Analysis II: Background

Jesse Ratzkin

September 16, 2009

In these notes we establish some background material which will prove useful throughout the course. Hopefully, much of this will be review for the reader.

Before we begin, we list some useful references. The classical textbook covering basic analysis is *Principles of Mathematical Analysis* by W. Rudin. This book is very dense, but very much worth the effort it will take to work through it. There are several other good textbooks, including *The Way of Analysis* by R. Strichartz and *Elementary Classical Analysis* by J. Marsden.

One can also find decent lecture notes online (for free). J. Taylor has a set of notes available at
`http://www.math.utah.edu/~taylor/foundations.html`
N. Korevaar has scanned lecture notes at
`http://www.math.utah.edu/~korevaar/5210spring09/5210lectures.html`
You can browse many online lecture notes at the MIT OpenCourseWare website:
`http://ocw.mit.edu/OcwWeb/Mathematics/index.htm`
This last website is highly recommended; it's usually a good idea to learn from the experts.

1 Topology of $\mathbb{R}^n$

Here we review the topology of the Euclidean space $\mathbb{R}^n$. Recall that Euclidean space comes equipped with a distance: for $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{R}^n$, we denote this distance by

$$|x - y| = \text{dist}(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

The distance function in turn determines a topology, in which a set $U \subset \mathbb{R}^n$ is open if and only if for any $x \in U$ there is an $\epsilon > 0$ such that the metric ball $B_\epsilon(x) = \{y : |x - y| < \epsilon\}$ is contained in U. Equivalently, a set is open if and only if one can write it as a union of metric balls.

With open sets come a host of other useful tools, such as closed sets, compact sets, and connected sets. It is important to remember the Heine-Borel characterization of compact sets in $\mathbb{R}^n$: a set $K \subset \mathbb{R}^n$ is compact if and only if it is both closed and bounded.

The main purpose of this section is to remind the reader of convergence properties. Let $\{x_k : k = 1, 2, 3, \ldots\}$ be a sequence of points in $\mathbb{R}^n$. (Note that here the subscript serves to index the points in $\mathbb{R}^n$, and does not denote components of an individual point. We have two equivalent definitions of convergence: the limit $\bar{x}$ of the sequence $\{x_k\}$ satisfies either one of the following conditions:

- for all $\epsilon > 0$ there is a natural number N such that $|x_k - \bar{x}| < \epsilon$ for all $k > N$.
- for all open sets U containing $\bar{x}$ there is a natural number N such that $x_k \in U$ for all $k > N$.

We prove that these two conditions are equivalent. Suppose the first condition holds and pick an open set U containing $\bar{x}$. There is a metric ball $B_\epsilon(\bar{x})$ centered at $\bar{x}$, for some $\epsilon > 0$, which is contained in U. Then, for $k > N$, $|x_k - \bar{x}| < \epsilon$, which implies $x_k \in B_\epsilon(\bar{x}) \subset U$. Conversely, suppose the second condition holds and choose $\epsilon > 0$. The ball $B_\epsilon(\bar{x})$ is an open set containing $\bar{x}$, so for $k > N$ we have $x_k \in B_\epsilon(\bar{x})$. This in turn implies $|x_k - \bar{x}| < \epsilon$. □

We write the limit in two equivalent ways:

$$lim_{k\to\infty} x_k = \bar{x}, \qquad x_k \to \bar{x}.$$

Notice that, if $x_k \to \bar{x}$, then *all* the components of x_k must approach the corresponding component of $\bar{x}$.

Lemma 1. *A sequence $\{x_k\}$ converges to $\bar{x}$ if and only if $\langle x_k, e_j\rangle \to \langle \bar{x}, e_j\rangle$ for all standard basis elements e_j.*

Proof. Let $x_k \to \bar{x}$. We suppose $\langle x_k, e_j\rangle \not\to \langle \bar{x}, e_j\rangle$ and derive a contradiction. In this case there is some $M > 0$ such that

$$|\langle x_k, e_j\rangle - \langle \bar{x}, e_j\rangle| > M.$$

Then for any choice $\epsilon < M$ we have $|x_k - \bar{x}| \geq M > \epsilon$, so $\{x_k\}$ cannot converge to $\bar{x}$.

Conversely, suppose $\langle x_k, e_j\rangle \to \langle \bar{x}, e_j\rangle$. For any $\epsilon > 0$, there is $N \in \mathbb{N}$ such that if $k \geq N$ then

$$|\langle x_k, e_j\rangle - \langle \bar{x}, e_j\rangle| \leq \sqrt{\frac{\epsilon}{n}}$$

for all $j = 1, 2, \ldots, n$. (We can choose the same N for all j by taking a maximum as j varies over the finitely many values it can assume.) Then, for $k \geq N$,

$$|x_k - \bar{x}| = \sqrt{\sum_{j=1}^{n} |\langle x_k, e_j\rangle| - \langle \bar{x}, e_j\rangle|^2} \leq \epsilon,$$

and so $x_k \to \bar{x}$. □

This leads to the Bolzano-Weierstrass theorem:

Theorem 2. *Any bounded sequence in* $\mathbb{R}^n$ *has a convergent subsequence.*

Recall that a sequence $\{x_k\}$ has a convergent subsequence if one can an increasing sequence k_j of indices, $j = 1, 2, \dots$, and a point $\bar{x} \in \mathbb{R}^n$ such that $x_{k_j} \to \bar{x}$.

Proof. The sequence $\{x_k\}$ is bounded, so there is a finite, positive number M such that $|\langle x_k, e_j\rangle| \leq M$ for all $k = 1, 2, \dots$ and $j = 1, \dots, n$. Now apply the one-dimensional version of the Bolzano-Weierstrass theorem to each component $\langle x_k, e_j\rangle$. □

For the reader's convenience, we recall the proof of the Bolzano-Weierstrass theorem on the real line.

Proof. We will first need to prove the following nested intervals lemma: if $\{I_k = (a_k, b_k)\}$ be a sequence of nested intervals, with $I_{k+1} \subset I_k$, then the intersection $\cap_k I_k$ is nonempty. The sequences $\{a_k\}$ and $\{b_k\}$ are both bounded monotone sequences ($\{a_k\}$ is an increasing sequence and $\{b_k\}$ is decreasing) so they both converge. Call the limits $a = \lim_{k\to\infty} a_k$ and $b = \lim_{k\to\infty} b_k$. Also, because for all $k, m \in \mathbb{N}$, $a_k < b_m$, we must have $a \leq b$ (equality is possible). Then, for any $c \in [a, b]$ and $k \in \mathbb{N}$, we have

$$a_k \leq a \leq c \leq b \leq b_k \Rightarrow c \in I_k,$$

and so $c \in \cap_k I_k$.

Now let $\{a_k\}$ be some bounded sequence. Then there is some $M > 0$ such that $a_k \in [-M, M] = I_1$ for all k. We will construct a nested sequence of intervals $\{I_k\}$ by successively choosing I_{k+1} to be either the left half or the right half of I_k. First observe that the length $|I_k| = 2^{2-k}M \to 0$ as $k \to \infty$. Now suppose I_k contains infinitely many elements in the sequence $\{a_k\}$ for some k. Either the left half or the right half of I_k will also contain infinitely many of the a_k's, so choose I_{k+1} to be one half that does. Next, we choose the subsequence $\{a_{k_m}\}$ as follows. Let $k_1 = 1$, and then choose k_m to be the least $k > k_{m-1}$ such that $a_{k_m} \in I_k$. This is always possible because each interval I_k contains infinitely many elements of the original sequence.

Finally, we let $a \in \cap_k I_k$. Indeed, if there is another $\tilde{a} \in \cap I_k$, then

$$|\tilde{a} - a| \leq |I_k| = \frac{M}{2^{k-2}} \to 0,$$

so $\tilde{a} = a$ and this choice of a is unique. Further, if $\epsilon > 0$ we can choose k large enough so that $2^{k-2} > M/\epsilon$, and then (using $k_m \geq k$) $|a_{k_m} - a| < \epsilon$. Thus $\lim_{m\to\infty} a_{k_m} = a$, completing the proof. □

As an exercise, it is worth the reader's time to determine which of these sequences converges.

1. $x_k = (\frac{1}{k}, \frac{1}{k^2})$
2. $x_k = (\frac{1}{k}, k)$
3. $x_k = (\frac{k}{k+1}, \log(k))$
4. $x_k = (\frac{\log(k)}{k}, \frac{k}{k+1})$

For the reader's convenience, we collect some other useful facts about the topology of $\mathbb{R}^n$. All of these should be familiar.

- A set $U \subset \mathbb{R}^n$ is open if and only if it is a union of open metric balls, if and only if for every $x \in U$ there is an $\epsilon > 0$ such that the open metric ball $B_\epsilon(x) \subset U$.
- A set $K \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded. (This is the Heine-Borel theorem.)
- A set $E \subset \mathbb{R}^n$ is connected if and only if it is path-connected. In particular, the only connected subsets of $\mathbb{R}$ are intervals.

We end this section with a brief discussion of convexity.

Definition 1. *If $p, q \in \mathbb{R}^n$ we denote the line segment joining p to q by $[p, q]$. A set $U \subset \mathbb{R}^n$ is convex if for every pair $p, q \in U$ we also have $[p, q] \subset U$. In general, the convex hull of a set U, written* $\text{conv}(U)$, *is the smallest convex set containing U.*

Observe that the intersection of two convex sets is still convex. Thus, for any $U \subset \mathbb{R}^n$, we can write

$$\text{conv}(U) = \cap\{K : U \subset K, K \text{ convex}\};$$

that is, the convex hull of a set is the intersection of all convex sets containing it. Next, observe that convex sets are connected. In fact, if U is convex then one can connect and $p \in U$ to any other $q \in U$ by a straight line segment, and so U is path-connected, which in turn implies U is connected. We concluded that the only convex subsets of the real line $\mathbb{R}$ are line segments.

2 Functions of several variables

Let $U \subset \mathbb{R}^n$ be open and connected; and all such a set a domain. A function$f : U \to \mathbb{R}^m$ is an assinment of a point $y = f(x) \in \mathbb{R}^m$ to each point $x \in U$. The space $\mathbb{R}^m$ is the target of the function f, and the image of f is the collection of all the points $\{y = f(x) : x \in U\}$.

We present two equivalent notions of continuity.

Definition 2. *A function $f : U \to \mathbb{R}^m$ is continuous on the set U if either*

- $\lim_{\tilde{x}\to x} f(\tilde{x}) = f(x)$ *for all* $x \in U$
- *If $O \subset \mathbb{R}^m$ is an open set containing then so is its preimage $f^{-1}(O)$.*

We must show these two definitions are equivalent.

Proof. Suppose $\lim_{\tilde{x}\to x} f(\tilde{x}) = f(x)$ for all $x \in U$ and choose an open set $O \subset \mathbb{R}^m$. For any $y \in f(U)\cap O$, there is an $\epsilon > 0$ such that $B_\epsilon(y) \subset O$. Using the definition of a limit, we see that we can choose $\delta > 0$ such that $|\tilde{x}-x| < \delta$ implies $|f(\tilde{x})-y| < \epsilon$, and so $B_\delta(x) \subset f^{-1}(O)$. There was nothing particular about our choice of x, other than $x \in U$, so we conclude that $f^{-1}(O)$ is contained in the union of metric balls and is hence open.

Conversely, suppose the preimage of every open set is open, choose $x \in U$, and let $y = f(x)$. This time we choose $\epsilon > 0$ and consider the open set $O = B_\epsilon(y)$. Because $x \in f^{-1}(B_\epsilon(y))$ and $f^{-1}(B_\epsilon(y))$ is open, we can choose $\delta > 0$ such that $B_\delta(x) \subset f^{-1}(B_\epsilon(y))$ Unraveling what this means, if $|\tilde{x} - x| < \delta$ then $f(\tilde{x}) \in B_\epsilon(y)$, which in turn implies $|f(\tilde{x}) - f(x)| < \epsilon$. Thus $\lim_{\tilde{x}\to x} f(\tilde{x}) = x$, as claimed. □

Observe that, by taking complements, we could have just as easily proved that $f : U \to \mathbb{R}^m$ is continuous if and only if the preimage of every closed set is closed.

It is worthwhile to remember which direction the inclusions go. In particular, the image of a closed set under a continuous function may not be closed. For instance, consider $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x$, and the closed set $C = \{(x, \frac{1}{x}) : x > 0\}$. The image of C under the continuous function f is the open half-line $(0,\infty)$, and it is not certainly not closed.

The following theorem proves one of the more useful properties of continuous functions.

Theorem 3. *The image of a compact set under a continuous function is compact.*

Proof. Let $f : U \to \mathbb{R}^m$ be continuous and choose a compact subset $K \subset U$. Choose any open covering $\{O_\alpha\}$ of the set $f(K)$. By continuity, each preimage $V_\alpha = f^{-1}(O_\alpha)$ is open, so the collection $\{V_\alpha\}$ is an open cover of K. By compactness, we can choose a finite subcover $\{V_1, \ldots, V_k\}$, whose union covers K. We claim $\{O_1, \ldots, O_k\}$ is the finite subcover fo $f(K)$ we're looking for, where $V_j = f^{-1}(O_j)$. Indeed, if $y \in f(K)$ then $y = f(x)$ for some $x \in K$. This implies $x \in V_j$ for some $j \in \{1, 2, \ldots, k\}$, and so $y \in O_j$. Thus $\{O_1, \ldots, O_k\}$ is indeed an open cover of $f(K)$. □

In particular, consider a continuous function $f : K \to \mathbb{R}$, where $K \subset \mathbb{R}^n$ is compact. The image is a compact subset of $\mathbb{R}$, so it must contain its supremum and infimum. We have just proved the following corollary:

Corollary 4. *Any continuous, real-valued function on a compact set attains both its maximum and minimum.*

Again, we have to be careful of the direction in which we apply the theorem. The example $f(x) = \arctan(x)$, mapping $\mathbb{R}$ to itself shows that the preimage of a compact set may not be compact.

The theorem and corollary above begin to explain why compact sets are important. Another supporting argument for the importance of compact sets is the following result regarding continuity and uniform continuity.

Theorem 5. *Let $f : K \to \mathbb{R}^m$ be continuous, with $K \subset \mathbb{R}^n$ compact. Then f is in fact uniformly continuous.*

Recall that the function f is uniformly continuous if for every ϵ there is $\delta > 0$ such that $|x - \tilde{x}| < \delta$ implies $|f(x) - f(\tilde{x})| < \epsilon$. The important characteristic of this definition (and what distinguished uniform continuity from regular continuity) is the fact that δ **does not depend of** x **or** $\tilde{x}$. The proof below is more general, and applies to continuous functions between metric spaces.

Proof. Pick $\epsilon > 0$. For any $x \in K$, there is $\delta = \delta(x) > 0$ such that $|x - \tilde{x}| < \delta(x)$ implies $|f(x) - f(\tilde{x})| < \frac{\epsilon}{2}$. Now the collection of metric balls $\{B_{\frac{\delta(x)}{2}}(x) : x \in K\}$ is an open cover of K, so (by compactness), we can choose an open subcover

$$\{B_{\frac{\delta_1}{2}}(x_1), \ldots, B_{\frac{\delta_k}{2}}(x_k)\},$$

and set

$$\delta = \min\{\delta_1, \ldots, \delta_k\}.$$

Now suppose we have two points $x, \tilde{x} \in K$, with $|x - \tilde{x}| < \delta$. There is some j such that $x \in B_{\frac{\delta_j}{2}}(x_j)$, which is the same as saying $|x - x_j| < \frac{\delta_j}{2}$. Further,

$$|\tilde{x} - x_j| \leq |\tilde{x} - x| + |x - x_j| < \delta + \frac{\delta_j}{2} \leq \delta.$$

Thus

$$|f(x) - f(\tilde{x})| \leq |f(x) - f(x_j)| + |f(x_j) - f(\tilde{x})| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

We leave the following facts as exercises to the reader.

- If $U \subset \mathbb{R}^n$ is not closed, then there is a continuous function $f : U \to \mathbb{R}$ which is not bounded. (Hint: consider a point $\bar{x} \notin U$ which is an accumulation point of U.)
- If $U \subset \mathbb{R}^n$ is not compact, then there is a continuous function $f : U \to \mathbb{R}$ which is not bounded. (Hint: use the Heine-Borel theorem.)
- If $U \subset \mathbb{R}^n$ is not compact, then there is a function $f : U \to \mathbb{R}$ which is not uniformly continuous.

3 Linear Algebra

This section covers some useful tools and techniques from linear algebra, assuming the reader is familiar with basic concepts, such as

$$\dim(T(V)) + \dim(\ker(T)) = \dim(V)$$

where $T : V \to W$ is linear (or, equivalently, rank + nullity $= n$). If you don't remember this formula you might want to review your favorite linear algebra textbook, such as Anton's book and Bretscher's book.

To make this discussion a little more concrete, we will consider a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$. So long as the domain V and target W are both finite dimensional, this is really the same thing as an abstract linear transformation $T : V \to W$ between vector spaces. (Note: if you're not nodding your head in agreement right now, review change of bases for linear transformations.) We'll also write the standard basis of $\mathbb{R}^n$ as $\{e_1, \ldots, e_n\}$, where e_j has a 1 in the jth slot and zeroes everywhere else.

We'd like to understand the geometry associated to the linear transformation T, so we will need to first define a way to measure distances. Denote the Euclidean length of a vector v as

$$|v| = |(v_1, \ldots, v_n)| = \sqrt{v_1^2 + \cdots v_n^2},$$

and recall that the unit sphere $\mathbf{S}^{n-1}$ is given by

$$\mathbf{S}^{n-1} = \{v \in \mathbb{R}^n : |v| = 1\}.$$

Definition 3. *If $T : \mathbb{R}^n \to \mathbb{R}^m$ linear, write its operator norm as*

$$\|T\| = \sup_{u \in \mathbf{S}^{n-1}} \{|T(u)|\}.$$

We collect some properties of the operator norm.

- In the definition of the operator norm, we can replace the supremum with a maximum.

  The unit sphere $\mathbf{S}^{n-1}$ is a closed, bounded set in $\mathbb{R}^n$, so by the Heine-Borel property it is compact. The restriction of T to $\mathbf{S}^{n-1}$ is a continuous function on a compact space, so it much achieve its maximum (and minimum).

- $\|T\| = 0$ only if T is the zero transformation.

  If $T \neq 0$ then there a nonzero $v \in \mathbb{R}^n$ such that $T(v) \neq 0$. Then

$$\|T\| \geq |T(\frac{v}{|v|})| > 0.$$

- $\|T\|$ is finite.

  Write $u \in \mathbf{S}^{n-1}$ as $u = \sum_{j=1}^{n} u_j e_j$. Because $u \in \mathbf{S}^{n-1}$ each coefficient u_j is between -1 and 1, so

  $$|T(u)| = |T(u_1e_1+\cdots+u_ne_n)| \le |u_1||T(e_1)|+\cdots+|u_n||T(e_n)| \le |T(e_1)|+\cdots+|T(e_n)|.$$

  This provides a uniform bound for $|T(u)|$ as u ranges over $\mathbf{S}^{n-1}$. Notice that the bound depends on T. Indeed, it is easy to construct a sequence of linear transformations $T_k : \mathbb{R}^n \to \mathbb{R}^m$ with $\|T_k\| \to \infty$. (Try it.)

- If $\lambda \in \mathbb{R}$ and $u \in \mathbf{S}^{n-1}$ then $|\lambda T(u)| = |\lambda||T(u)|$, and so $\|\lambda T\| = |\lambda|\|T\|$.

- If $S : \mathbb{R}^n \to \mathbb{R}^m$ is another linear transformation and $u \in \mathbf{S}^{n-1}$ then

  $$|(S+T)(u)| = |S(u) + T(u)| \le |S(u)| + |T(u)| \le \|S\| + \|T\|,$$

  and so $\|S+T\| \le \|S\| + \|T\|$.

- Now let $S : \mathbb{R}^m \to \mathbb{R}^k$ be linear, and consider the compostion $S \circ T$. If $u \in \mathbf{S}^{n-1}$ then

  $$|S(T(u))| \le \|S\||T(u)| \le \|S\|\|T\|,$$

  and so $\|S \circ T\| \le \|S\|\|T\|$.

- This last inequality can be strict. Let $S, T : \mathbb{R}^3 \to \mathbb{R}^3$ be orthogonal projections onto the xy plane and the z axis, respectively. Then $\|S\| = \|T\| = 1$, but the composition is the zero transformation.

Write the matrix representation of T as

$$[T] = [a_{ij}].$$

We can estimate the operator norm of T in terms of the coefficients a_{ij}. If $v = v_1e_1 + \cdots + v_ne_n$ then

$$|T(v)| \le \sum_{j=1}^{n} |v_j||T(e_j)| \le \sum_{j=1}^{n}\sum_{i=1}^{m} |v_j||a_{ij}||e_j| = \sum_{j=1}^{n}\sum_{i=1}^{m} |v_j||a_{ij}|.$$

Since there are mn terms in the last sum, we conclude

$$\|T\| \le mn \max_{i,j}\{|a_{ij}|\}.$$

On the other hand, we can also find a lower bound. Let $a_{i_0j_0} = \max_{i,j} |a_{ij}|$. Then $a_{i_0j_0}e_{i_0}$ is one component of $T(e_{j_0})$, so $|a_{i_0,j_0}| \le |T(e_{j_0})|$. We conclude

$$\max_{i,j} |a_{ij}| \le \|T\| \le mn \max_{i,j} |a_{ij}|.$$

Geometrically, $\|T\|$ measures the maximum stretch factor of T on the unit sphere. If $v \neq 0$ in $\mathbb{R}^n$, then

$$|T(v)| = |v||T(\frac{v}{|v|})| \leq \|T\||v|.$$

Thus we see that T maps the unit ball into a ball of radius $\|T\|$. Put another way, if $\mathbf{B}$ is the unit ball in $\mathbb{R}^n$, the circumradius of $T(\mathbf{B})$ is $\|T\|$.

We can refine this picture by introducing the following semi-norm. (A semi-norm has most of the properties of a norm, except that it might be zero on some nonzero elements.)

Definition 4. *If $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, we define the conorm*

$$\mathbf{m}(T) = \inf_{u \in \mathbf{S}^{n-1}} \{|T(u)|\}.$$

The conorm measures the minimum stretch of T, when applied to the unit sphere. If $\mathbf{m}(T) \neq 0$, the ratio,

$$\frac{\|T\|}{\mathbf{m}(T)} \geq 1,$$

is called the condition number of T, and it measure how much T distorts the unit sphere. (Exercise: what can you say if the condition number is 1?)

Proposition 6. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear. The conorm $\mathbf{m}(T) > 0$ if and only if T is one-to-one. If $S : \mathbb{R}^n \to \mathbb{R}^m$ is also linear then $\mathbf{m}(S + T) \geq \mathbf{m}(s) - \|T\|$. Finally, if $S : \mathbb{R}^m \to \mathbb{R}^k$ then $\mathbf{m}(S \circ T) \geq \mathbf{m}(S)\mathbf{m}(T)$.*

Proof. We first show $\mathbf{m}(T) > 0$ if and only if T is one-to-one. If $\mathbf{m}(T) = 0$ then $T(u) = 0$ for some $u \in \mathbf{S}^{n-1}$. But because T is linear, we also have $T(0) = 0$, so T is not one-to-one. Now suppose T is not one-to-one, say $T(v_1) = T(v_2)$ for some $v_1 \neq v_2$. Then let

$$u = \frac{v_1 - v_2}{|v_1 - v_2|} \in \mathbf{S}^{n-1}.$$

By our choice, $T(u) = 0$, and so $\mathbf{m}(T) = 0$.

Now let $S : \mathbb{R}^n \to \mathbb{R}^m$, and choose $u \in \mathbf{S}^{n-1}$. Then

$$|S(u) + T(u)| \geq |S(u)| - |T(u)| \geq \mathbf{m}(S) - \|T\|.$$

Finally, let $S : \mathbb{R}^m \to \mathbb{R}^k$ and choose $u \in \mathbf{S}^{n-1}$. Then

$$|S(T(u))| \geq \mathbf{m}(S)|T(u)| \geq \mathbf{m}(S)\mathbf{m}(T).$$

□

We consider the case $m = n$.

Proposition 7. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$. Then $\mathbf{m}(T) > 0$ if and only if T is invertible. Furthermore, in this case $\mathbf{m}(T) = \|T^{-1}\|^{-1}$.*

Proof. The first statement follows from the previous proposition and general facts about linear transformations. Now let T be invertible and denote its inverse by S. Then

$$\mathrm{Id} = S \circ T \Rightarrow 1 = \mathbf{m}(S \circ T) \geq \mathbf{m}(S)\mathbf{m}(T) \Rightarrow \mathbf{m}(T) \geq \frac{1}{\mathbf{m}(T^{-1})} \geq \frac{1}{\|T^{-1}\|}.$$

Next choose $w \in \mathbf{S}^{n-1}$ such that $|T^{-1}(w)| = \|T^{-1}\|$ and let $v = T^{-1}(w)$. Then

$$1 = |w| = |T(v)| \geq \mathbf{m}(T)|v| = \mathbf{m}(T)\|T^{-1}\|.$$

Rearranging this inequality gives $\mathbf{m}(T) \leq \|T^{-1}\|^{-1}$, completing the proof. □

Finally, we restrict ourselves further to the set $Gl(n, \mathbb{R})$ of invertible linear transformations on $\mathbb{R}^n$, inside the space $L(n, \mathbb{R}) = \mathbb{R}^{n^2}$ of linear transformations from $\mathbb{R}^n$ to itself. The operator norm $\| \cdot \|$ gives us a natural way to measure distances on $\mathbb{R}^{n^2}$, and so we can make sense of things like open sets and continuous maps.

Proposition 8. *$Gl(n, \mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$, and the inversion map $T \mapsto T^{-1}$ is continuous on it.*

Proof. Let $T \in \mathbb{R}^{n^2}$ be invertible, with matrix coefficients a_{ij}. The determinant $\det(T)$ is a degree n polynomial in the coefficients a_{ij}, so it is a continuous function. The linear transformation T is invertible if and only if $\det(T) \neq 0$, so $Gl(n, \mathbb{R})$ is the preimage of the open set $\mathbb{R}\backslash\{0\}$ under the continuous function det, and is hence open. Now let $S, T \in Gl(n, \mathbb{R})$ and observe

$$S^{-1} - T^{-1} = S^{-1} \circ (T - S) \circ T^{-1}.$$

Then

$$\|S^{-1} - T^{-1}\| \leq \|S^{-1}\|\|T - S\|\|T^{-1}\| \to 0$$

as $\|T - S\| \to 0$. □

There is a trick to remember the factorization we used in the second part of this proof. Start with $S^{-1} - T^{-1}$ and factor out S^{-1} on the left to get $S^{-1} \circ (\mathrm{Id} - S \circ T^{-1})$. Now factor out T^{-1} on the right to get the formula above.