$\begin{array}{lll} \text{SP1.} & < v, w >=< w, v >, \ \forall \ v, w \in V & \text{(symmetry)} \\ \text{SP2.} & < u, v + w >=< u, v > + < u, w >, \ \forall \ u, v, w \in V & \\ \text{SP3.} & < \alpha v, w >= \alpha < v, w >, \ \forall \ v, w \in V \ \text{ and } \ \forall \alpha \epsilon \mathbb{R} & \\ \text{SP4.} & < v, v > \geq 0, \forall \ v \in V \ \text{ and } \ < v, v >= 0 \iff v = 0 & \textit{(positive definite)} \end{array}$$

Axioms SP1,2,3 say $< v, w >$ is **linear** in each variable holding the other variable fixed. Axiom SP4 says the scalar product is **positive definite**. We will always want to assume SP4 because we want to be able to get a norm out of the scalar product.

Definition 5 *If* V *is a vector space with a (positive definite) scalar product* $< v, w >$ *for* $v, w \in V$*, define the associated* ***norm*** *by* $\|v\| = \sqrt{< v, v >}$*, for all* $v \in V$

Before proving that this really gives a norm, let's look at some examples.
Example 1. In $\mathbb{R}^3$ the scalar product is

$$< x, y >= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \bullet \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = x_1y_1 + x_2y_2 + x_3y_3.$$

It is easy to check the axioms. For example, the positive definiteness follows from the fact that squares of real numbers are ≥ 0 and sums of non-negative numbers are non-negative:

$$< x, x >= x_1^2 + x_2^2 + x_3^2 \geq 0.$$

And $0 =< x, x > \ \geq x_i^2$ implies $x_i = 0$ for all i and thus $x = 0$.

Example 2.

$$C[a, b] = \{f : [a, b] \to \mathbb{R} \mid \ f \ \text{ is } \ \text{continuous}\}$$

For f, g in $C[a, b]$, define the **scalar product** by

$$< f, g >= \int_a^b fg.$$

Once more, it is not hard to use the properties of the integral to check axioms SP1,SP2,SP3 (extra credit exercise). To see SP4, note that $f(x)^2 \geq 0$ for all $x \in [a, b]$ implies by the fact that integrals preserve $\geq$ that

$$< f, f >= \int_a^b f(x)^2 dx \geq 0.$$

Now suppose that $< f, f >= 0$. Then by the positivity property of the integral, we know that $f(x)^2 = 0$ for all $x \in [a, b]$ which says that f is the 0 function (the identity for addition in our vector space $C[a, b]$).

Then the norm associated to this scalar product is $\|f\|_2 = \left(\int_a^b |f(x)|^2 \, dx \right)^{1/2}$.

The following theorem is so useful people from lots of countries got their names attached.

Theorem 6 ***Cauchy-Schwarz (Bunyakovsky) Inequality***

Suppose that V *is a vector space with scalar product* $< v, w >$. *Then using our definition of the norm* $\|v\| = \sqrt{< v, v >}$, *we have for all* $v, w \in V$:

$$|< v, w >| \leq \|v\| \, \|w\| .$$

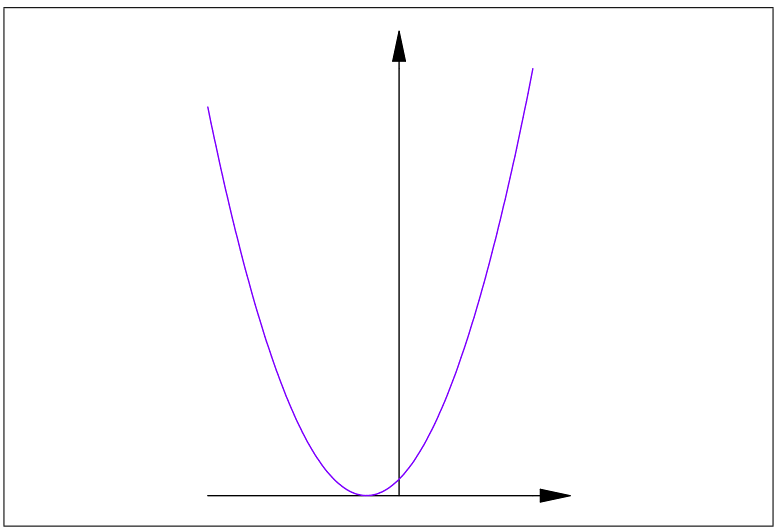
Proof. (Compare Problem 8 in Homework 8).

Let $t \in \mathbb{R}$ and look at

$$f(t) =< v + tw, v + tw > .$$

By properties of the scalar product, we have $0 \leq f(t) =< v, v > +2t < v, w > +t^2 < w, w >$.

As a function of f, we see that $f(t) = At^2 + Bt + C$, where $A =< w, w >$, $B = 2 < v, w >$ and $C =< v, v >$. So the graph of $f(t)$ is that of a parabola above or touching the t-axis. For example, we have drawn a parabola touching the t-axis at one point.



Recall the quadratic formula for the roots $r_\pm$ of $f(t) = At^2 + Bt + C = 0$,

$$r_\pm = \frac{B \pm \sqrt{B^2 - 4AC}}{2A}.$$

Since we have at most one real root it follows that

$$B^2 - 4AC \leq 0.$$

Now plug in $A =< w, w >$, $B = 2 < v, w >$ and $C =< v, v >$. This gives the Cauchy-Schwarz inequality. ∎

Corollary 7 *Under the hypotheses of the preceding theorem,* $\|v\| = \sqrt{< v, v >}$ *defines a norm on* V.

Proof. We must prove:

N1. $\|v\| \geq 0 \ \ \forall \, v \in V$ and $\|v\| = 0$ if and only if $v = 0$.

N2. $\|\alpha v\| = |\alpha| \, \|v\|$, $\forall \, v \in V$ and $\forall \alpha \in \mathbb{R}$.

N3. Triangle Inequality. $\|u + v\| \leq \|u\| + \|v\|$, $\forall \, u, v \in V$.

We get N2 from SP3.For then $\|\alpha v\|^2 =< \alpha v, \alpha v >= \alpha^2 < v, v >= |\alpha|^2 \|v\|^2$, $\forall \ v \in V$ and $\forall \alpha \epsilon \mathbb{R}$.

We get N1 from SP4. This says $\|v\|^2 = <v,v> \geq 0, \forall \ \ v \in V$ and $\|v\|^2 = <v,v> = 0 \iff v = 0$.

To prove the triangle inequality N3, we need to use the Cauchy-Schwarz inequality. This proof goes as follows. By the linearity and symmetry of the scalar product we see that

$$\begin{aligned} \|u+v\|^2 &= <u+v, u+v> = <u,u> +2 <u,v> + <v,v> \\ &= \|u\|^2 + 2<u,v> + \|v\|^2 \\ &\leq \|u\|^2 + 2\,|<u,v>| + \|v\|^2 \quad \text{as} \ \ x \leq |x| \\ &\leq \|u\|^2 + 2\,\|u\|\,\|v\| + \|v\|^2 \ \ \text{by} \ \ \text{Cauchy} - \text{Schwarz} \\ &= (\|u\| + \|v\|)^2 . \end{aligned}$$

Now use the fact that the square root $\surd$ preserves inequalities to finish the proof of the triangle inequality. ■

What's the good of all this? Now we can happily definite limits of sequences of vectors $\{v_n\}$in our normed vector space V. **Can you guess the definition of** $\lim\limits_{n\to\infty} v_n = L \in V$?

Answer: $\forall \varepsilon > 0$, $\exists N_\varepsilon \in \mathbb{Z}^+$ s.t. $n \geq N_\varepsilon$ implies $\|v_n - L\| < \varepsilon$. That is, just replace absolute value in the old definition of limit with the norm.

Similarly we can define a Cauchy sequence $\{v_n\}$ in the normed vector space V.

Another use of the scalar product is to define orthogonal vectors in a vector space V with a scalar product.

Definition 8 *Two vectors $v, w \in V$, a vector space V with scalar product $<,>$, are defined to be* ***orthogonal*** *if the scalar product $<v,w> = 0$.*

In a vector space with scalar product, you can also **define the angle** θ **between 2 vectors** $v, w \in V$, by

$$<v,w> = \|v\|\,\|w\| \cos\theta.$$

You can draw the same picture you would draw in the plane since 2 vectors determine a plane and then use the cosine law from high school trig to see this.

What is the cosine law? Using the triangles in Figure 2, it says

$$\|v-w\|^2 = \|v\|^2 - 2\,\|v\|\,\|w\|\cos\theta + \|w\|^2 .$$

You also need to see, using the axioms for scalar product, that

$$\|v-w\|^2 = <v-w, v-w> = \|v\|^2 - 2<v,w> + \|w\|^2 .$$

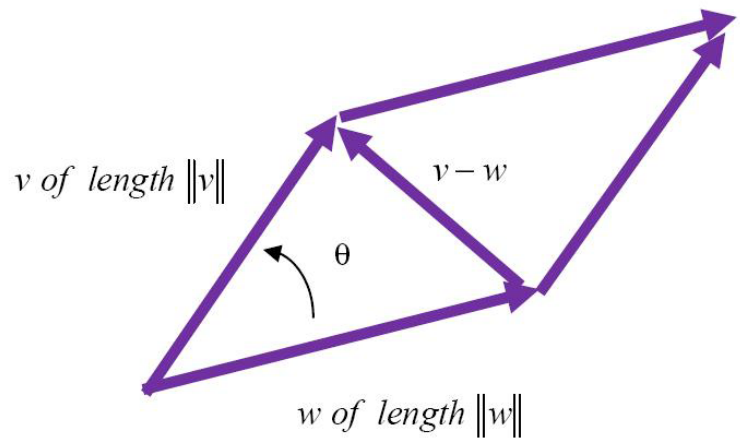


Figure 1: Visualizing vectors in a normed vector space and the angle between them.

4 Comparison of Norms

Definition 9 *Suppose $\{v_n\}$ is a sequence in a normed vector space V with norm $\|\|$, We say*

$$\lim_{n\to\infty} v_n = L$$

and "v_n converges to $L \in V$" $\Longleftrightarrow$

$$\lim_{n\to\infty} \|v_n - L\| = 0.$$

Note that the last limit is a sequence of real numbers.

Question. We know that there are lots of norms on V. How can we guarantee that 2 different norms $\|\|_\alpha$ and $\|\|_\beta$ produce the same convergent sequences in V?

The answer is that equivalent norms produce the same convergent sequences where we define equivalent as follows.

Definition 10 *2 norms $\|\|_\alpha$ and $\|\|_\beta$ on the vector space V are equivalent iff there are constants $A, B > 0$ such that for all $v \in V$ we have*

$$A \|v\|_\alpha \leq \|v\|_\beta \leq B \|v\|_\alpha .$$

In the preceding definition we are assuming that the constants A and B are independent of $v \in V$.

Why do equivalent norms lead to the same convergent sequences?

Answer. Suppose $\{v_n\}$ is a convergent sequence for the $\|\|_\alpha$-norm; i.e., for some $L \in V$ we have $\lim\limits_{n\to\infty} \|v_n - L\|_\alpha = 0$. And suppose $\|\|_\beta$ is an equivalent norm. Then

$$A \|v_n - L\|_\alpha \leq \|v_n - L\|_\beta \leq B \|v_n - L\|_\alpha .$$

Since the outside sequences go to 0 as $n \to \infty$, it follows by the squeeze lemma that the guy in the middle has to go to 0 as well. For limits preserve inequalities and we would have

$$0 \leq \lim_{n\to\infty} \|v_n - L\|_\beta \leq 0.$$

which implies $\lim\limits_{n\to\infty} \|v_n - L\|_\beta = 0$.

Similarly $\lim\limits_{n\to\infty} \|v_n - L\|_\beta = 0$ implies $\lim\limits_{n\to\infty} \|v_n - L\| = 0$.

Moral. It does not matter which of 2 equivalent norms you use to test a sequence for convergence.

Theorem 11 *All norms on $\mathbb{R}^n$ are equivalent.*

Proof. See Lang p. 145. ■

Thus, for our purposes, it does not matter which norm you use on finite dimensional vector spaces. You get the same definition of convergence of sequences. However, things are very different for infinite dimensional vector spaces. Look at Exercise 6 in Homework 9 for example. There you see a sequence of functions f_n in $C[0,1]$ such that

$$\lim_{n\to\infty} \|f_n - 0\|_1 = 0 \quad \text{using the norm} \quad \|f\|_1 = \int_0^1 |f(x)|\, dx.$$

but

$$\|f_n - 0\|_\infty = 1 \text{ for all n, using the norm} \quad \|f\|_\infty = \max_{a\leq x\leq b} |f(x)| .$$

It follows that the norms $\|\|_1$ and $\|\|_\infty$ on $C[0,1]$ are not equivalent.

Extra Credit Exercise. Using the same example, show that $\lim\limits_{n\to\infty} \|f_n - 0\|_2 = 0$ using the norm $\|f\|_2 = \left(\int_0^1 |f(x)|^2 \, dx \right)^{1/2}$.

It follows that the norms $\|\|_2$ and $\|\|_\infty$ on $C[0,1]$ are not equivalent.

Proposition 12 *The norms* $\|f\|_1 = \int_a^b |f(x)|\,dx$ *and* $\|f\|_2 = \left(\int_a^b f(x)^2 dx\right)^{1/2}$ *are not equivalent. However, we do have the inequality*

$$\|f\|_1 \le \sqrt{b-a}\,\|f\|_2\,.$$

Proof. To prove the inequality, use the Cauchy-Schwarz inequality on the functions $|f|$ and $g(x) = 1$ for all $x \in [a,b]$. This gives

$$|< f, g >| \le \|f\|_2\,\|g\|_2\,.$$

So we have $|< |f|\,, 1 >| \le \||f|\|_2\,\|1\|_2$. Now this really means

$$\begin{aligned}\|f\|_1 &= \int_a^b |f(x)|\,dx \le \left(\int_a^b f(x)^2 dx\right)^{1/2} \left(\int_a^b 1 dx\right)^{1/2} \\ &= \sqrt{b-a}\left(\int_a^b f(x)^2 dx\right)^{1/2} = \sqrt{b-a}\,\|f\|_2\,.\end{aligned}$$

To see that $\|\|_1$ and $\|\|_2$ are not equivalent norms, we take $a = 0$ and $b = 1$. Then we look at the following example. Define as in Lang, p. 147:

$$g_n(x) = \begin{cases} \sqrt{n}, & \text{for } 0 \le x \le \frac{1}{n}; \\ \frac{1}{\sqrt{x}}, & \text{for } \frac{1}{n} \le x \le 1. \end{cases}$$

Note that g_n is continuous. Why? **Extra Credit Exercise.** Show

$$\|g_n\|_1 = 2 - \frac{1}{\sqrt{n}}$$

and

$$\|g_n\|_2 = \sqrt{1 + \log n}.$$

It follows that there cannot be a constant $C > 0$ such that $\|v\|_2 \le C\,\|v\|_1$ at least on the interval $[0,1]$. Can you extend this idea to arbitrary intervals $[a,b]$? ■