The Contraction Mapping Theorem

Andrew Paul

3/22/2020

1 Homotheties as Contraction Mappings

We begin by setting some intuitive ground for what we've set out to prove. Consider a homothety. (also known as a dilation). This takes all points in some space[1], and multiplies all of their distances to some point defined as the *center* by a scale factor k. For example, here is a homothety in the plane.

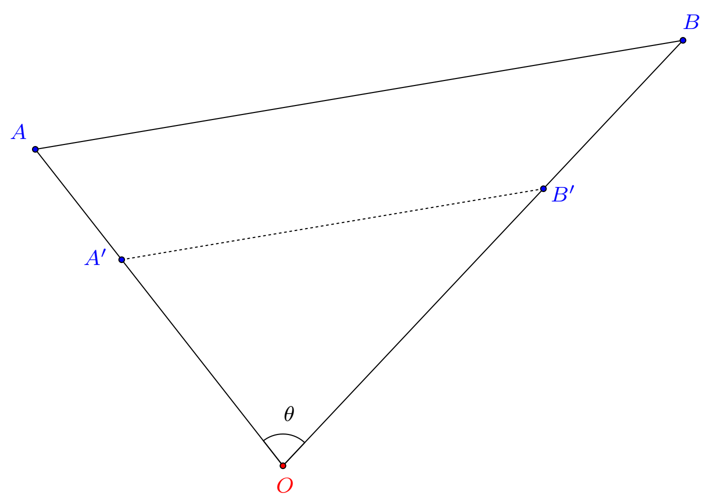

Figure 1: This is a homothety centered at O with $0 < k < 1$.

Observe that O is a fixed point in this homothety. This makes sense, since intuitively we can see that since we are forcing everything to "collapse" into O, O itself must not move. More formally, since $0 \cdot k = 0$, and due to the uniqueness of the limit, the only point that is at a distance of 0 from itself is itself.

A homothety is just one type of *contraction mapping*. As its name suggests, a contraction mapping is a function that "squishes together" the points in its domain. We define a contraction mapping as follows.

[1] To be more precise, we are talking about a *complete metric space*. We will have more to say about this later.

Definition. *A contraction mapping is a continuous function* $f\colon \overline{U} \to \overline{U}$ *for which* $\exists c \in (0,1)$ *such that*

$$|f(\vec{x}) - f(\vec{y})| \le c|\vec{x} - \vec{y}|.$$

From this definition, it is easy to see that a homothety with scale factor $|k| < 1$ is a contraction mapping.

Proof. Consider a homothety h centered at O and two arbitrary points A and B. WLOG let $0 < k < 1$ as shown in Figure 1. The case for $-1 < k < 0$ is essentially the same and is geometrically just a reflection of the case that we are considering.

By the law of cosines, in the preimage we have

$$d(A,B)^2 = d(A,O)^2 + d(B,O)^2 - 2d(A,O)d(B,O)\cos\theta.$$

Since homotheties preserve angles, in the image we must have that

$$\begin{aligned} d(h(A),h(B))^2 &= d(h(A),h(O))^2 + d(h(B),h(O))^2 - 2d(h(A),h(O))d(h(B),h(O))\cos\theta \\ &= k^2[d(A,O)^2 + d(B,O)^2 - 2d(A,O)d(B,O)\cos\theta] \\ &= k^2 d(A,B)^2, \end{aligned}$$

so $d(h(A),h(B)) = kd(A,B)$. Hence, we can choose $c = k + \frac{1-k}{2}$, and thus h is a contraction mapping. □

Here, we have essentially proven the old eighth-grade geometry adage: a dilation with (positive) scale factor k multiplies the perimeter by k.

So then what is the contraction mapping theorem?

Theorem (Contraction Mapping Theorem[2])**.** *Let* $f\colon \overline{U} \to \overline{U}$ *be a contraction mapping. Then there exists a fixed point in* $\overline{U}$*.*

As we observed before, O, the center of a homothety with $|k| < 1$, is a fixed point. Since we have shown that such a homothety is just a type of contraction mapping, the contraction mapping theorem essentially generalizes this result. The homothety makes it clear how to visualize what the contraction mapping theorem says. Essentially, a more general "squishing function" still has a "center" that does not move.

This (poorly drawn) diagram illustrates this principle.

[2]This theorem is sometimes called the *Banach fixed-point theorem*.

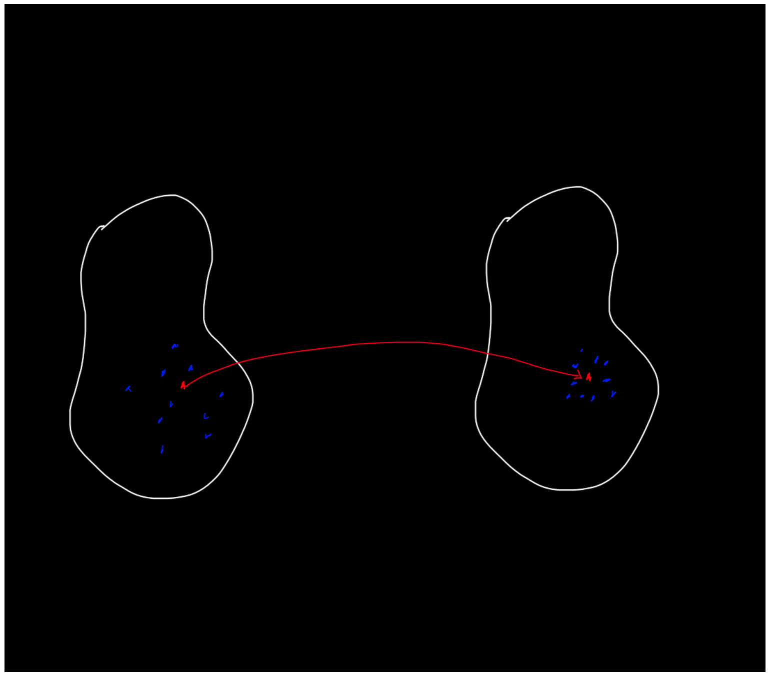

Figure 2: This is an example of a contraction mapping. Shown is a fixed point (in red) and some points in a neighborhood of it (in blue). CMT guarantees that such a red point exists.

Let us prove this.

2 Completeness (Space Is Not Spongy)

There is an important property of *space* that we must discuss before diving in to CMT. This is the completeness of the metric space that we are working in.

For instance, consider the real number line. The real numbers are *Cauchy complete*. Informally, this means that there are no "holes" in the real number line. The same cannot be true of, say, the set of rational numbers. For instance, consider the sequence

$$1, 1.4, 1.41, 1.414, ..., \sqrt{2},$$

where the n^{th} term is given by the n^{th} decimal approximation of $\sqrt{2}$. Clearly, this sequence converges to $\sqrt{2}$, but this is irrational. To be more precise, we introduce the notions of a *metric space* and a *Cauchy sequence*.

Definition. *Let V be a set. V is a metric space if and only if there exists a function $d\colon V \times V \to [0, \infty)$ such that*

(1) $d(x, y) = d(y, x)$.

(2) $d(x, y) \geq 0$.

(3) $d(x, y) = 0 \Leftrightarrow x = y$.

(4) $d(x, z) \leq d(x, y) + d(y, z)$.

The function d is called a metric.

Observe that the metric simply encodes the intuitive notions of distance. In particular, the distance between points is always nonnegative, the distance between points A and B is the distance between points B and A, and the only point that is at a distance of 0 from another point is that point itself. The last inequality is simply the triangle inequality.

Definition. *Consider the sequence $i \mapsto a_i \in V$ where V is a metric space. The sequence is Cauchy if and only if $\forall \epsilon > 0$, $\exists N > 0$ such that $\forall m, n > N$, we have*

$$|a_m - a_n| < \epsilon.$$

This looks very similar to the condition for the convergence of a series. A notable difference is that the classical definition of convergence uses the limit in the definition. In fact, a Cauchy sequence is equivalent to a convergent sequence in a *complete metric space*.

Definition. *Let V be a metric space. If every Cauchy sequence in V also converges in V, then V is complete.*

Intuitively, a complete metric space has no "holes" in it. As an example, the real number line $\mathbb{R}$ is complete. This is because every Cauchy sequence of real numbers converges to a real number.

In particular, it is important to note that there is no distinction between a Cauchy sequence and a convergent sequence in a complete metric space. We can use this to our advantage to prove that a sequence converges without knowing its limit!

This is all important because if the space we are working in had "holes" in it, CMT would not work. In fact, the very notion of continuity falls apart if space was spongy.

3 The Main Proof

Suppose that $f : \overline{U} \to \overline{U}$ is a contraction mapping. Let us define a sequence in $\overline{U}$ by $a_0 \in \overline{U}$ and $a_i = f(a_{i-1})$. Since f is a contraction mapping, we have that

$$|a_{i+1} - a_i| = |f(a_i) - f(a_{i-1})| \leq c|a_i - a_{i-1}|.$$

For some $c \in (0, 1)$. This implies that

$$|a_{i+1} - a_i| < c|a_i - a_{i-1}| < c^2|a_{i-1} - a_{i-2}| < ... < c^i|a_1 - a_0|.$$

Hence, $\forall m > n$,

$$
\begin{aligned}
|a_m - a_n| &= |a_m - a_{m-1} + a_{m-1} - a_{m-2} + a_{m-2} - \ldots - a_{n+1} + a_{n+1} - a_n| \\
&\leq |a_m - a_{m-1}| + |a_{m-1} + a_{m-2}| + \ldots + |a_{n+1} - a_n| \\
&\leq \left(c^{m-1} + c^{m-2} + \ldots + c^n\right) |a_1 - a_0| \\
&= c^n \left(1 + c + \ldots + c^{m-n-1}\right) |a_1 - a_0| \\
&\leq \frac{c^n}{1-c} |a_1 - a_0|.
\end{aligned}
$$

In the second line, we applied the triangle inequality, and in the last line we use the fact that an infinite geometric series with a positive common ratio has monotonically increasing partial sums. It follows that we can choose

$$N > \left\lceil \log_c \left(\frac{1-c}{|a_1 - a_0|} \epsilon \right) \right\rceil,$$

for any positive ϵ and have $|a_m - a_n| < \epsilon$, $\forall m, n > N$. Hence, the sequence $i \mapsto a_i$ is Cauchy. Since $\overline{U}$ is a complete metric space, the sequence is also convergent.

Let $x = \lim_{i \to \infty} a_i$. Since f is continuous, the sequence $a_1, a_2, \ldots = f(a_0), f(a_1), \ldots$ converges to $f(x)$. But the limit of a subsequence of a particular sequence must converge to the same limit as the sequence, and this limit is unique. In particular, we have that

$$x = f(x).$$

Furthermore, since $\overline{U}$ is closed, we have that $x \in \overline{U}$. Hence x is a fixed point of f. QED.

4 Uniqueness of The Fixed Point

From here, it is not hard to show that the fixed point that we have found must be unique. In other words, there must be exactly *one* center to any contraction mapping. Intuitively, this makes sense. Consider the following vector field.

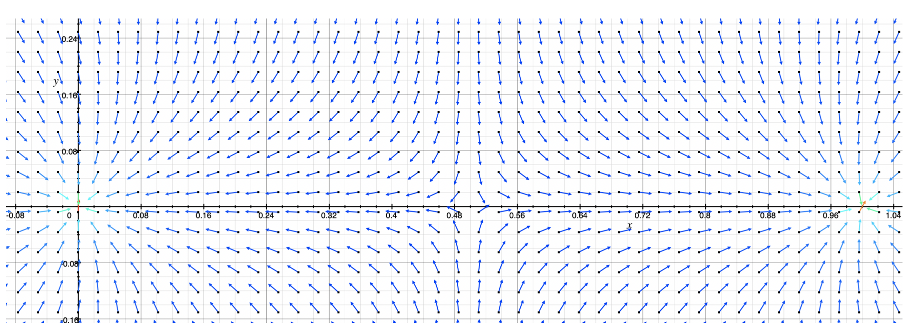

Figure 3: This vector field has a negative divergence with two "sinks". It does not represent a contraction mapping.

In this case, we have two "centers", one at the origin and one at $(1, 0)$, where both centers have the same “strength”. But the problem becomes apparent when we look at a small neighborhood of $\left(\frac{1}{2}, 0\right)$. These points hardly move. Moreover, it is definitely possible to choose two points in that neighborhood so that one of them gets closer to the origin and the other gets closer to $(1, 0)$, clearly contradicting the definition of a contraction mapping.

Intuitively, if two centers have to “compete” with each other, there will exist points in between that are not squished together in the mapping - instead they are pulled apart.

Formally, we can show that the fixed point must be unique as follows.

Proof. Suppose that x_1 and x_2 are fixed points of the contraction mapping f. Then,

$$|f(x_1) - f(x_2)| = |x_1 - x_2| \leq c|x_1 - x_2|.$$

Since $c \in (0, 1)$, we are forced to have $|x_1 - x_2| = 0$ or $x_1 = x_2$. □

5 Conclusion

CMT is not only a very natural result - it is very powerful too. Among other things, it can guarantee the existence of solutions to certain differential and integral equations (for instance, via the Picard-Lindelöf theorem) and it can be used to prove the inverse function theorem in calculus. The techniques used to prove CMT are in the style of analysis, and it certainly provides a good δ-ϵ workout.

Two prominent questions remain.

(a) What are some examples of contraction mappings besides homotheties? Figure 3 was motivated by the gravitational field of two point masses, though it seems like the gravitational field of even a single point mass is still not a contraction mapping. If not inverse-square, what “force functions” produce vector fields that represent contraction mappings?

(b) Can the intuitive reasoning of “competing centers can’t constitute contraction mappings” be formalized? If not, it still makes for some great alliteration.

Any progress on these questions will be duly reported!