COMPLETE METRIC SPACES AND THE CONTRACTION MAPPING THEOREM

A metric space (M, d) is a set M with a metric $d(x, y) \geq 0$, $x, y \in M$ that has the properties

$$d(x, y) = d(y, x) \ , x, y \in M$$

$$d(x, y) \leq d(x, z) + d(z, y) \ , x, y, z \in M \ \text{(triangle inequality)}$$

and

$$d(x, y) = 0 \iff x = y$$

Most of you have seen these definitions and so I will not go into any details.

A sequence of points $x_n \in M$ is a **Cauchy Sequence** if for any $\varepsilon > 0$ there exists $N(\varepsilon)$ such that

$$d(x_n, x_m) < \varepsilon$$

for all $n, m > N(\varepsilon)$. Accordingly we say that a complete metric space is **complete** if every Cauchy Sequence converges to some element $x \in M$, i.e., for every $\varepsilon > 0$ there exists $N(\varepsilon)$ such that

$$d(x_n, x) < \varepsilon$$

for all $n > N(\varepsilon)$.

A function $f : M \to M$ is a **contraction** if there exists a constant $0 \leq \alpha < 1$ such that for all $x, y \in M$

$$d(f(x), f(y)) \leq \alpha d(x, y) \ .$$

A simple consequence of these definitions is the **Banach fixed point theorem**:

Theorem 0.1. *Let (M, d) be a complete metric space and $f : M \to M$ a contraction. Then the equation*

$$x = f(x)$$

has a unique solution $\overline{x}$. Moreover, if $x_0 \in M$ is any initial point and $x_{n+1} = f(x_n), n = 0, 1, \ldots,$ then

$$d(\overline{x}, x_n) \leq \frac{\alpha^n}{1-\alpha} d(f(x_0), x_0)$$

Proof. If $\overline{x}$ and $\overline{y}$ are two solutions then

$$d(\overline{x}, \overline{y}) = d(f(\overline{x}), f(\overline{y})) \leq \alpha d(\overline{x}, \overline{y})$$

and hence $d(\overline{x}, \overline{y}) = 0$ and therefore $\overline{x} = \overline{y}$. Pick any $n > m$ and use the triangle inequality to find

$$d(x_n, x_m) \leq \sum_{k=m}^{n-1} d(x_{k+1}, x_k)$$

Moreover,

$$d(x_{k+1}, x_k) = d(f(x_k), f(x_{k-1})) \leq \alpha d(x_k, x_{k-1}) \leq \alpha^k d(f(x_0), x_0)$$

and so

$$\sum_{k=m}^{n-1} d(x_{k+1}, x_k) \leq \sum_{k=m}^{n-1} \alpha^k d(f(x_0), x_0) = \alpha^m \frac{1-\alpha^{n-m}}{1-\alpha} d(f(x_0), x_0) \leq \frac{\alpha^m}{1-\alpha} d(f(x_0), x_0) \ .$$

Hence

$$d(x_n, x_m) \leq \frac{\alpha^m}{1-\alpha} d(f(x_0), x_0)$$

which proves that the sequence x_n is a Cauchy sequence. The completeness of (M, d) guarantees that x_n has a limit $\overline{x}$. We have to show that $\overline{x}$ solves the equation. Pick $n \geq 1$ arbitrary and use that

$$d(f(\overline{x}), \overline{x}) \leq d(f(\overline{x}), x_n) + d(x_n, \overline{x}) = d(f(\overline{x}), f(x_{n-1})) + d(x_n, \overline{x})$$

which is bounded above by

$$\alpha d(\overline{x}, x_{n-1}) + d(x_n, \overline{x}) \leq \frac{2\alpha^n}{1-\alpha} d(f(x_0), x_0) \ .$$

Because n is arbitrary $d(f(\overline{x}), \overline{x})$ is smaller than any positive number and hence equal to 0. □

Example 1: For a simple example of a metric space that is not necessarily complete consider any set $S \subset \mathbb{R}^d$ and consider the Euclidean distance

$$d(x, y) = |x - y| = \sqrt{\sum_{j=1}^{d} (x_j - y_j)^2} \ .$$

It is obvious that $d(x, y) = d(y, x) \geq 0$ for all $x, y \in S$. Further

$$0 = d(x, y) = |x - y|$$

implies that x and y have the same components and hence are equal. The triangle inequality follows from the following facts:

Schwarz' inequality

$$x \cdot y := \sum_{j=1}^{d} x_j y_j \leq \sqrt{\sum_{j=1}^{d} x_j^2} \sqrt{\sum_{j=1}^{d} y_j^2} = |x||y|$$

Minkowski's inequality

$$|x + y| \leq |x| + |y|$$

This follows easily from Schwarz' inequality. Thus, we find that for any $x, y, z \in S$

$$d(x, y) = |x - y| = |x - z + z - y| \leq |x - z| + |z - y| = d(x, z) + d(z, y) \ .$$

In this context, the following is an interesting application of the contraction mapping theorem. We start first with an easy case. A map is called **Lipschitz**, if there exists a constant L such that for all x_1, x_2

$$|f(x - 1) - f(x_2)| \leq L|x_1 - x_2| \ .$$

Thus, a contraction is a Lipschitz map with Lipschitz constant $L < 1$.

Given a map

$$f : \mathbb{R}^d \to \mathbb{R}^d$$

and consider the map

$$h : \mathbb{R}^d \to \mathbb{R}^d$$

given by $h(x) = x + f(x)$. Assume that f is a contraction, i.e.,

$$|f(x) - f(y)| \leq \alpha |x - y|$$

for some constant $\alpha < 1$. We claim that h has an inverse which is also a contraction

To see this we have to show two things.

a) h is injective.

This follows from the fact that

$$x_1 + f(x_1) = x_2 + f(x_2)$$

entails that

$$|x_1 - x_2| = |f(x_2) - f(x_1)| \leq \alpha |x_1 - x_2|$$

which yields $x_1 = x_2$.

b) Next we have to show that h is onto. For any given $y \in \mathbb{R}^d$ we consider the equation

$$y = x + f(x)$$

which we rewrite as

$$x = y - f(x) := \phi(x) \ .$$

The map ϕ is a contraction

$$|\phi(x_1) - \phi(x_2)| = |f(x_1) - f(x_2)| \leq \alpha |x_1 - x_2|$$

and hence there exists a unique fixed point $a \in \mathbb{R}^d$, i.e.,

$$a = \phi(a) = y - f(a) \ .$$

Hence h has an inverse, which we denote by $g : \mathbb{R}^d \to \mathbb{R}^d$. To show that g is Lipschitz we write $y_i = h(x_i), i = 1, 2$ and note that

$$|x_1 - x_2| \leq |y_1 - y_2| + |f(x_1) - f(x_2)| \leq |y_1 - y_2| + \alpha |x_1 - x_2|$$

so that

$$|x_1 - x_2| \leq \frac{1}{1 - \alpha} |y_1 - y_2|$$

which shows that g is Lipschitz with Lipschitz constant $\frac{1}{1-\alpha}$. This argument can be adapted to a more general situation.

Theorem 0.2. *Imagine an open set $S \subset \mathbb{R}^d$ and let*

$$f : S \to \mathbb{R}^d$$

be a contraction with contraction constant $\alpha < 1$. Then for the map

$$h : S \to h(S) \ , h(x) = x + f(x)$$

$h(S)$ is open and the map h has an inverse $g : h(S) \to S$ which is Lipschitz with Lipschitz constant $\frac{1}{1-\alpha}$.

Proof. The fact that h is injective has the same proof as before. A priori we do not know much about the set $h(S)$. We prove that this set is open in $\mathbb{R}^d$. Pick any $y_0 \in h(S)$. Then, by definition, there exists a point $x_0 \in S$ so that $h(x_0) = y_0$. To arrange things in a convenient way we set

$$U(x) = h(x_0 + x) - y_0 = x + f(x_0 + x) + x_0 - y_0 = x + V(x)$$

so that $U(0) = 0$ i.e., U fixes the origin. Hence

$$U : S - x_0 \to h(S) - y_0 \ ,$$

and our goal is to show that $U(S-x_0)$ is an open set. Pick $r>0$ so that the closed ball $\overline{B}_r(0) \subset S - x_0$ and note that

$$|V(x)| = |V(x) - V(0)| = |f(x+x_0) - f(x_0)| \le \alpha |x|$$

so that V maps the ball $\overline{B}_r(0)$ into the ball $\overline{B}_{\alpha r}(0) \subset B_r(0)$. Such a radius r exists, because $S - x_0$ is open. Indeed pick r' so that the open ball $B_{r'}(0) \subset S - x_0$ and pick any $0 < r < r'$ which assures that $\overline{B}_r(0) \subset S - x_0$. Hence, V is a map of the metric space $\overline{B}_r(0)$ into itself. Moreover, V is a contraction on $\overline{B}_r(0)$. Indeed for $x_1, x_2 \in \overline{B}_r(0)$ we have that

$$|V(x_1) - V(x_2)| = |f(x_1 + x_0) - f(x_2 + x_0)| \le \alpha |x_1 - x_2| \ .$$

If we can show that any point $y \in \overline{B}_r(0)$ is of the form $U(z)$ for some $z \in \overline{B}_r(0)$ we are done. Thus, we have to find $z \in \overline{B}_r(0)$ so that

$$y = z + V(z)$$

i.e., the map $y - V(x)$ has a fixed point in $\overline{B}_r(0)$. Note that $\overline{B}_r(0)$ is closed and hence is a complete metric space. Thus, by the fixed point theorem there exists $z \in \overline{B}_r(0)$ with the desired properties. Denoting the inverse by $g : h(S) \to S$ we have for $y_1, y_2 \in h(S)$, setting $x_i = g(y_i), i = 1, 2$,

$$|x_1 - x_2| = |(y_1 - f(x_1)) - (y_2 - f(x_2))| \le |y_1 - y_2| + |f(x_1) - f(x_2)| \le |y_1 - y_2| + \alpha |x_1 - x_2|$$

so that

$$|x_1 - x_2| \le \frac{1}{1-\alpha} |y_1 - y_2| \ ,$$

which shows that g is a Lipschitz map with Lipschitz constant $\frac{1}{1-\alpha}$. □

A consequence of this Theorem is the inverse function theorem.

Theorem 0.3. *Let $S \subset \mathbb{R}^d$ be an open set and $F : S \to \mathbb{R}^d$ a map that is continuously differentiable. Assume that the Jacobi matrix $DF(x_0)$, $x_0 \in S$, is invertible. Then there exists an open set $U \subset \mathbb{R}^d$ with $x_0 \in U$ such that $F(U)$ is open and there exists a map $g : F(U) \to U$ such that $g \circ F = id$. Moreover, g is differentiable at $F(x_0)$ and we have that*

$$Dg(F(x_0)) = DF(x_0)^{-1}$$

Proof. We have to construct U. First we normalize things conveniently. By replacing $F(x)$ by $DF(x_0)^{-1}F(x)$ we may assume that $DF(x_0) = I$. Further, replacing $F(x)$ by $F(x+x_0) - F(x_0)$ we may assume that $x_0 = 0$ and $F(x_0) = 0$. Let's denote this renormalized map by h. Since h is continuously differentiable we have that

$$h(x) - h(0) = \int_0^1 \frac{d}{dt} h(tx) dt = \int_0^1 Dh(tx) dt \cdot x$$

which leads to

$$h(x) = x + f(x)$$

where

$$f(x) := \int_0^1 (Dh(tx) - I) dt \cdot x \ .$$

It is convenient to set

$$M_{i,j}(x) = (Dh(x) - I)_{i,j} \ .$$

Since $Dh(x)$ is continuous at 0 we can find $r > 0$ so that

$$\max_{i,j} \sup_{|x|\le 3r} |M_{i,j}(x)| \le \frac{1}{2d} \ .$$

Hence we have that

$$|f(x)| = \sqrt{\sum_i (\sum_j \int_0^t M_{i,j}(tx)dtx_j)^2} \le \frac{1}{2d}\sqrt{\sum_i(\sum_j |x_j|)^2} = \frac{1}{2d}\sqrt{d(\sum_j |x_j|)^2} \le \frac{1}{2d}\sqrt{d^2(\sum_j |x_j|^2)} = \frac{1}{2}|x|$$

and we see that $f(B_{2r}(0)) \subset B_r(0)$ and in particular $f(\overline{B}_r(0)) \subset \overline{B}_r(0)$. Further for $x_1, x_2 \in B_r(0)$ we have that

$$f(x_1) - f(x_2) = h(x_1) - h(x_2) - x_1 + x_2 = \int_0^1 [Dh((1-t)x_2 + tx_1) - I]\, dt(x_1 - x_2)$$

and

$$|(1-t)x_2 + tx_1| \le |x_2| + t|x_1 - x_2| \le 3r$$

and hence

$$|f(x_1) - f(x_2)| \le \frac{1}{2}|x_1 - x_2| \ .$$

Thus, $f : \overline{B}_r(0) \to \overline{B}_r(0)$ is a contraction and therefore $V = f(B_r(0))$ is open and $h : V \to \overline{B}_r(0)$ has a Lipschitz continuous inverse, which we denote again by g. To see that g is differentiable at 0 we shall show that

$$|g(x) - x| = o(|x|) \ .$$

This implies that g is differentiable at 0 and $Dg(0) = I$ as it should be. Pick any sequence $x_n \to 0$, set $y_n = g(x_n)$. Note that $y_n \to 0$ as well and we compute

$$\frac{|g(x_n) - x_n|}{|x_n|} = \frac{|y_n - h(y_n)|}{|y_n|}\frac{|g(x_n)|}{|x_n|} \ .$$

Because $g(0) = 0$, we have that

$$\frac{|g(x_n)|}{|x_n|} \le 2$$

and hence

$$\lim_{n\to\infty} \frac{|g(x_n) - x_n|}{|x_n|} = 0 \ .$$

□