Lecture Notes on Metric Spaces

Math 117: Summer 2007
John Douglas Moore

Our goal of these notes is to explain a few facts regarding metric spaces not included in the first few chapters of the text [1], in the hopes of providing an easier transition to more advanced texts such as [2].

1 The dot product

If $\mathbf{x} = (x_1 \ldots, x_n)$ and $\mathbf{y} = (y_1 \ldots, y_n)$ are elements of $\mathbb{R}^n$, we define their *dot product* by

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n.$$

The dot product satisfies several key axioms:

1. it is symmetric: $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$;
2. it is bilinear: $(a\mathbf{x} + \mathbf{x}') \cdot \mathbf{y} = a(\mathbf{x} \cdot \mathbf{y}) + \mathbf{x}' \cdot \mathbf{y}$;
3. and it is positive-definite: $\mathbf{x} \cdot \mathbf{x} \geq 0$ and $\mathbf{x} \cdot \mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

We define the *length* of $\mathbf{x}$ by

$$|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

Note that the length of $\mathbf{x}$ is always ≥ 0.

Cauchy-Schwarz Theorem. *The dot product satisfies*

$$-1 \leq \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|} \leq 1.$$

Sketch of proof: Expand the inequality

$$0 \leq (\mathbf{x}(\mathbf{y} \cdot \mathbf{y}) - \mathbf{y}(\mathbf{x} \cdot \mathbf{y})) \cdot (\mathbf{x}(\mathbf{y} \cdot \mathbf{y}) - \mathbf{y}(\mathbf{x} \cdot \mathbf{y}))$$

and simplify. (Exercise: Work this out!)

The importance of the Cauchy-Schwarz Theorem is that it allows us to define angles between vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. Given an number $t \in [-1, 1]$, there is a unique angle θ such that

$$\theta \in [0, \pi] \qquad \text{and} \qquad \cos\theta = t.$$

Thus we can define the angle between two nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ by requiring that

$$\theta \in [0, \pi] \qquad \text{and} \qquad \cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|}.$$

Thus we can say that two vectors vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ are *perpendicular* or *orthogonal* if $\mathbf{x} \cdot \mathbf{y} = 0$. This provides much intuition for dealing with vectors in $\mathbb{R}^n$.

Corollary of Cauchy-Schwarz Theorem. *If* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, *then*

$$|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|.$$

Proof: It suffices to check that

$$|\mathbf{u} + \mathbf{v}|^2 \leq (|\mathbf{u}| + |\mathbf{v}|)^2$$

or

$$|\mathbf{u}|^2 + 2\mathbf{u} \cdot \mathbf{v} + |\mathbf{v}|^2 \leq |\mathbf{u}|^2 + 2|\mathbf{u}||\mathbf{v}| + |\mathbf{v}|^2.$$

But this follows immediately from the Cauchy-Schwarz inequality.

2 Metric spaces

Definition 2.1. A *metric space* is a set X together with a function $d : X \times X \to \mathbb{R}$ such that

1. $d(x, y) \geq 0$ and $d(x, y) = 0 \Leftrightarrow x = y$,
2. $d(x, y) = d(y, x)$, and
3. $d(x, z) \leq d(x, y) + d(y, z)$.

The last of these conditions is known as the *triangle inequality*.

Example 1. The most basic example is $\mathbb{R}$ with

$$d : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \quad \text{defined by} \quad d(x, y) = |x - y|.$$

Example 2. The example needed for multivariable calculus is $\mathbb{R}^n$ with

$$d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \quad \text{defined by} \quad d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|,$$

the length of $\mathbf{x} - \mathbf{y}$ being defined as in the preceding section. In verifying that this really is a metric space, the only difficulty is checking the triangle inequality. But this follows from the Corollary in the preceding section when $\mathbf{u} = \mathbf{x} - \mathbf{y}$ and $\mathbf{v} = \mathbf{y} - \mathbf{z}$. Of course, this example includes the previous one as a special case.

Example 3. Let S be a subset of $\mathbb{R}^n$. If $\mathbf{x}$ and $\mathbf{y}$ are elements of S, we let $d(\mathbf{x}, \mathbf{y})$ be the distance from $\mathbf{x}$ to $\mathbf{y}$ in $\mathbb{R}^n$. This provides numerous examples of metric spaces. For example, the unit sphere

$$S^{n-1} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : (x_1^2 + \cdots + x_n^2 = 1\}$$

inherits a metric from $\mathbb{R}^n$.

Definition 2.2. If (X_1, d_1) and (X_2, d_2) are metric spaces, a function $f : X_1 \to X_2$ is *continuous* if for every $c \in X_1$ and every $\epsilon > 0$ there exists a $\delta_c > 0$ such that

$$x \in X_1 \quad \text{and} \quad d_1(x, c) < \delta_c \quad \Rightarrow \quad d_2(f(x), f(c)) < \epsilon. \tag{1}$$

A function $f : X_1 \to X_2$ is *uniformly continuous* if for every $\epsilon > 0$ there exists a δ such that

$$x, y \in X_1 \quad \text{and} \quad d_1(x, y) < \delta \quad \Rightarrow \quad d_2(f(x), f(y)) < \epsilon.$$

Definition 2.3. If (X, d) is a metric space and $\epsilon > 0$ is given the ϵ-neighborhood $N(x; \epsilon)$ of a point $x \in X$ is

$$N(x; \epsilon) = \{y \in X : d(x, y) < \epsilon\}.$$

We can then rewrite (1) as

$$x \in N(x; \delta_c) \quad \Rightarrow \quad f(x) \in N(f(c); \epsilon). \tag{2}$$

Definition 2.4. A subset U of X is said to be *open* if

$$x \in U \quad \Rightarrow \quad N(x; \epsilon) \subset U, \quad \text{for some } \epsilon > 0.$$

A subset C of X is *closed* if its complement $X - C$ is open.

Proposition 2.1. *The collection of open sets satisfies the following:*

1. *X is open and the empty set $\emptyset$ is open.*
2. *An arbitrary union of open sets is open.*
3. *A finite intersection of open sets is open.*

Proof: Exercise.

Proposition 2.2. *Let (X, d) be a metric space and let S be a subset of X, which is a metric space in its own right. Then a subset U of S is open if and only if $U = V \cap S$, where V is open in X.*

Proof: Exercise.

Proposition 2.3. *If (X_1, d_1) and (X_2, d_2) are metric spaces and a function $f : X_1 \to X_2$ is continuous, then whenever U is an open subset of X_2, $f^{-1}(U)$ is an open subset of X_1.*

Proof: Suppose that $c \in f^{-1}(U)$. Then $f(c) \in U$ and since U is open, $N(f(c); \epsilon) \subset U$ for some $\epsilon > 0$. Since f is continuous, there exists $\delta_c > 0$ such that (2) holds; thus $f(N(c : \delta_c)) \subset N(f(c); \epsilon)$. Let

$$V = \bigcup \{N(c, \delta_c) : c \in f^{-1}(U)\}.$$

Then $V = f^{-1}(U)$, and since V is the union of open sets, it is open.

Remark. The converse of this Proposition is also true: If $f^{-1}(U)$ is an open subset of X_1 whenever u is an open subset of X_2, then f is continuous. We leave this as an exercise for you to prove.

3 Compact sets

Definition. Suppose that (X, d) is a metric space. A subset K of X is *compact* if every open cover of K has a finite subcover.

Proposition 3.1. *Any closed subset of a compact set is compact.*

Proof: You did this as one of the earlier exercises.

Main Theorem. *If (X_1, d_1) and (X_2, d_2) are metric spaces, $f : X_1 \to X_2$ is a continuous function and K is a compact subset of X_1, then $f(K)$ is compact.*

Proof: Suppose that

$$\mathcal{F} = \{U_\alpha : \alpha \in A\}$$

is an open cover of $f(K)$. For each U_α, there is an open subset $V_\alpha \subset X_1$ such that $V_\alpha \cap K = f^{-1}(U_\alpha)$. We then have

$$K \subset \bigcup \{V_\alpha : \alpha \in A\}.$$

In other words, $\{V_\alpha : \alpha \in A\}$ is an open cover of K. Since K is compact, there exists a finite subcover $\{V_{\alpha_1}, \ldots, V_{\alpha_n}\}$; in other words,

$$K \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}.$$

But then

$$f(K) \subset f(V_{\alpha_1}) \cup \cdots \cup f(V_{\alpha_n}) \subset U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}.$$

So $\{U_{\alpha_1}, \ldots, U_{\alpha_n}\}$ is a finite subcover of $\mathcal{F}$. Since $\mathcal{F}$ was an arbitrary open cover, we have shown that ANY open cover of $f(K)$ has a finite subcover, so $f(K)$ is indeed compact, as needed.

Corollary. *Suppose that (X, d) is a metric space and $f : X \to \mathbb{R}$ is a continuous function. If K is a compact subset of X, then f assumes its maximum and minimum values on K.*

Proof: By the previous theorem, $f(K)$ is compact and therefore closed and bounded by the Heine-Borel Theorem.

Let $S = f(K)$ and let m be the least upper bound of S, which exists by the Completeness Axiom for real numbers. It suffices to show that $m \in S$.

If $m \notin S$, then for any $\epsilon > 0$ there exists and element $s \in S$ such that $m - \epsilon < s < m$, because otherwise $m - \epsilon$ would be an upper bound, contradicting the fact that m is the least upper bound.

Thus m is an accumulation point of S. But since S is closed it contains all of its accumulation points. Therefore $m \in S$. Thus m is a maximum for S and f achieves its maximum value on K.

The argument that f achieves its minimum value is similar.

Example 4. We can now give another important example of metric space. Let

$$X = \mathcal{C}([a,b],\mathbb{R}) = \{f : [a,b] \to \mathbb{R} : \ f \text{ is continuous } \},$$

and define

$$d : \mathcal{C}([a,b],\mathbb{R}) \times \mathcal{C}([a,b],\mathbb{R}) \to \mathcal{R} \quad \text{by} \quad d(f,g) = \sup\{|f(x) - g(x)| : x \in [a,b]\}.$$

The supremum exists by the above corollary. It is straightforward to check that this distance function satisfies the axioms for a metric space.

Using this example, we can apply many of the techniques that we have learned for dealing with the real numbers to spaces of functions. This leads to an important subject—functional analysis—that plays a key role in proving existence of solutions to differential equations.

For example, we can consider Cauchy sequences in $\mathcal{C}([a,b],\mathbb{R})$, and we can prove:

Completeness Theorem. *Every Cauchy sequence in $\mathcal{C}([a,b],\mathbb{R})$ converges.*

This gives a strategy for proving the existence of solutions to differential equitons. Suppose that we want to solve the initial value problem

$$\frac{d}{dx}(y(x)) = f(x, y(x)), \quad y(x_0) = y_0. \tag{3}$$

We can rewrite this as an integral equation

$$y(x) = y_0 + \int_{x_0}^{x} f(\xi, y(\xi)) d\xi.$$

If $y(x) \in \mathcal{C}([x_0, x_0 + \epsilon], \mathbb{R})$, where $\epsilon > 0$ is small, we can define

$$T(y)(x) \in \mathcal{C}([x_0, x_0 + \epsilon], \mathbb{R}) \quad \text{by} \quad T(y)(x) = y_0 + \int_{x_0}^{x} f(\xi, y(\xi)) d\xi.$$

Starting with an approximate solution $y_0(x)$, one can define a sequence of solutions $y_n(x)$ by

$$y_{n+1}(x) = T(y_n)(x).$$

With some effort, one can sometmes show that $(y_n(x))$ is a Cauchy sequence in $\mathcal{C}([a,b],\mathbb{R})$, and hence by the Completeness Theorem converges to a limit $y_\infty(x)$. This is then a solution to the initial-value problem (3).

The key point is that the notion of metric spaces provides an avenue for extending many of the theorems used in the foundations of calculus to settings that allow us to find solutions to differential equations, both ordinary and partial.

4 Compact subsets of $\mathbb{R}^n$

Recall that the Heine-Borel Theorem states that a subset of $\mathbb{R}$ is compact if and only if it is closed and bounded. It is important to realize that the Heine-Borel Theorem also holds for $\mathbb{R}^n$, when $n \geq 2$ and $\mathbb{R}^n$ is given the metric presented in Example 2.

General Heine-Borel Theorem. *A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.*

The proof is beyond the scope of this course. It is clearly useful for several variable calculus, and one of the key theorems proven in Math 118 or Math 145.

References

[1] S. Lay, *Analysis with an introduction to proof*, Fourth edition, Pearson Prentice Hall, Upper Saddle River, NJ, 2005.

[2] W. Rudin, *Principles of mathematical analysis*, Third edition, McGraw-Hill, New York, 1976.