Linear maps

Throughout these lecture notes, when a result is stated without proof, there is an implied exercise to work out the proof.

Before we can study derivatives of functions of several variables, we need to first do some analysis of linear maps (functions) between normed spaces. Although I’m assuming that you have a familiarity with linear algebra, and that you have seen at least a little bit about normed spaces, we’ll do a quick review, including a couple of things you might not have seen.

Unless otherwise specified, \(X, Y, Z, \ldots\) will denote normed spaces\(^1\). We will always take the scalar field to be the real numbers, although most of what we do carries over essentially verbatim to complex scalars. We will see that if \(X\) is \(n\)-dimensional, then \(X\) is “essentially the same” as \(\mathbb{R}^n\) (and this will be made precise.)

Recall that a norm on a vector space \(X\) is a function \(x \mapsto \|x\|\) from \(X\) to \(\mathbb{R}\) such that:

- \(\|x\| \geq 0\), and \(\|x\| = 0\) if and only if \(x = 0\);
- \(\|cx\| = |c|\|x\|\) if \(c \in \mathbb{R}\);
- \(\|x + y\| \leq \|x\| + \|y\|\).

**Example 1.** On \(\mathbb{R}^n\), the Euclidean norm on \(\mathbb{R}^n\) is given by

\[
\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2},
\]

the 1-norm (or \(\ell^1\)-norm) is

\[
\|x\|_1 = \sum_{i=1}^{n} |x_i|,
\]

and the \(\infty\)-norm (or \(\ell^\infty\)-norm, or max-norm) is

\[
\|x\|_\infty = \max_i |x_i|.
\]

The default for \(\mathbb{R}^n\) will be the Euclidean norm. \(\mathbb{R}^n\) with the Euclidean norm, the 1-norm, and the \(\infty\)-norm are sometimes denoted respectively by \(\ell^2_n\), \(\ell^1_n\), and \(\ell^\infty_n\).

**Definition 2.** Let \((x_n)\) be a sequence in \(X\). For \(k \in \mathbb{N}\) define \(s_k = \sum_{n=1}^{k} x_n\). The series with \(n\)th term \(x_n\) and \(k\)th partial sum \(s_k\) is the sequence \((s_k)\), and is written \(\sum_{n=1}^{\infty} x_n\). A series \(\sum_{n=1}^{\infty} x_n\) in \(X\) converges absolutely if \(\sum_{n=1}^{\infty} \|x_n\|\) converges in \(\mathbb{R}\).

Recall that a Banach space is a complete normed space. In a Banach space, every absolutely convergent series converges, because the sequence of partial sums is Cauchy. It is important to know that the converse is also true:

\(^1\)usually, but not always, finite-dimensional
Proposition 3. Let $X$ be a normed space in which every absolutely convergent series converges. Then $X$ is complete.

Proof. Let $(s_k)$ be a Cauchy sequence in $X$. It suffices to show that $(s_k)$ has a convergent subsequence. Choose $k_1$ such that $\|s_j - s_l\| < 2^{-1}$ for all $j, l \geq k_1$. Next choose $k_2 > k_1$ such that $\|s_j - s_l\| < 2^{-2}$ for all $j, l \geq k_2$. Continue inductively, getting a subsequence $(s_{k_j})$ such that

$$\|s_{k_{j+1}} - s_{k_j}\| < 2^{-j} \quad \text{for all } j.$$

Define

$$x_n = \begin{cases} s_{k_1} & \text{if } n = 1 \\ s_{k_n} - s_{k_{n-1}} & \text{if } n > 1. \end{cases}$$

Then $(s_{k_j})$ is the sequence of partial sums of the series $\sum_{n=1}^{\infty} x_n$. By the above, this series converges absolutely, hence so does the subsequence $(s_{k_j})$ of $(s_k)$. □

Notation and Terminology. If $T : X \to Y$ is linear, we write $Tx = T(x)$, and if $S : Y \to Z$ is also linear, we write $ST$ for $S \circ T$.

Definition 4. The operator norm of a linear map $T$ is

$$\|T\| := \sup \{\|Tx\| : \|x\| = 1\},$$

and $T$ is bounded if $\|T\|$ is finite.

Warning: “boundedness” for a linear map is different from the usual notion of boundedness for a map between arbitrary metric spaces. Unless a linear map is $0$, its range is a nontrivial subspace, so can’t be a bounded subset of the codomain normed space. For linear maps it’s only boundedness on the unit sphere that is of interest. We’ll soon see that boundedness is automatic in finite dimensions.

Lemma 5. If $T$ is a bounded linear map then

$$\|T\| = \sup\{\|Tx\| : \|x\| \leq 1\} = \min\{c \geq 0 : \|Tx\| \leq c\|x\| \text{ for all } x \in X\}.$$ 

Proof. The first equality is left as an exercise. For the second, put

$$S = \{c \in \mathbb{R} : \|Tx\| \leq c\|x\| \text{ for all } x \in X\}.$$

If $c \in S$ then $\|Tx\| \leq c$ for all $\|x\| = 1$, so $\|T\| \leq c$. Thus $\|T\|$ is a lower bound for $S$, so it remains to verify that $\|T\| \in S$, and it suffices to check the inequality for $x \neq 0$:

$$\|Tx\| = \|x\| \left\| T \left( \frac{x}{\|x\|} \right) \right\| \leq \|x\| \|T\|.$$

□

Corollary 6. Let $T : X \to Y$ be linear. Then $T$ is continuous if and only if it is bounded.

Proof. First assume that $T$ is continuous. Choose $\delta > 0$ such that if $\|x\| < \delta$ then $\|Tx\| < 1$. Then for $\|x\| = 1$ we have $\|(\delta/2)x\| = \delta/2 < \delta$, so

$$1 > \left\| T \left( \frac{\delta}{2} x \right) \right\| = \frac{\delta}{2} \|Tx\|,$$

and hence $T$ is bounded.
Conversely, assume that $T$ is bounded. Then for all $x, y \in X$ we have

$$\|Tx - Ty\| = \|T(x - y)\| \leq \|T\|\|x - y\|,$$

so $T$ is continuous. \(\square\)

**Corollary 7.** If $T : X \to Y$ is a linear isomorphism (i.e., is 1-1 onto), then $T^{-1}$ is bounded if and only if

$$\inf \{\|Tx\| : \|x\| = 1\} > 0.$$

In general, if a proof is missing that means you are supposed to do it as an exercise. In this case the exercise is not exactly trivial, but it only requires some elementary manipulation of the definitions.

**Corollary 8.** Let $T, S : X \to Y$ and $R : Y \to Z$ be linear maps. Then:

1. $\|cT\| = |c|\|T\|$ if $c \in \mathbb{R}$;
2. $\|T\| = 0$ if and only if $T = 0$;
3. $\|T + S\| \leq \|T\| + \|S\|$;
4. $\|RT\| \leq \|R\|\|T\|$.

**Theorem 9.** Let $T : X \to Y$ be linear, with $\dim X < \infty$. Then:

1. $T$ is bounded;
2. If $T$ is a linear isomorphism then it is a homeomorphism.

*Proof.* We first prove both (1) and (2) for the case $X = \mathbb{R}^n$. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$. Then for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we have

$$\|Tx\| = \left\| T \sum_{i=1}^{n} x_i e_i \right\| = \left\| \sum_{i=1}^{n} x_i T e_i \right\| \leq \sum_{i=1}^{n} |x_i|\|Te_i\| \leq \sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} \|Te_i\|^2} \quad \text{Cauchy-Schwartz Inequality}$$

$$= c\|x\|,$$

where $c = \sqrt{\sum_{i=1}^{n} \|Te_i\|^2}$. Thus $T$ is bounded.

Still assuming $T : \mathbb{R}^n \to Y$ is linear, but now further assuming also that $T$ is an isomorphism, put

$$S = \{x \in \mathbb{R}^n : \|x\| = 1\}.$$

By the Heine-Borel Theorem, $S$ is compact since it is closed and bounded. Since $T$ is 1-1 we have $\|Tx\| > 0$ for all $x \in S$. By what we have already proved above, $T$ is bounded, hence

\(^2\)and in fact Lipschitz, hence uniformly continuous
continuous. Also, the norm function $\| \cdot \|$ on $Y$ is continuous. Thus, the continuous function $x \mapsto \|Tx\|$ has a positive minimum on $S$, so $T^{-1}$ is bounded.

Now for the general case $T : X \to Y$, it suffices to prove (1), because then (2) will follow by applying (1) to $T^{-1}$. Let $n = \dim X$, and choose a linear isomorphism $U : X \to \mathbb{R}^n$. Define $S : \mathbb{R}^n \to Y$ by $S = TU^{-1}$. The situation is summarized in the following diagram:

$$\begin{align*}
\begin{array}{ccc}
X & \xrightarrow{T} & Y \\
\uparrow & & \uparrow \\
\mathbb{R}^n & \xrightarrow{U} & S \\
\downarrow & & \downarrow \\
& & T = SU
\end{array}
\end{align*}$$

It follows from the above arguments that both $S$ and $U$ are bounded, and therefore so is $T$.

**Definition 10.** Two norms $\| \cdot \|$ and $\| \cdot \|$' on a vector space $X$ are equivalent if there exist $a, b > 0$ such that

$$a \|x\| \leq \|x\|' \leq b \|x\| \quad \text{for all } x \in X.$$

Thus two norms $\| \cdot \|$ and $\| \cdot \|$' on $X$ are equivalent if and only if the identity map on $X$, when regarded as a linear map between the normed spaces $(X, \| \cdot \|)$ and $(X, \| \cdot \|')$, is bounded and has bounded inverse.

**Example 11.** The $\infty$-norm and the Euclidean norm on $\mathbb{R}^n$ are equivalent, because

$$\max\{|x_1|, \ldots, |x_n|\} \leq \sqrt{n} \sum_{i=1}^{n} x_i^2 \leq \sqrt{n} \max\{|x_1|, \ldots, |x_n|\}.$$

Similarly, the 1-norm and the $\infty$-norm are equivalent because

$$\max\{|x_1|\} \leq \sum_{i=1}^{n} |x_i| \leq n \max\{|x_i|\}.$$

But in fact the above equivalences are no accident:

**Corollary 12.** On a finite-dimensional vector space, any two norms are equivalent.

**Corollary 13.** If $X$ is a finite-dimensional normed space, then:

1. $X$ is complete;
2. every bounded sequence in $X$ has a convergent subsequence;
3. a subset of $X$ is compact if and only if it is closed and bounded.

**Notation and Terminology.** We write $L(X, Y)$ for the set of all linear maps from $X$ to $Y$, and $L(X)$ for $L(X, X)$.

**Corollary 14.** If $X$ and $Y$ are finite-dimensional then $L(X, Y)$ is a Banach space.

**Proof.** We have not quite proven that the operator norm really is a norm; we should check that if $\|T\| = 0$ then $T = 0$, but this is obvious. Now the corollary follows from the previous results.
Every $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ is represented by an $m \times n$ matrix $A$ with respect to the standard bases, and in fact we usually identify $T$ and $A$. This won’t cause any confusion.

**Lemma 15.** Let $T, T_1, T_2, \cdots \in L(\mathbb{R}^n, \mathbb{R}^m)$, and identify the linear maps with matrices:

$$T_n = (a^n_{ij}) \quad \text{and} \quad T = (a_{ij}).$$

Then $T_n \to T$ in $L(\mathbb{R}^n, \mathbb{R}^m)$ if and only if

$$\lim_{n \to \infty} a^n_{ij} = a_{ij} \quad \text{for all} \ i, j.$$

**Proof.** We can use the norm

$$\|T\| = \max \{|a_{ij}| : i = 1, \ldots, m, j = 1, \ldots, n\},$$

and then the result is obvious. □

**Proposition 16.** Let $T \in L(X)$. If $\|T\| < 1$, then $I - T$ is invertible and

$$(I - T)^{-1} = \sum_{n=0}^{\infty} T^n.$$

**Proof.** The series $\sum_{n=0}^{\infty} T^n$ converges absolutely (because $\|T^n\| \leq \|T\|^n$), hence converges since $L(X)$ is complete. Since left multiplication by $T$ is a bounded linear map on $L(X)$, we have

$$T \sum_{n=0}^{\infty} T^n = \sum_{n=0}^{\infty} T^{n+1} = \sum_{n=1}^{\infty} T^n = \sum_{n=0}^{\infty} T^n - I,$$

so

$$(I - T) \sum_{n=0}^{\infty} T^n = I,$$

and similarly (or consequently, since dim $X$ is finite)

$$\sum_{n=0}^{\infty} T^n (I - T) = I.$$

**Corollary 17.** Let $T \in L(X)$ with $\|T\| < 1$. Then:

(1) $$\|(I + T)^{-1} - I\| \leq \frac{\|T\|}{1 - \|T\|};$$

(2) $$\|(I + T)^{-1} - I + T\| \leq \frac{\|T\|^2}{1 - \|T\|}.$$

**Proof.** (1) We have

$$(I + T)^{-1} = \sum_{n=0}^{\infty} (-1)^n T^n,$$

so

$$(I + T)^{-1} - I = \sum_{n=1}^{\infty} (-1)^n T^n,$$

hence

$$\|(I + T)^{-1} - I\| \leq \sum_{n=1}^{\infty} \|(-1)^n T^n\| = \sum_{n=1}^{\infty} \|T^n\| \leq \sum_{n=1}^{\infty} \|T\|^n = \frac{\|T\|}{1 - \|T\|}.$$
Then verification of (2) is similar, using

\[(I + T)^{-1} - I + T = \sum_{n=2}^{\infty} (-1)^n T^n.\] □

**Corollary 18.**

1. The set of invertible linear operators on \(X\) is open in \(L(X)\), and
2. the map \(T \mapsto T^{-1}\) is a homeomorphism on this set.

**Proof.** (1) Let \(T \in L(X)\) be invertible, and let \(A \in L(X)\) with

\[\|A\| < \frac{1}{\|T^{-1}\|}.\]

Note that

\[T + A = T(I + T^{-1}A).\]

Since

\[\|T^{-1}A\| \leq \|T^{-1}\|\|A\| < 1,\]

\(I + T^{-1}A\) is invertible, hence so is \(T + A\).

(2) With \(T\) and \(A\) as above, we have

\[(T + A)^{-1} - T^{-1} = (I + T^{-1}A)^{-1}T^{-1} - T^{-1} = ((I + T^{-1}A)^{-1} - I)T^{-1},\]

so

\[\|(T + A)^{-1} - T^{-1}\| \leq \|(I + T^{-1}A)^{-1} - I\|\|T^{-1}\|

\[\leq \frac{\|T^{-1}A\|\|T^{-1}\|}{1 - \|T^{-1}A\|}\] (by Corollary 17)

\[\xrightarrow{A \to 0} 0,\]

since \(T^{-1}A \to 0\) as \(A \to 0\). Therefore the inversion map is continuous at \(T\), hence is a homeomorphism since it is its own inverse. □

**Corollary 19.** The set of invertible \(n \times n\) matrices is open in \(M_n\).

Although this last corollary could be derived more quickly using properties of determinants, it is more instructive from the point of view of linear maps to prove it using operator norms.