As mentioned in the introduction, the main idea in analysis is to take limits. In we learned to take limits of sequences of real numbers. And in we learned to take limits of functions as a real number approached some other real number.

We want to take limits in more complicated contexts. For example, we might want to have sequences of points in 3-dimensional space. Or perhaps we wish to define continuous functions of several variables. We might even want to define functions on spaces that are a little harder to describe, such as the surface of the earth. We still want to talk about limits there.

Finally, we have seen the limit of a sequence of functions in . We wish to unify all these notions so that we do not have to reprove theorems over and over again in each context. The concept of a metric space is an elementary yet powerful tool in analysis. And while it is not sufficient to describe every type of limit we can find in modern analysis, it gets us very far indeed.

Definition: Metric Space

Let \((X, d)\) be a set and let \(d : X \times X \to \mathbb{R}\) be a function such that

\[ d(x, y) \geq 0 \] for all \(x, y\) in \(X\),

\[ d(x, y) = 0 \] if and only if \(x = y\),

\[ d(x, y) = d(y, x) \] (symmetry),

\[ d(x, z) = d(x, y) + d(y, z) \] (triangle inequality).

Then the pair \((X, d)\) is called a metric space. The function \(d\) is called the metric or sometimes the distance function. Sometimes we just say \(X\) is a metric space if the metric is clear from context.
The geometric idea is that \(d(x)\) is the distance between two points. Items \([\text{metric:pos}]-[\text{metric:com}]\) have obvious geometric interpretation: distance is always nonnegative, the only point that is distance 0 away from \(x\) is \(x\) itself, and finally that the distance from \(x\) to \(y\) is the same as the distance from \(y\) to \(x\). The triangle inequality \([\text{metric:triang}]\) has the interpretation given in

For the purposes of drawing, it is convenient to draw figures and diagrams in the plane and have the metric be the standard distance. However, that is only one particular metric space. Just because a certain fact seems to be clear from drawing a picture does not mean it is true. You might be getting sidetracked by intuition from euclidean geometry, whereas the concept of a metric space is a lot more general.

Let us give some examples of metric spaces.

The set of real numbers \(\{\mathbb{R}\}\) is a metric space with the metric \(d(x,y) := |x-y|\). Items \([\text{metric:pos}]-[\text{metric:com}]\) of the definition are easy to verify. The triangle inequality \([\text{metric:triang}]\) follows immediately from the standard triangle inequality for real numbers: \(d(x,z) = |x-z| = |x-y+y-z| \leq |x-y| + |y-z| = d(x,y) + d(y,z).\) This metric is the standard metric on \(\{\mathbb{R}\}\). If we talk about \(\{\mathbb{R}\}\) as a metric space without mentioning a specific metric, we mean this particular metric.

We can also put a different metric on the set of real numbers. For example take the set of real numbers \(\{\mathbb{R}\}\) to-gether with the metric \(d(x,y) := \frac{|x-y|}{|x-y|+1}\). Items \([\text{metric:pos}]-[\text{metric:com}]\) are again easy to verify. The triangle inequality \([\text{metric:triang}]\) is a little bit more difficult. Note that \(d(x,y) = \varphi(|x-y|)\) where \(\varphi(t) = \frac{t}{t+1}\) and note that \(\varphi\) is an increasing function (positive derivative) hence \(d(x,z) = \varphi(|x-z|) = \varphi(|x-y+y-z|) = \frac{|x-y+y-z|}{|x-y+y-z|+1} = \frac{|x-y|+|y-z|}{|x-y+y-z|+1} = \frac{|x-y|}{|x-y+y-z|+1} + \frac{|y-z|}{|x-y+y-z|+1} \leq d(x,y) + d(y,z)\). Here we have an example of a nonstandard metric on \(\{\mathbb{R}\}\). With this metric we can see for example that \(d(x,y) < 1\) for all \(x,y \in \{\mathbb{R}\}\). That is, any two points are less than 1 unit apart.

An important metric space is the \((n)\)-dimensional euclidean space \(\{\mathbb{R}\}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}\). We use the following notation for points: \((x = (x_1,x_2,\ldots,x_n) \in \{\mathbb{R}\}^n)\). We also simply write \(0 \in \{\mathbb{R}\}^n\) to mean the vector \(0 = (0,0,\ldots,0)\). Before making \(\{\mathbb{R}\}^n\) a metric space, let us prove an important inequality, the so-called Cauchy-Schwarz inequality.

Take \((x = (x_1,x_2,\ldots,x_n) \in \{\mathbb{R}\}^n)\) and \((y = (y_1,y_2,\ldots,y_n) \in \{\mathbb{R}\}^n)\). Then \(\text{biggl(}\sum_{j=1}^n x_j y_j \text{biggr)}^2 \leq \text{biggl(}\sum_{j=1}^n x_j^2 \text{biggr)} \text{biggl(}\sum_{j=1}^n y_j^2 \text{biggr)}\).\)

Any square of a real number is nonnegative. Hence any sum of squares is nonnegative: \(0 \leq \text{biggl(}\sum_{j=1}^n x_j y_j \text{biggr)}^2 \leq \text{biggl(}\sum_{j=1}^n x_j^2 \text{biggr)} \text{biggl(}\sum_{j=1}^n y_j^2 \text{biggr)}\).\)

UC Davis ChemWiki is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License.
\[ y_k \bigl) & = \biggl( \sum_{j=1}^n x_j^2 \biggr) \biggl( \sum_{k=1}^n y_k^2 \biggr) + 2 \biggl( \sum_{j=1}^n y_j \biggr) \biggl( \sum_{k=1}^n x_k \biggr) \biggr) - 2 \biggl( \sum_{j=1}^n x_j y_j \biggr) \biggl( \sum_{k=1}^n x_k y_k \biggr) \end{split} \]

We relabel and divide by 2 to obtain \[ 0 \leq \biggl( \sum_{j=1}^n x_j^2 \biggr) \biggl( \sum_{j=1}^n y_j^2 \biggr) - \left( \sum_{j=1}^n x_j y_j \right)^2 ,\] which is precisely what we wanted.

Let us construct standard metric for \((\mathbb{R}^n)\). Define \[ d(x,y) := \sqrt{ \sum_{j=1}^n (x_j-y_j)^2 } = \sqrt{ (x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_n-y_n)^2 } ,\] For \(n=1\), the real line, this metric agrees with what we did above. Again, the only tricky part of the definition to check is the triangle inequality. It is less messy to work with the square of the metric. In the following, note the use of the Cauchy-Schwarz inequality.

\[
\begin{split}
d(x,z)^2 & = \sum_{j=1}^n (x_j-z_j)^2 \\
& = \sum_{j=1}^n (x_j-y_j+y_j-z_j)^2 \\
& = \sum_{j=1}^n \left( (x_j-y_j)^2 + (y_j-z_j)^2 + 2(x_j-y_j)(y_j-z_j) \right) \\
& = \sum_{j=1}^n (x_j-y_j)^2 + \sum_{j=1}^n (y_j-z_j)^2 + 2 \sqrt{ \sum_{j=1}^n (x_j-y_j)^2 \sum_{j=1}^n (y_j-z_j)^2 } \\
& = \left( \sqrt{ \sum_{j=1}^n (x_j-y_j)^2 } + \sqrt{ \sum_{j=1}^n (y_j-z_j)^2 } \right)^2 = \left( d(x,y) + d(y,z) \right)^2 .
\end{split}
\]

Taking the square root of both sides we obtain the correct inequality.

An example to keep in mind is the so-called discrete metric. Let \(\mathcal{X}\) be any set and define \[ d(x,y) := \begin{cases} 1 & \text{if } x \not= y \\ 0 & \text{if } x = y \end{cases} .\] That is, all points are equally distant from each other. When \(\mathcal{X}\) is a finite set, we can draw a diagram, see for example . Things become subtle when \(\mathcal{X}\) is an infinite set such as the real numbers.

While this particular example seldom comes up in practice, it is gives a useful “smell test.” If you make a statement about metric spaces, try it with the discrete metric. To show that \((\mathcal{X},d)\) is indeed a metric space is left as an exercise.

An example to keep in mind is the so-called discrete metric. Let \(\mathcal{X}\) be any set and define \[ d(x,y) := \begin{cases} 1 & \text{if } x \not= y \\ 0 & \text{if } x = y \end{cases} .\] That is, all points are equally distant from each other. When \(\mathcal{X}\) is a finite set, we can draw a diagram, see for example . Things become subtle when \(\mathcal{X}\) is an infinite set such as the real numbers.

While this particular example seldom comes up in practice, it is gives a useful “smell test.” If you make a statement about metric spaces, try it with the discrete metric. To show that \((\mathcal{X},d)\) is indeed a metric space is left as an exercise.

[example:msC01] Let \(C([a,b])\) be the set of continuous real-valued functions on the interval \([a,b]\). Define the metric on \(C([a,b])\) as \[ d(f,g) := \sup_{x \in [a,b]} |f(x)-g(x)| \] Let us check the properties. First, \(d(f,g)\) is finite as \(\sup_{x \in [a,b]} |f(x)-g(x)|\) is a continuous function on a closed bounded interval \([a,b]\), and so is bounded. It is clear that \(d(f,g) \geq 0\), it is the supremum of nonnegative numbers. If \(f=g\) then \(\sup_{x \in [a,b]} |f(x)-g(x)| = 0\) for all \(x\) and hence \(d(f,g) = 0\). Conversely if \(d(f,g) = 0\), then for any \(x\) we have \(\sup_{x \in [a,b]} |f(x)-g(x)| = 0\) and hence \(|f(x)-g(x)| = 0\) for all \(x\) and \((f=g)\). That \(d(f,g) = d(g,f)\) is equally trivial. To show the triangle inequality we use the standard triangle inequality. \[ \begin{split}
d(f,h) & = \sup_{x \in [a,b]} |f(x)-h(x)+h(x)-g(x)| \\
& = \sup_{x \in [a,b]} |(f(x)-h(x)) + (h(x)-g(x))| \\
& = \sup_{x \in [a,b]} |f(x)-g(x)| + |h(x)-g(x)| \\
& \leq |f(x)-h(x)| + |h(x)-g(x)| \\
& = d(f,g) + d(g,h) \\
\end{split} \] When treat \((C([a,b]))\) as a metric space without mentioning a metric, we mean this particular metric.

This example may seem esoteric at first, but it turns out that working with spaces such as \((C([a,b]))\) is really the meat of a large part of modern analysis. Treating sets of functions as metric spaces allows us to abstract away a lot of the grubby detail and prove powerful results such as Picard’s theorem with less work.

Oftentimes it is useful to consider a subset of a larger metric space as a metric space. We obtain the following proposition, which has a trivial proof.
Let \((X,d)\) be a metric space and \((Y \subset X)\), then the restriction \((d|_{Y \times Y})\) is a metric on \((Y)\).

If \((X,d)\) is a metric space, \((Y \subset X)\), and \((d' := d|_{Y \times Y})\), then \((Y,d')\) is said to be a **subspace** of \((X,d)\).

It is common to simply write \((d|_Y)\) for the metric on \((Y)\), as it is the restriction of the metric on \((X)\). Sometimes we will say that \((d|_Y)\) is the **subspace metric** and that \((Y)\) has the **subspace topology**.

A subset of the real numbers is bounded whenever all its elements are at most some fixed distance from 0. We can also define bounded sets in a metric space. When dealing with an arbitrary metric space there may not be some natural fixed point 0. For the purposes of boundedness it does not matter.

Let \((X,d)\) be a metric space. A subset \((S \subset X)\) is said to be **bounded** if there exists a \((p \in X)\) and a \((B \in \mathbb{R})\) such that 
\[d(p,x) \leq B \quad \text{for all } x \in S.\]
We say that \((X,d)\) is bounded if \((X)\) itself is a bounded subset.

For example, the set of real numbers with the standard metric is not a bounded metric space. It is not hard to see that a subset of the real numbers is bounded in the sense of if and only if it is bounded as a subset of the metric space of real numbers with the standard metric.

On the other hand, if we take the real numbers with the discrete metric, then we obtain a bounded metric space. In fact, any set with the discrete metric is bounded.

**Exercises**

Show that for any set \((\mathbb{X})\), the discrete metric \((d(x,y) = 1)\) if \((x \neq y)\) and \((d(x,x) = 0))\) does give a metric space \((\mathbb{X},d))\).

Let \((\mathbb{X} := \{ 0 \})\) be a set. Can you make it into a metric space?

Let \((\mathbb{X} := \{ a, b \})\) be a set. Can you make it into two distinct metric spaces? (define two distinct metrics on it)

Let the set \((\mathbb{X} := \{ A, B, C \})\) represent 3 buildings on campus. Suppose we wish to our distance to be the time it takes to walk from one building to the other. It takes 5 minutes either way between buildings \((\mathbb{A})\) and \((\mathbb{B})\). However, building \((\mathbb{C})\) is on a hill and it takes 10 minutes from \((\mathbb{A})\) and 15 minutes from \((\mathbb{B})\) to get to \((\mathbb{C})\). On the other hand it takes 5 minutes to go from \((\mathbb{C})\) to \((\mathbb{A})\) and 7 minutes to go from \((\mathbb{C})\) to \((\mathbb{B})\), as we are going downhill. Do these distances define a metric? If so, prove it, if not say why not.

Suppose that \((\mathbb{(X,d)})\) is a metric space and \((\mathbb{\varphi} \mathbb{\colon} [0,\mathbb{\text{infty}}] \to \mathbb{\mathbb{R}})\) is an increasing function such that \((\mathbb{\varphi}(t) \geq 0)\) for all \((t)\) and \((\mathbb{\varphi}(t) = 0)\) if and only if \((t=0)\). Also suppose that \((\mathbb{\varphi})\) is **subadditive**, that is \((\mathbb{\varphi}(s+t) \leq \mathbb{\varphi}(s) + \mathbb{\varphi}(t))\). Show that with \((d'(x,y) := \mathbb{\varphi}(d(x,y)))\), we obtain a new metric space \((\mathbb{(X,d')}))\).

Let \((\mathbb{(X,d_X)})\) and \((\mathbb{(Y,d_Y)})\) be metric spaces.
a) Show that \( (X \times Y, d) \) with \( d((x_1, y_1), (x_2, y_2)) := d_X(x_1, x_2) + d_Y(y_1, y_2) \) is a metric space.

b) Show that \( (X \times Y, d) \) with \( d((x_1, y_1), (x_2, y_2)) := \max \{ d_X(x_1, x_2), d_Y(y_1, y_2) \} \) is a metric space.

Let \( (X) \) be the set of continuous functions on \([0,1]\). Let \( \phi : [0,1] \to (0,\infty) \) be continuous. Define \[ d(f,g) := \int_0^1 |f(x) - g(x)| \phi(x) \, dx \] Show that \( ((X), d) \) is a metric space.

Let \( (X, d) \) be a metric space. For nonempty bounded subsets \( A \) and \( B \) let \[ d(x,B) := \inf \{ d(x,b) : b \in B \} \quad \text{and} \quad d(A,B) := \sup \{ d(a,B) : a \in A \} \] Now define the Hausdorff metric as \[ d_H(A,B) := \max \{ d(A,B), \ d(B,A) \} \] Note: \( d_H \) can be defined for arbitrary nonempty subsets if we allow the extended reals.

a) Let \( (Y, d_H) \) be the set of bounded nonempty subsets. Show that \( ((Y, d_H)) \) is a metric space. b) Show by example that \( ((d)) \) itself is not a metric. That is, \( (d) \) is not always symmetric.

---

**Contributors and Attributions**

- [Template:ContribLebl](#)