Additional lecture notes for Math 63CM, Version 2020

Lenya Ryzhik*

January 14, 2020

1 Complete metric spaces

1.1 Basic properties of metric spaces

As a warm-up, let us briefly recall some very basic properties of metric spaces covered in 61CM.

Definition 1.1 *A metric space (X, d) is a set X together with a map $d : X \times X \to \mathbb{R}$ (called a distance function) such that*

1. *$d(x, y) \geq 0$ for all $x, y \in X$, and $d(x, y) = 0$ if and only if $x = y$.*
2. *$d(x, y) = d(y, x)$ for all $x, y \in X$,*
3. *(Triangle inequality) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.*

Recall that a norm $\|.\| : V \to \mathbb{R}$ on a vector space V over $\mathbb{R}$ (or $\mathbb{C}$) is a map that is

1. positive definite, i.e. $\|x\| \geq 0$ for all $x \in V$ with equality if and only if $x = 0$,
2. absolutely homogeneous, i.e. $\|\lambda x\| = |\lambda| \, \|x\|$ for $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) and $x \in V$,
3. satisfies the triangle inequality, i.e. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$.

Then one easily checks that every normed vector space is a metric space with the induced metric $d(x, y) = \|x - y\|$; for instance the triangle inequality for the metric follows from

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

where the inequality in the middle is the triangle inequality for norms.

There are many interesting metric spaces that are *not* normed vector spaces, for the simple reason that the distance function does not require that X is a vector space. For instance, for any set X we may define a metric on it setting $d(x, y) = 0$ if $x = y$, $d(x, y) = 1$ if $x \neq y$.

Another example of a metric space we will use often is the space $C(K)$ of continuous real-valued functions $f : K \to \mathbb{R}$, where K is a compact subset of $\mathbb{R}^n$. The norm on $C(K)$ is

$$\|f\|_{C(K)} = \sup_{x \in K} |f(x)|,$$

and the corresponding distance is defined as

$$d(f, g) = \sup_{x \in K} |f(x) - g(x)|.$$

*Department of Mathematics, Stanford University, Stanford, CA 94305, USA; ryzhik@math.stanford.edu

We will often simply look at $K = [a, b]$, a closed interval on $\mathbb{R}$. However, we can define different norms that lead to different distances on this space. For example, we can define the norm of a continuous function defined on a Riemann measurable set K as

$$\|f\|_{L^1(K)} = \int_K |f(x)|dx,$$

and the corresponding distance between two continuous functions f and g as

$$d_1(f, g) = \int_K |f(x) - g(x)|dx.$$

We will denote the space of continuous functions on K with this metric as $\tilde{L}_1(K)$.

Exercise 1.2 Show that d_1 is, indeed, a metric on the set of continuous functions on K. Next, show that if the set K is bounded, then there exists a constant $C > 0$ such that $d_1(f, g) \leq Cd(f, g)$ for all continuous functions f and g defined on K. Finally, show that if K is a bounded Riemann measurable set, and a sequence f_n converges in $C(K)$, then it converges also in $\tilde{L}_1(K)$.

Convergence in metric spaces is defined exactly as in $\mathbb{R}^n$, except we use the metric $d(x, y)$ rather than $\|x - y\|$.

Definition 1.3 *A sequence x_n of points in a metric space (X, d) converges to $x \in X$, denoted as $\lim x_n = x$, if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $d(x_n, x) < \varepsilon$.*

1.2 The Cauchy sequences

Existence of the limit of a sequence is often difficult to establish using the definition of the limit directly. A very useful tool is the Cauchy criterion for convergence that is phrased purely in terms of the elements of the sequence.

Definition 1.4 *We say that a_n in a metric space (X, d) is a Cauchy sequence if for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ so that $d(x_n, x_m) < \varepsilon$ for all $n, m \geq N$.*

It is easy to see that a convergent sequence is always Cauchy.

Theorem 1.5 *If a sequence a_n converges in a metric space (X, d), then a_n is a Cauchy sequence.*

Proof. Assume that a_n converges and $A = \lim_{n\to\infty} a_n$. Given $\varepsilon > 0$ we can find $N \in \mathbb{N}$ so that $d(a_n, A) < \varepsilon/2$ for all $n \geq N$. Then, for all $n, m \geq N$ we have

$$d(a_n, a_m) \leq d(a_n, A) + d(A, a_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

hence a_n is a Cauchy sequence. □

Next, let us first show that on $\mathbb{R}$ the converse is also true: if a sequence is Cauchy then it converges.

Theorem 1.6 *A sequence $a_n \in \mathbb{R}$ converges if and only if a_n is a Cauchy sequence.*

Proof. We have already shown that a convergent sequence is Cauchy in any metric space, so we only need to show that on $\mathbb{R}$ every Cauchy sequence is convergent. Let us assume that a_n is a Cauchy sequence. First, we claim that a_n is bounded. Indeed, taking $\varepsilon = 1$, we can find N such that $|a_n - a_m| < 1$ for all $n, m \geq N$. In particular, we have $a_N - 1 < a_m < a_N + 1$ for all $m \geq N$.

In addition, there are only finitely many elements of the sequence a_n with $n < N$, so the set $\{a_1, a_2, \dots, a_{N-1}\}$ is a bounded set. Hence, $\{a_n\}$ is a union of two bounded sets, hence it is also bounded, and the sequence a_n is bounded as well. Thus, for each $n \in \mathbb{N}$ we can define

$$x_n = \inf_{k \geq n} a_k, \quad y_n = \sup_{k \geq n} a_k.$$

It is clear from the definition that $x_n \leq x_{n+1} \leq y_{n+1} \leq y_n$ for all $n \in \mathbb{N}$, so that the sequence x_n is increasing and the sequence y_n is decreasing. By the nested intervals theorem there exists a point A common to all intervals $[x_n, y_n]$:

$$x_n \leq A \leq y_n \text{ for all } n \in \mathbb{N}.$$

In addition, we have

$$x_n \leq a_n \leq y_n \text{ for all } n \in \mathbb{N},$$

and it follows that

$$|A - a_n| \leq |x_n - y_n|. \tag{1.1}$$

However, given any $\varepsilon > 0$ we can find N so that for all $m, N \geq N$ we have

$$|a_n - a_m| < \frac{\varepsilon}{10},$$

and in particular,

$$|a_n - a_N| < \frac{\varepsilon}{10}.$$

Now, it follows from the definition of x_n and y_n that for all $n \geq N$ we have

$$|x_n - a_N| \leq \frac{\varepsilon}{10}, \quad |y_n - a_N| \leq \frac{\varepsilon}{10},$$

hence

$$|x_n - y_n| \leq \frac{2\varepsilon}{10} < \varepsilon.$$

We conclude from this and (1.1) that $|A - a_n| < \varepsilon$ for all $n \geq N$, thus a_n converges to A as $n \to +\infty$. $\Box$

Exercise 1.7 Use the Cauchy criterion to show that the sequence

$$a_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}$$

does not converge. Hint: show that $a_{2n} - a_n \geq 1/2$ for all n.

Exercise 1.8 Show that the claim of Theorem 1.6 does not hold for the set $\mathbb{Q}$ of rational numbers. In other words, show that a Cauchy sequence of rational numbers does not necessarily converge to a rational number.

Exercise 1.9 Show that if a sequence a_k in a metric space (X, d) is Cauchy and a subsequence a_{n_k} converges to $A \in X$, then the sequence a_n converges to X. In other words, a Cauchy sequence that has a converging subsequence must converge.

Let us give another example of a Cauchy sequence that does not converge. Consider the following sequence of continuous functions

$$f_n(x) = \begin{cases} 1, & 0 \le x \le 1/2 - 1/n, \\ 2n(1/2 - x - 1/(2n)), & 1/2 - 1/n \le x \le 1/2 - 1/(2n), \\ 0, & 1/2 - 1/(2n) \le x \le 1. \end{cases}$$

The sequence f_n is not Cauchy in the space $C[0,1]$. This is because for $m > 2n+1$ we have $f_m(1/2 - 1/(2n)) = 1$ but $f_n(1/2 - 1/(2n)) = 0$, so that $\|f_n - f_m\|_{C[0,1]} = 1$. However, it is Cauchy in the space $\tilde{L}_1[0,1]$. Indeed, for $m \ge n$ we have

$$\|f_n - f_m\|_{L^1([0,1])} \le \int_{1/2-1/n}^{1/2} |f_n - f_m| dx \le \int_{1/2-1/n}^{1/2} 1 \cdot dx \le \frac{1}{n}.$$

It follows that f_n is a Cauchy sequence in the space $\tilde{L}_1[0,1]$. However, it can not converge in this metric to a limit that is a continuous function on $[0,1]$. Indeed, assume that f_n converges to a function g in $\tilde{L}_1[0,1]$. We claim that then we must have both that $g(x) = 0$ for all $x \in (1/2, 1]$ and $g(x) = 1$ for all $x \in [0, 1/2)$. To see the former, write

$$\int_{1/2}^{1} |g(x)| dx = \int_{1/2}^{1} |f_n(x) - g(x)| dx \le \int_0^1 |f_n(x) - g(x)| dx = \|f_n - g\|_{L^1([0,1])} = d_1(f_n, g).$$

As $d_1(f_n, g) \to 0$ as $n \to +\infty$, it follows that

$$\int_{1/2}^{1} |g(x)| dx = 0.$$

As $g(x)$ is continuous on $[0,1]$, it follows that $g(x) = 0$ for all $x \ge 1/2$. Similarly, if we fix any $a \in [0, 1/2)$ and take n sufficient;y large so that $a < 1/2 - 1/n$, then we have

$$\int_0^a |g(x) - 1| dx = \int_0^a |g(x) - f_n(x)| dx \le \int_0^1 |f_n(x) - g(x)| dx = \|f_n - g\|_{L^1([0,1])} = d_1(f_n, g).$$

For the same reason, it follows that $g(x) = 1$ for all $x \in [0, a]$. As $a < 1/2$ is arbitrary, it follows that $g(x) = 1$ for all $x \in [0, 1/2)$. However, there is no continuous function $g(x)$ such that $g(x) = 0$ for all $x \ge 1/2$ and $g(x) = 1$ for all $x \in [0, 1/2)$. It follows that the sequence f_n can not have any limit in $\tilde{L}^1([0,1])$.

1.3 Uniform convergence of functions

Let us make a slight digression on the uniform convergence of functions. Let (X, d_X) and (Y, d_Y) be two metric spaces.

Definition 1.10 *A sequence of functions $f_n : X \to Y$ converges to a function $f : X \to Y$ uniformly on X if for any $\varepsilon > 0$ there exists N so that $d_Y(f(x), f_n(x)) < \varepsilon$ for all $n \ge N$ and all $x \in X$. We sometimes use the notation $f_n \rightrightarrows f$ on X.*

The sequence of constant functions $f_n(x) \equiv 1/n$ converges uniformly to $f = 0$ on $\mathbb{R}$, or on any other set $E \subset \mathbb{R}$. On the other hand, the sequence $f_n(x) = x/n$ converges uniformly on $[0,1]$ to $f \equiv 0$ but is not uniformly convergent on $\mathbb{R}$. Thus the notion of uniform convergence very much depends on the set E on which we consider the convergence.

Definition 1.11 *A sequence of functions $f_n : X \to Y$ is uniformly Cauchy on X if for any $\varepsilon > 0$ there exists N so that $d_Y(f_m(x), f_n(x)) < \varepsilon$ for all $n, m \geq N$ and all $x \in X$.*

Here is a useful criterion for the uniform convergence for functions into $\mathbb{R}$.

Theorem 1.12 *A sequence of functions $f_n : X \to \mathbb{R}$ is uniformly convergent to $f : X \to \mathbb{R}$ if and only if f_n is uniformly Cauchy on X.*

Proof. First, let $f_n \rightrightarrows f$ on X. Then for any $\varepsilon > 0$ there exists N so that $|f(x) - f_n(x)| < \varepsilon/10$ for all $n \geq n$ and all $x \in X$. The triangle inequality implies that for any $n, m \geq N$ and all $x \in X$ we have

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \leq \frac{\varepsilon}{10} + \frac{\varepsilon}{10} < \varepsilon,$$

hence f_n is a uniformly Cauchy sequence on X.

Next, assume that f_n is a uniformly Cauchy sequence on X. Then, for each $x \in X$ fixed, the sequence of numbers $f_n(x)$ is Cauchy, thus it converges to some limit that we denote by $f(x)$. As the sequence f_n is uniformly Cauchy on X, given $\varepsilon > 0$ there exists N so that for all $n, m \geq n$ and all $x \in X$ we have

$$|f_n(x) - f_m(x)| < \frac{\varepsilon}{10}.$$

Fixing $m > N$ and letting $n \to \infty$ above, we deduce that

$$|f(x) - f_m(x)| < \frac{\varepsilon}{10}, \text{ for all } m \geq n \text{ and all } x \in X,$$

which means that $f_n \rightrightarrows f$ on X. □

This result can be generalized to functions taking values not in $\mathbb{R}$ but in a complete metric space.

Theorem 1.13 *Let (X, d_X) and (Y, d_Y) be metric spaces and assume that (Y, d_Y) is a complete metric space. Then, a sequence of functions $f_n : X \to Y$ is uniformly convergent to $f : X \to Y$ if and only if f_n is uniformly Cauchy on X.*

Exercise 1.14 Modify the proof of Theorem 1.12 to prove Theorem 1.13.

Let us recall the definition of a continuous map between two metric spaces.

Definition 1.15 *Suppose (X, d_X), (Y, d_Y) are metric spaces. A function $f : X \to Y$ is continuous at a point $a \in X$ if for all $\varepsilon > 0$ there exists $\delta > 0$ such for all $x \in X$ such that $d_X(x, a) < \delta$ we have $d_Y(f(x), f(a)) < \varepsilon$. A function is called continuous on X if it is continuous at all $a \in X$.*

Uniform convergence preserves continuity.

Theorem 1.16 *Let (X, d_X), (Y, d_Y) be two metric spaces, and $f_n : X \to Y$ be continuous functions. Assume that f_n converge uniformly on X to $f : X \to Y$, then f is continuous.*

Proof. As f_n converge uniformly to f on E, given $\varepsilon > 0$,we can find N so that

$$d_Y(f_n(x), f(x)) < \frac{\varepsilon}{10}$$

for all $n \geq N$ and $x \in E$. Given $x \in E$, as the function $f_N(x)$ is continuous at x, we can find $\delta > 0$ so that

$$d_Y(f_N(x), f_N(y)) < \frac{\varepsilon}{10},$$

if $d_X(x, y) < \delta$. It follows that for all $y \in E$ such that $d(x, y) < \delta$ we have

$$d_Y(f(x), f(y)) < d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(y)) + d_Y(f_N(y), f(y)) < \frac{\varepsilon}{10} + \frac{\varepsilon}{10} + \frac{\varepsilon}{10} < \varepsilon.$$

We used the uniform convergence of f_n to f in the first and the last terms above, and continuity of f_N at x in the second term. Now, continuity of $f(x)$ at the point x follows. □

1.4 Complete metric spaces

In order to distinguish the spaces in which Cauchy sequences converge, we make the following definition: a metric space X is complete if any Cauchy sequence in X converges. Theorem 1.6 shows that the spaces $\mathbb{R}^n$ are complete. Let us show that the space $C[0,1]$ with the norm

$$\|f\| = \sup_{0\le x\le 1} |f(x)|$$

is a complete metric space.

Theorem 1.17 *The metric space $C[0,1]$ is complete.*

Proof. Let f_n be a Cauchy sequence in $C[0,1]$. This means that for any $\varepsilon > 0$ there exists N so that for any $n, m \ge N$ we have

$$\|f_n - f_m\| < \varepsilon.$$

In other words, we have

$$\sup_{0\le x\le 1} |f_n(x) - f_m(x)| < \varepsilon,$$

so that

$$|f_n(x) - f_m(x)| < \varepsilon,$$

for all $n, m \ge N$ and all $x \in [0,1]$. This means that the sequence f_n is uniformly Cauchy on $[0,1]$. Recall that Theorem 1.12 implies that then f_n is uniformly convergent, and by Theorem 1.16 the limit is a continuous function. Thus, the space $C[0,1]$ is a complete metric space. □

More generally, given a metric space X and a compact subset $K \subset X$, and a normed space Y we can consider the space $C(K;Y)$ of continuous functions on K with values in Y with the norm

$$\|f\|_{C(K;Y)} = \sup_{x\in K} \|f(x)\|_Y.$$

A normed space Y is complete if the corresponding metric space with the metric $d(x,y) = \|x-y\|_Y$ is complete.

Theorem 1.18 *Assume that K is a compact subset of a metric space (X, d_X) and that Y is a complete normed space. Then, the metric space $C(K;Y)$ is complete.*

Exercise 1.19 Adapt the proof of Theorem 1.18) to prove Theorem 1.18.

2 The contraction mapping principle

The contraction mapping principle is the most basic tool that can be used to prove existence of solutions to all kinds of equations in many situations in analysis. Many problems can be formulated in the form

$$f(x) = y_0, \tag{2.1}$$

where y_0 is an element of some metric space X, f is a mapping from X to X, and x is the unknown that we need to find. We can reformulate it as

$$f(x) + x - y_0 = x.$$

The advantage of the latter formulation is that now we have what is known as a fixed point problem. These are equations of the form

$$F(x) = x, \tag{2.2}$$

where F is a mapping from a metric space X to itself, and x is an unknown point $x \in X$. A solution to (2.2) is known as a fixed point of the mapping F. In other words, (2.1) is equivalent to (2.2) with

$$F(x) = f(x) + x - y_0. \tag{2.3}$$

Of course, there is a serious difference: to pass from (2.1) to (2.2) we need to introduce the mapping F in (2.3). This requires an addition structure on X: we need X to be a vector space to be able to do that. However, often X is a vector space, so that issue is not a problem.

An important class of mappings of a metric space X onto itself are contractions. We say that a mapping $f: X \to X$ is a contraction if there exists a number $q \in (0,1)$ so that for an $x_1, x_2 \in X$ we have

$$d(f(x_1), f(x_2)) \leq q d(x_1, x_2). \tag{2.4}$$

The contraction mapping principle says the following.

Theorem 2.1 *Let X be a complete metric space, and $f : X \to X$ be a contraction, then f has a unique fixed point a in X. Moreover, for any $x_0 \in X$, the sequence defined recursively by $x_{n+1} = f(x_n)$, with $x_1 = f(x_0)$, converges to a as $k \to +\infty$. The rate of convergence can be estimated by*

$$d(x_n, a) \leq \frac{q^n}{1-q} d(x_1, x_0). \tag{2.5}$$

Note that the theorem provides an algorithm to compute the unique fixed point, and that the rate of convergence in (2.5) depends on how close q is to 0 or 1: it gets faster for q close to 0 and slower for q close to 1.

Proof. We will show that the sequence x_k is a Cauchy sequence. Note that

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq q d(x_n, x_{n-1}), \tag{2.6}$$

so that an induction argument shows that

$$d(x_{n+1}, x_n) \leq q^n d(x_1, x_0). \tag{2.7}$$

Now, by the triangle inequality and (2.7), we have

$$\begin{aligned} d(x_{n+k}, x_n) &\leq d(x_{n+k}, x_{n+k-1}) + d(x_{n+k-1}, x_{n+k-2}) + \cdots + d(x_{n+1}, x_n) \\ &\leq (q^{n+k-1} + q^{n+k-2} + \cdots + q^n) d(x_1, x_0) \leq \frac{q^n}{1-q} d(x_1, x_0). \end{aligned} \tag{2.8}$$

It follows that if $0 < q < 1$, then the sequence x_n is Cauchy. Since the space X is complete, the limit of x_n exists, and we set

$$a = \lim_{n \to \infty} x_n.$$

Now, as f is a continuous map, passing to the limit $n \to \infty$ in the recursion relation $x_n = f(x_{n-1})$, we arrive at $a = f(a)$, hence a is a fixed point of f.

The reason the fixed point is unique is that f is a contraction. Indeed, if a_1 and a_2 are two fixed points, so that $a_1 = f(a_1)$ and $a_2 = f(a_2)$, then by the contraction property we have

$$d(f(a_1), f(a_2)) \leq q d(a_1, a_2).$$

However, as both a_1 and a_2 are fixed points of f, the left side above equals $d(a_1, a_2)$. Since $q \in (0,1)$, we deduce that $d(a_1, a_2) = 0$ and $a_1 = a_2$. □

Existence theorem for ordinary differential equations

Let us consider a system of ordinary differential equations (ODE) for an unknown vector-valued function $y(x)$ with values in $\mathbb{R}^n$:

$$y' = f(x, y) \tag{2.9}$$

supplemented by the initial condition

$$y(x_0) = y_0, \tag{2.10}$$

with some given $x_0 \in \mathbb{R}$ and $y_0 \in \mathbb{R}^n$. We assume that the vector-valued function $f(x, y) \in \mathbb{R}^n$ is continuous in $(x, y) \in \mathbb{R}^{n+1}$ and Lipschitz in y: there exists a constant $M > 0$ so that

$$\|f(x, y_1) - f(x, y_2)\|_{\mathbb{R}^n} \leq M\|y_1 - y_2\|_{\mathbb{R}^n} \text{ for all } x \in \mathbb{R} \text{ and } y_1, y_2 \in \mathbb{R}^n. \tag{2.11}$$

Theorem 2.2 *Under the above assumptions on $f(x, y)$, the problem (2.9)-(2.10) has a unique continuous solution $y(x)$ on the interval $(x_0 - \delta_0, x_0 + \delta_0)$ with $\delta_0 = 1/(2M)$. This solution is also continuously differentiable.*

An immediate corollary of this theorem is that under the above assumptions a unique solution to (2.9)-(2.10) exists for all $x \in \mathbb{R}$ and not just in a small interval around x_0.

Corollary 2.3 *Under the above assumptions on $f(x, y)$, the problem (2.9)-(2.10) has a unique continuously differentiable solution $y(x)$ defined for all $x \in \mathbb{R}$.*

Proof of Corollary 2.3. Note that δ_0 in Theorem 2.2 does not depend on x_0 or y_0. Therefore, after we construct the solution on the interval $[x_0 - \delta_0, x_0 + \delta_0]$, we may consider (2.9) with the initial condition $y(x_1) = y_1$, with $x_1 = x_0 + \delta_0$ and $y_1 = y(x_0 + \delta_0)$, obtained from solving our problem on $[x_0 - \delta_0, x_0 + \delta_0]$ during the first step. As δ_0 does not depend on the initial conditions (x_1, y_1), we may apply Theorem 2.2 again to this new initial value problem, and extend the solution to the interval $[x_1, x_1 + \delta_0] = [x_0 + \delta_0, x_0 + 2\delta_0]$, with the same interval length δ_0. Continuing this process, we will define the solution for all $x \geq x_0$. Similarly, we can proceed to define it for all $x < x_0$. □

Exercise 2.4 Fill in the gaps on how you proceed for $x < x_0$ in the above proof.

Proof of Theorem 2.2. If $y(x)$ is a continuous (and hence continuously differentiable) solution to (2.9)-(2.10), we can re-write (2.9)-(2.10), using the fundamental theorem of calculus as

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt. \tag{2.12}$$

Conversely, if $y(x)$ is a continuos solution to (2.12) then it is continuously differentiable and satisfies (2.9)-(2.10).

Exercise 2.5 Check this and, in particular, justify why a continuous solution $y(x)$ to (2.12) is not just continuous but also continuously differentiable.

Let us define the map A that maps a continuous function $y(x)$ to a function $A[y]$ via

$$(A[y])(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt. \tag{2.13}$$

Then (2.12) can be written as

$$y(x) = A[y](x), \tag{2.14}$$

so that $y(x)$ is a solution to (2.12), or, equivalently, to (2.9)-(2.10) if and only if the function y is a fixed point of the mapping A. The first step is to show that A maps the space $C([x_0 - \delta, x_0 + \delta]; \mathbb{R}^n)$ to itself, for any $\delta > 0$.

Lemma 2.6 *If $y \in C[x_0-\delta, x_0+\delta]$ for some $\delta > 0$, then $A[y]$ is also in $C([x_0-\delta, x_0+\delta];\mathbb{R}^n)$.*

Proof of Lemma. Let $y(x)$ be an $\mathbb{R}^n$-valued continuous function on the closed interval $[x_0-\delta, x_0+\delta]$ for some given $\delta > 0$. Then for any $x_1, x_2 \in [x_0-\delta, x_0+\delta]$, we have

$$A[y](x_1) - A[y](x_2) = y_0 + \int_{x_0}^{x_1} f(t, y(t))dt - y_0 - \int_{x_0}^{x_2} f(t, y(t))dt = \int_{x_2}^{x_1} f(t, y(t))dt. \quad (2.15)$$

The function y is continuous on $[x_0-\delta, x_0+\delta]$, hence it is bounded on that interval: there exists K such that $\|y(t)\|_{\mathbb{R}^n}| \le K$ for all $x_0-\delta \le t \le x_0+\delta$. As the function f is continuous on $\mathbb{R}^{n+1}$, there exists $L > 0$ so that $\|f(x,y)\|_{\mathbb{R}^n} \le L$ for all $x \in [x_0-\delta, x_0+\delta]$ and $\|y\|_{\mathbb{R}^n} \le K$. It follows that $\|f(t, y(t))\|_{\mathbb{R}^n} \le L$ for all $t \in [x_0-\delta, x_0+\delta]$. Using this in (2.15) gives

$$\|A[y](x_1) - A[y](x_2)\|_{\mathbb{R}^n} \le \int_{x_2}^{x_1} \|f(t, y(t))\|_{\mathbb{R}^n} dt \le L|x_1 - x_2|. \quad (2.16)$$

It follows that the function $A[y]$ is continuous. □

We return to the proof of the theorem. Our goal is to show that if δ is sufficiently small, then A is a contraction on $C([x_0-\delta, x_0+\delta];\mathbb{R}^n)$. Let us take two functions $y_1, y_2 \in C([x_0-\delta, x_0+\delta];\mathbb{R}^n)$ and write

$$A[y_1](x) - A[y_2](x) = y_0 + \int_{x_0}^{x} f(t, y_1(t))dt - y_0 - \int_{x_0}^{x} f(t, y_2(t))dt = \int_{x_0}^{x} [f(t, y_1(t)) - f(t, y_2(t))]dt. \quad (2.17)$$

We will now use the Lipschitz property (2.11) of the function $f(x,y)$:

$$\|A[y_1](x) - A[y_2](x)\|_{\mathbb{R}^n} \le \int_{x_0}^{x} \|f(t, y_1(t)) - f(t, y_2(t))\|_{\mathbb{R}^n} dt \le \int_{x_0}^{x} M\|y_1(t) - y_2(t)\|_{\mathbb{R}^n} dt. \quad (2.18)$$

Note that for all $t \in [x_0, x]$ we have

$$\|y_1(t) - y_2(t)\|_{\mathbb{R}^n} \le \sup_{x_0-\delta \le x \le x_0+\delta} \|y_1(x) - y_2(x)\| = \|y_1 - y_2\|_C. \quad (2.19)$$

Here, we use an abbreviation

$$|y_1 - y_2\|_C := \|y_1 - y_2\|_{C([x_0-\delta, x_0+\delta];\mathbb{R}^n)}.$$

Using (2.19) in (2.18), we arrive at

$$\begin{aligned} \|A[y_1](x) - A[y_2](x)\|_{\mathbb{R}^n} &\le \int_{x_0}^{x} M\|y_1(t) - y_2(t)\|_{|Rm^n} dt \le \int_{x_0}^{x} M\|y_1 - y_2\||_C dt \\ &= M|x - x_0|\|y_1 - y_2\|_C \le M\delta\|y_1 - y_2\|_C. \end{aligned} \quad (2.20)$$

Taking supremum over all $x \in [x_0-\delta, x_0+\delta]$ gives

$$\|A[y_1] - A[y_2]\|_C = \sup_{x \in [x_0-\delta, x_0+\delta]} \|A[y_1](x) - A[y_2](x)\|_{\mathbb{R}^n} \le M\delta\|y_1 - y_2\|_C. \quad (2.21)$$

Therefore, if $M\delta < 1$ then A is a contraction on $C([x_0-\delta, x_0+\delta];\mathbb{R}^n)$. In particular, we may take $\delta = \delta_0 = 1/(2M)$. It follows that A has a unique fixed point y in $C([x_0-\delta_0, x_0+\delta_0];\mathbb{R}^n)$. This means that the function $y(x)$ satisfies (2.12):

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt. \quad (2.22)$$

It follows immediately that $y(x_0) = y_0$. Moreover, as the function $y(t)$ is continuous, and $f(t, y)$ is continuous in both variables, it follows that $p(t) = f(t, y(t))$ is continuous in t. The fundamental theorem of calculus implies then that $y(x)$ is differentiable and

$$y'(x) = f(x, y(x)). \tag{2.23}$$

This finishes the proof. □

Let us modify the above proof to functions $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ that are not uniformly Lipschitz as in (2.11) but rather locally Lipschitz. This means the following: for any compact set $K \subset \mathbb{R}^{n+1}$ there exists a constant M_K so that

$$\|f(x, y_1) - f(x, y_2)\|_{\mathbb{R}^n} \le M\|y_1 - y_2\|_{\mathbb{R}^n} \text{ for all } (x, y_1) \in K \text{ and } (x, y_2) \in K. \tag{2.24}$$

A standard example of a locally Lipschitz but not uniformly Lipschitz function on $\mathbb{R}$ is $f(y) = y^2$, or any other polynomial $p(x)$, or the exponential function $f(y) = e^y$. We have the following result for this case. Consider the initial value problem

$$y' = f(x, y) \tag{2.25}$$

supplemented by the initial condition

$$y(x_0) = y_0, \tag{2.26}$$

with some given $x_0 \in \mathbb{R}$ and $y_0 \in \mathbb{R}^n$.

Theorem 2.7 *Assume that the vector-valued function $f(x, y) \in \mathbb{R}^n$ is continuous in $(x, y) \in \mathbb{R}^{n+1}$ and locally Lipschitz in y. Then, there exists $\delta_0 > 0$ that may depend both on x_0 and y_0 so that the problem (2.9)-(2.10) has a unique continuous solution $y(x)$ on the interval $(x_0 - \delta_0, x_0 + \delta_0)$. This solution is also continuously differentiable.*

Proof. The proof is very similar to that of Theorem 2.2 but has important modifications. As before, $y(x)$ is a continuous (and hence continuously differentiable) solution to (2.25)-(2.26) if and only if it satisfies the integral equation

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt. \tag{2.27}$$

Also as in the previous proof, we define the map $\mathcal{A}$ that maps a continuous function $y(x)$ to a function $\mathcal{A}[y]$ as in (2.13):

$$(\mathcal{A}[y])(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt, \tag{2.28}$$

so that (2.27) can be written as

$$y(x) = \mathcal{A}[y](x). \tag{2.29}$$

Thus, $y(x)$ is a solution to (2.27), or, equivalently, to (2.25)-(2.26) if and only if the function y is a fixed point of the mapping $\mathcal{A}$. Exactly as in the proof of Lemma 2.6, we can show that for any $\delta > 0$, the map $\mathcal{A}$ maps the space $C([x_0 - \delta, x_0 + \delta]; \mathbb{R}^n)$ to itself. The proof is verbatim the same.

The next step in the proof is slightly different from that in the proof of Theorem 2.2. Consider the sets

$$B = \{y \in \mathbb{R}^n : \ \|y - y_0\|_{\mathbb{R}^n} \le 1\} \subset \mathbb{R}^n,$$

and

$$K = \{(x, y) \in \mathbb{R}^{n+1} : \ |x - x_0| \le 1, \ y \in B\} \subset \mathbb{R}^{n+1}.$$

As the set K is compact and $f(x,y)$ is a continuous function, there exists a constant C_K so that

$$\|f(x,y)\|_{\mathbb{R}^n} \le C_K, \text{ for all } (x,y) \in K. \tag{2.30}$$

In addition, as $f(x,y)$ is locally Lipschitz, and K is compact, there exists a constant M_K so that

$$\|f(x,y_1) - f(x,y_2)\|_{\mathbb{R}^n} \le M_K, \text{ for all } (x,y_1) \in K \text{ and } (x,y_2) \in K. \tag{2.31}$$

The main difference with the previous proof is in how the contraction mapping principle is applied. Take $\delta > 0$ and let X be the metric space of all continuous $\mathbb{R}^n$-valued functions $z(x)$ defined on an interval $[x_0-\delta, x_0+\delta]$ such that $z(x) \in B$ for all $x \in [x_0-\delta, x_0+\delta]$, with the distance

$$d(z_1,z_2) = \sup_{x\in[x_0-\delta,x_0+\delta]} \|z_1(x) - z_2(x)\|_{\mathbb{R}^n}. \tag{2.32}$$

Exercise 2.8 Show that the metric space X is complete. Be careful to check that if z_n is a Cauchy sequence in X that converges in the uniform norm to a limit z, then $z(x)$ is continuous and also takes values in B for all $x \in [x_0-\delta, x_0+\delta]$, and thus lies in X.

The key observation is the following lemma.

Lemma 2.9 *There exists $\delta_0 > 0$ so that $\mathcal{A}$ is a contraction on X for all $\delta \in (0,\delta_0)$.*

Proof. We need to check two things: first, that $\mathcal{A}$ maps X onto itself, and, second, that it is a contraction. For the former, given a function $z \in X$ we already know that $\mathcal{A}[z]$ is a continuous function, so we only need to check that $\mathcal{A}[z](x)$ belongs to B for all $x \in [x_0-\delta, x_0+\delta]$ if δ is sufficiently small. To this end, note that if $\delta < 1$, then we have

$$\|\mathcal{A}[z](x) - y_0\|_{\mathbb{R}^n} = \Big\| \int_{x_0}^{x} f(t,z(t))dt \Big\|_{\mathbb{R}^n} \le \int_{x_0}^{x} \|f(t,z(t))\|_{\mathbb{R}^n} dt \le C_K|x-x_0|. \tag{2.33}$$

The last inequality above follows from the fact that $z \in X$, so that $f(t,z(t)) \in B$ for all t in the domain of integration, together with the definition of C_K in (2.30). Therefore, if we choose $\delta \in (0,1)$ so that

$$C_K\delta \le 1, \tag{2.34}$$

then we have, for all $x \in [x_0-\delta, x_0+\delta]$:

$$\|\mathcal{A}[z](x) - y_0\|_{\mathbb{R}^n} \le C_K|x-x_0| \le C_K\delta \le 1, \tag{2.35}$$

and thus $\mathcal{A}[z](x)$ lies in the set B for all $x \in [x_0-\delta, x_0+\delta]$. It follows that $\mathcal{A}$ maps X into itself.

The next step is to check that $\mathcal{A}$ is a contraction on X if δ is sufficiently small. Let us take two functions $z_1, z_2 \in X$ and write

$$\mathcal{A}[z_1](x) - \mathcal{A}[z_2](x) = y_0 + \int_{x_0}^{x} f(t,z_1(t))dt - y_0 - \int_{x_0}^{x} f(t,z_2(t))dt = \int_{x_0}^{x} [f(t,z_1(t)) - f(t,z_2(t))]dt. \tag{2.36}$$

As the functions z_1 and z_2 are in X, the points $(t,z_1(t))$ and $(t,z_2(t))$ lie in K for all t in the domain of integration above. hence, we may use the local Lipschitz constant M_K defined in (2.31):

$$\|\mathcal{A}[z_1](x) - \mathcal{A}[z_2](x)\|_{\mathbb{R}^n} \le \int_{x_0}^{x} \|f(t,z_1(t)) - f(t,z_2(t))\|_{\mathbb{R}^n} dt \le \int_{x_0}^{x} M_K\|z_1(t) - z_2(t)\|_{\mathbb{R}^n} dt. \tag{2.37}$$

Note that for all $t \in [x_0, x]$ we have

$$\|z_1(t) - z_2(t)\|_{\mathbb{R}^n} \le \sup_{x_0-\delta\le x\le x_0+\delta} \|z_1(x) - z_2(x)\| = d_X(z_1, z_2). \tag{2.38}$$

Using (2.38) in (2.37), we arrive at

$$\|\mathcal{A}[z_1](x) - \mathcal{A}[z_2](x)\|_{\mathbb{R}^n} \le M_K|x - x_0|d_X(z_1, z_2) \le M_K\delta d_X(z_1, z_2). \tag{2.39}$$

Taking supremum over all $x \in [x_0 - \delta, x_0 + \delta]$ gives

$$d_X(\mathcal{A}[z_1], \mathcal{A}[z_2]) = \sup_{x\in[x_0-\delta,x_0+\delta]} \|A[z_1](x) - A[z_2](x)\|_{\mathbb{R}^n} \le \delta M_K d_X(z_1, z_2). \tag{2.40}$$

Therefore, if $\delta M_K < 1$ then A is a contraction on X. Keeping in mind that we also need (2.34) to hold, to ensure that $\mathcal{A}$ maps X into itself, and that we also need $\delta < 1$ because of the way the set K is defined, we see that if we take

$$\delta = \frac{1}{2}\min\Big(\frac{1}{M_K}, \frac{1}{C_K}, 1\Big), \tag{2.41}$$

then $\mathcal{A}$ is a contraction on X. Thsi finishes the proof of Lemma 2.9. $\Box$

Together with the result of Exercise 2.8, Lemma 2.9 implies that $\mathcal{A}$ has a unique fixed point y in X. This means that the function $y(x)$ satisfies (2.27):

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt. \tag{2.42}$$

It follows immediately that $y(x_0) = y_0$. Moreover, as the function $y(t)$ is continuous, and $f(t, y)$ is continuous in both variables, it follows that $p(t) = f(t, y(t))$ is continuous in t. The fundamental theorem of calculus implies then that $y(x)$ is differentiable and

$$y'(x) = f(x, y(x)). \tag{2.43}$$

This finishes the proof. $\Box$

Exercise 2.10 We have actually only proved that solution of (2.25)-(2.26) is unique in X, that is, among functions that take values in B. Exaplain why uniqueness in $C[x_0 - \delta, x_0 + \delta]; \mathbb{R}^n)$ follows.

Some examples

Let us now combine the existence theorem for ODEs with the construction of the fixed point of a contraction mapping in the proof of the existence theorem of a fixed point, to show how a solution can be constructed iteratively. Consider an ODE

$$y'(t) = y(t), \quad y(0) = 1,$$

and write it, as in (2.12), in the form

$$y(t) = 1 + \int_0^t y(s)ds.$$

The mapping A is now defined via

$$A[y](t) = 1 + \int_0^t y(s)ds.$$

Consider the recursive sequence $y_n(t)$, with $y_0(t) = 1$, and

$$y_{n+1}(t) = 1 + \int_0^t y_n(s)ds.$$

Exercise 2.11 Show by induction that

$$y_n(t) = \sum_{k=1}^{n} \frac{t^k}{k!}.$$

We see that the unique solution is $y(t) = e^t$.

Example of non-uniqueness. The assumption that the function $f(x, y)$ is Lipschitz in y is crucial in Theorem 2.2 to ensure that solution of the initial value problem

$$y'(x) = f(x, y), \quad y(x_0) = y_0,$$

is unique. Without it uniqueness may simply fail. Indeed, consider an ordinary differential equation

$$y' = \sqrt{y},$$

with the initial condition $y(0) = 0$. It has one obvious solution $y_1(x) = 0$ for all $x \in \mathbb{R}$ but also another one: $y_2(x) = x^2/4$.

Example of a blow-up. Another natural question is whether solution to an initial value problem

$$y'(x) = f(x, y), \quad y(x_0) = y_0,$$

always exists for all $x \in \mathbb{R}$ if $f(x, y)$ is not uniformly Lipschitz in y. The answer is not necessarily: consider the initial value problem

$$y' = y^2, \tag{2.44}$$

with the initial condition $y(0) = 1$. Its solution is

$$y(x) = \frac{1}{1-x},$$

and it can not be extended past $x = 1$. On the other hand, if the initial condition for (2.44) is negative: $y(0) = y_0 < 0$ then we claim that solution to (2.44) such that $y(0) = y_0$ exists for all $x \geq 0$. Indeed, (2.44) has a special solution $y_1(x) = 0$ for all $x \geq 0$. Moreover, the function $f(y) = y^2$ is locally Lipschitz. Hence, solution to (2.44) with an initial condition $y(x_0') = y_0'$ is unique in a neighborhood of x_0', for all x_0' and y_0'. Going back to the solution to (2.44) with $y(0) = y_0 < 0$, we conclude that if there exists $x_0 \in \mathbb{R}$ such that $y(x_0) = 0$ then uniqueness would imply that $y(x) = 0$ for all $x \in \mathbb{R}$ (please think through why we can say that for all $x \in \mathbb{R}$ and not just in a neighborhood of x_0). This would contradict the initial condition $y(0) = y_0 < 0$. Hence, $y(x)$ is an increasing function such that $y_0 \leq y(x) < 0$ for all $x \geq 0$, as long as the solution exists. However, the function $f(y) = y^2$ is Lipschitz on $[y_0, 0]$ with some Lipschitz constant M. From that, existence of the solution for all $x \geq 0$ follows (please think through why this claim is true).

We will revisit all these issues in much greater detail later in the course.

3 The implicit function theorem

The implicit function theorem addresses the question of when an equation of the form $f(x, y) = 0$ uniquely defines x in terms of y, or y in terms of x. The former question means that given y, we are trying to "solve the equation $f(x, y) = 0$ for an unknown x". The latter asks when an equation of the form $f(x, y) = 0$ uniquely defines a function $y(x)$ – hence, the name the implicit function theorem. The two questions are totally equivalent but the points of view are slightly different.

Let us consider a very simple example: the equation

$$x^2 - y = 0. \tag{3.1}$$

Then, for $y > 0$ this equation has two solutions $x = \pm\sqrt{y}$, for $y = 0$ it has one solution $x = 0$, and for $y < 0$ it has no real solutions. Let, us change our perspective somewhat. Suppose we know a particular solution (x_0, y_0) – that is, $x_0^2 = y_0$ and we ask: given a y close to y_0, can we find a unique x close to x_0 so that $x^2 = y$? That is, if we perturb y_0 slightly, can we still find a unique solution of (3.1) close to the original solution x_0? In other words, is there a map that sends y to x that is stable under small perturbations? Note that we have $y_0 \geq 0$ automatically, simply because $x_0^2 = y_0$. The answer to our question is that if $y_0 > 0$, and, say, $x_0 > 0$ then, indeed, for y close to y_0 we still have a solution to (3.1) that is close to x_0: $x = \sqrt{y}$. Similarly, if we have $x_0 < 0$, then we still have a solution to (3.1) that is close to x_0: $x = -\sqrt{y}$. Note that existence of "the other solution": $x = -\sqrt{y}$ in the former case and $x = \sqrt{y}$ in the latter is irrelevant: we are only asking that the inversion is unique locally, not globally. On the other hand, if $x_0 = 0$ so that $y_0 = 0$, then in any interval $y \in (-\delta, \delta)$ around $y_0 = 0$ and any interval $(-\delta', \delta')$ around $x_0 = 0$ we can find $y < 0$ for which the equation $x^2 = y$ has no solutions and $y > 0$ for which $x^2 = y$ has two solutions in the interval $(-\delta', \delta')$ – we just need to take $y < (\delta')^2$. Thus, there is a qualitative difference between $x_0 = 0$, $y_0 = 0$ and other points on the graph of $y = x^2$ – we can locally invert the relationship around the latter but not the former. The implicit function theorem generalizes this trivial observation.

3.1 The inverse function theorem on $\mathbb{R}$

We begin with the inverse function theorem, that looks not at an "implicit" equation $f(x, y) = 0$ but at the simpler problem of solving an equation of the form $f(x) = y$. In one dimension the situation is quite simple.

Proposition 3.1 *Let $f(x)$ be continuously differentiable on an interval $[a, b]$, and set*

$$m = \inf_{a \leq x \leq b} f(x), \quad M = \sup_{a \leq x \leq b} f(x).$$

(i) Then f is a one-to-one map from $[a, b]$ to $[m, M]$ if and only if f is monotonic on $[a, b]$. (ii) In addition, if f is monotonic on $[a, b]$ then the inverse function $g = f^{-1} : [m, M] \to [a, b]$ is continuously differentiable at $y_0 \in [m, M]$ if and only if $f'(g(y_0)) \neq 0$. In that case, $g'(y_0) = 1/f'(g(y_0))$.

Exercise 3.2 Prove the first statement (i) in the above proposition.

To prove (ii), assume first that $x_0 \in (a, b)$ and $f'(x_0) \neq 0$. Without loss of generality, we may assume that $f'(x_0) > 0$. As the function f' is continuous at x_0, there exists $\delta > 0$ so that

$$f'(x) > f'(x_0)/2 \text{ for all } x \in (x_0 - \delta, x_0 + \delta), \tag{3.2}$$

thus f is strictly increasing on that interval. Let us set $\alpha = f(x_0 - \delta)$, $\beta = f(x_0 + \delta)$ and take some $y \in (\alpha, \beta)$, so that $y = f(x)$ for some $x \in (x_0 - \delta, x_0 + \delta)$, that is, $x = g(y)$ – recall that $g = f^{-1}$. Our goal is to show that the function $g(y)$ is differentiable at y_0 and $g'(y_0) = 1/f'(g(y_0))$. As f is differentiable on $(x_0 - \delta, x_0 + \delta)$, there exists c between x and x_0 so that we have

$$f(x) - f(x_0) = f'(c)(x - x_0), \tag{3.3}$$

and, because of (3.2), we know that $f'(c) > f'(x_0)/2 > 0$. Take now some $y \in (f(x_0 - \delta), f(x_0 + \delta))$, and use (3.3) with $x = g(y) \in (x_0 - \delta, x_0 + \delta)$. This gives

$$y - y_0 = f'(c)(g(y) - g(y_0)), \tag{3.4}$$

so that

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{1}{f'(c)}, \tag{3.5}$$

with c between $x = g(y)$ and $x_0 = g(y_0)$. We now let $y \to y_0$ in (3.5). As c lies between $g(y)$ and $g(y_0)$, and the function $g(y)$ is continuous, we know that $c \to g(y_0)$ as $y \to y_0$. Hence, the limit of the right side of (3.5) exists, and thus so does the limit of the left side, and

$$\lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \frac{1}{f'(x_0)}, \tag{3.6}$$

which means exactly that $g'(y_0) = 1/f'(g(y_0))$, as claimed. □

3.2 The inverse function theorem for maps $\mathbb{R}^n \to \mathbb{R}^n$: an outline

Before we proceed with the inverse function theorem for maps from $\mathbb{R}^n$ to $\mathbb{R}^n$, that is, $n \times n$ systems of equations

$$\begin{aligned} &F_1(x_1, \dots, x_n) = y_1, \\ &\dots\dots \\ &F_n(x_1, \dots, x_n) = y_n, \end{aligned} \tag{3.7}$$

let us first consider a special case when $F : \mathbb{R}^n \to \mathbb{R}^n$ is an affine map, that is, there exists an $n \times n$ matrix A and a vector $q \in \mathbb{R}^n$, so that

$$F(x) = Ax + q. \tag{3.8}$$

In that case, the equation $F(x) = y$ takes the form

$$Ax + q = y, \tag{3.9}$$

and has an explicit solution

$$x = A^{-1}(y - q),$$

provided that the matrix A is invertible.

A general differentiable map F can be approximated in a neighborhood of a point x_0 as

$$F(x) \approx F(x_0) + [DF(x_0)](x - x_0), \tag{3.10}$$

in the sense that

$$F(x) - (F(x_0) + [DF(x_0)](x - x_0)) = o(\|x - x_0\|). \tag{3.11}$$

Here, $DF(x_0)$ is the derivative matrix of F at x_0. Now, for y close to $y_0 = F(x_0)$, let us replace the exact equation

$$F(x) = y \tag{3.12}$$

by an approximate equation

$$F(x_0) + [DF(x_0)](x - x_0) = y. \tag{3.13}$$

Warning: at the moment, we do not really know that solutions of (3.12) and (3.13) are close, we are simply trying to understand informally what should be important for (3.12) to have a solution x

close to x_0. The question you may want to keep in the back of your mind is why the solution to an approximate equation (3.13) is close to a solution of the true equation (3.12) – this should become clear from the proof of Theorem 3.7 below. Note that the approximate equation (3.13) has the familiar form (3.9), and its solution is explicit:

$$x = x_0 + [DF(x_0)]^{-1}(y - y_0). \tag{3.14}$$

Recall that $y_0 = F(x_0)$. For (3.14) to make sense, we must know that the matrix $DF(x_0)$ is invertible. This is a generalization of the condition $f'(f^{-1}(y_0)) \neq 0$ in Proposition 3.1. A natural guess then is that a map $F : \mathbb{R}^n \to \mathbb{R}^n$ is invertible in a neighborhood of a point $y_0 = F(x_0)$ if the derivative matrix $DF(x_0)$ is an invertible matrix. This will be our goal.

3.3 Some preliminaries

We will need the following auxiliary lemmas for the proof of the inverse function theorem.

Lemma 3.3 *Let A be an $n \times n$ matrix, $U \subset \mathbb{R}^n$ be an open set, and a map $f : U \to \mathbb{R}^n$ be continuously differentiable on U. Set $g(x) = Af(x)$, then $g : U \to \mathbb{R}^n$ is a continuously differentiable map with*

$$Dg(x) = ADf(x). \tag{3.15}$$

Proof. Recall that the entries of the matrix Dg are

$$[Dg(x)]_{ij} = \frac{\partial g_i(x)}{\partial x_j},$$

and

$$g_i(x) = \sum_{k=1}^{n} A_{ik} f_k(x).$$

It follows that

$$\frac{\partial g_i(x)}{\partial x_j} = \sum_{k=1}^{n} A_{ik} \frac{\partial f_k(x)}{\partial x_j} = \sum_{k=1}^{n} A_{ik} [Df(x)]_{kj} = (ADf(x))_{ij},$$

thus $Dg(x) = ADf(x)$, and we are done. $\square$

Lemma 3.4 *Let A be an $n \times n$ matrix with entries such that $A_{ij} \leq \varepsilon$ for all $1 \leq i, j \leq n$, then for any vector $v \in \mathbb{R}^n$ we have $\|Av\| \leq n\varepsilon \|v\|$.*

Proof. First, we recall the inequality

$$(x_1 + \cdots + x_n)^2 \leq n(x_1^2 + \cdots + x_n^2). \tag{3.16}$$

To see that (3.16) holds, write

$$\begin{aligned}(x_1 + \cdots + x_n)^2 &= x_1^2 + \cdots + x_n^2 + 2 \sum_{1 \leq i < j \leq n} x_i x_j \leq x_1^2 + \cdots + x_n^2 + \sum_{1 \leq i < j \leq n} (x_i^2 + x_j^2) \\ &= x_1^2 + \cdots + x_n^2 + (n-1)(x_1^2 + \cdots + x_n^2) = n(x_1^2 + \cdots + x_n^2).\end{aligned}$$

Alternatively, one can use the Cauchy-Schwarz inequality: observe that

$$x_1 + x_2 + \cdots + x_n = (1, 1, , \ldots, 1) \cdot (x_1, x_2, \ldots, x_n),$$

thus

$$|x_1 + x_2 + \cdots + x_n|^2 \le \sqrt{\|(1,1,\dots,1)\|}\|x\| = \sqrt{n}\|x\|,$$

which is (3.16). Note that for each $1 \le i \le n$ we have

$$|(Av)_i| = \Big|\sum_{j=1}^n A_{ij}v_j\Big| \le \sum_{j=1}^n |A_{ij}||v_j| \le \varepsilon \sum_{j=1}^n |v_j|,$$

so that, using (3.16), we get

$$|(Av)_i|^2 \le \varepsilon^2 \Big(\sum_{j=1}^n |v_j|\Big)^2 \le n\varepsilon^2 \sum_{j=1}^n |v_j|^2 = n\varepsilon^2\|v\|^2.$$

Next, summing over i, we get

$$\|Av\|^2 = \sum_{i=1}^n |(Av)_i|^2 \le n^2\varepsilon^2\|v\|^2,$$

and the claim of the lemma follows. □

Lemma 3.5 *Let A be an $n \times n$ invertible matrix. Then there exist $\alpha_A > 0$ and $\beta_A > 0$ so that*

$$\beta_A\|x\| \le \|Ax\| \le \alpha_A\|x\|, \quad \text{for all } x \in \mathbb{R}^n. \tag{3.17}$$

Proof. The function $G(x) = \|Ax\|$ is continuous on $\mathbb{R}^n$, and the unit sphere $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$ is a compact subset of $\mathbb{R}^n$. Hence, G attains its maximum α_A and minimum β_A on S. Moreover, as the matrix A is invertible, $G(x) \ne 0$ for all $x \in S$, thus $\alpha_A > 0$ and $\beta_A > 0$. However, each $x \in \mathbb{R}^n$ can be written as $x = ry$, with $\|y\| = 1$, and $r = \|x\|$, so that $y \in S$ and $r \ge 0$. Then, we have

$$\|Ax\| = \|A(ry)\| = \|rA(y)\| = r\|A(y)\| = rG(y) = \|x\|G(y),$$

thus

$$\|Ax\| \le \|x\|\alpha_A,$$

and

$$\|Ax\| \ge \|x\|\beta_A,$$

finishing the proof. □

Corollary 3.6 *Let A be an $n \times n$ invertible matrix, then the map $F(x) = Ax$ maps any open set $U \in \mathbb{R}^n$ to an open set.*

Proof. Lemma 3.5 applied to the matrix A^{-1} implies that there exists $\alpha > 0$ so that

$$\|A^{-1}w\| \le \alpha\|w\| \text{ for all } w \in \mathbb{R}^n. \tag{3.18}$$

Let now U be any open set and $V = F[U]$ be the image of U under F. Take any $y_0 \in F[U]$, so that $y_0 = Ax_0$, with $x_0 \in U$. As the set U is open, it contains a ball $B(x_0, r)$ with some $r > 0$. Consider any $z \in B(y_0, \rho)$, with $\rho < r/\alpha$. Then $z = F(x)$, with $x = A^{-1}(z)$, and we have, using (3.18):

$$\|x - x_0\| = \|A^{-1}(z) - A^{-1}(y_0)\| = \|A^{-1}(z - y_0)\| \le \alpha\|z - y_0\| \le \alpha\rho < r, \tag{3.19}$$

so that $x \in B(x_0, r)$, and thus $x \in U$. It follows that $z \in F[U]$, thus the ball $B(y_0, \rho)$ is contained in V and the set V is open. □

3.4 The inverse function theorem for maps $\mathbb{R}^n \to \mathbb{R}^n$

We will now prove the inverse function theorem for maps from $\mathbb{R}^n$ to $\mathbb{R}^n$.

Theorem 3.7 *Let $U \in \mathbb{R}^n$ be an open set, and $x_0 \in U$. Let $f : U \to \mathbb{R}^n$ be a continuously differentiable map, and set $y_0 = f(x_0)$. Suppose that the derivative matrix $Df(x_0)$ is invertible. Then there exist an open set $V \subset U$ such that $x_0 \in V$, and an open set $W \subset \mathbb{R}^n$ such that $y_0 \in W$, so that f is a one-to-one map from V to W. Moreover, the inverse map $g = f^{-1} : W \to V$ is also continuously differentiable and for $y \in W$ we have $Dg(y) = [Df(g(y))]^{-1}$.*

Proof. Step 1. Reduction to the case $Df(x_0) = I$. We first note that it suffices to prove the theorem under an additional assumption that

$$Df(x_0) = I_n, \tag{3.20}$$

the $n \times n$ identity matrix I_n. Indeed, if $Df(x_0) \neq I_n$, we consider the map

$$\tilde{f}(x) = [Df(x_0)]^{-1} f(x).$$

Then the gradient matrix of $\tilde{f}$ at the point x_0 is I_n:

$$D\tilde{f}(x_0) = [Df(x_0)]^{-1} Df(x_0) = I_n,$$

as follows from Lemma 3.3. Since the matrix $Df(x_0)$ is invertible, the function f is one-to-one from a neighborhood V of x_0 to a neighborhood W of $y_0 = f(x_0)$ if and only if the function $\tilde{f}$ is a one-to-one map from V to $\tilde{W} = [Df(x_0)]^{-1}W$, and $\tilde{W}$ is a neighborhood of the point $\tilde{y}_0 = \tilde{f}(x_0)$. This is a consequence of Corollary 3.6. Again, as the matrix $Df(x_0)$ is invertible, and by Lemma 3.3 we have

$$D\tilde{f}(x) = [D\tilde{f}(x_0)]^{-1} Df(x),$$

the function $\tilde{f}$ is continuously differentiable if and only if f is continuously differentiable. Hence, we may assume without any loss of generality that f satisfies (3.20), and this is what we will do for the rest of the proof.

It is convenient to write

$$f(x) = x + E(x). \tag{3.21}$$

As

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij},$$

where δ_{ij} is the Kronecker delta: $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$, it follows from (3.20) that

$$DE(x_0) = 0. \tag{3.22}$$

Since the map $f(x)$ is continuously differentiable in U, so is $E(x)$, hence for any $\varepsilon > 0$ there exists $r > 0$ so that

$$\left|\frac{\partial E_i}{\partial x_j}\right| < \varepsilon \tag{3.23}$$

for all $x \in B(x_0, r)$. It follows that for every $v \in \mathbb{R}^n$ and all $x \in B(x_0, r)$ we have, by the triangle inequality

$$\|Df(x)v\| = \|(I + DE(x))v\| = \|v + DE(x)v\| \geq \|v\| - \|DE(x)v\|.$$

Now, Lemma 3.4, with $A = DE(x)$, implies that for every $v \in \mathbb{R}^n$ and all $x \in B(x_0, r)$ we have

$$\|Df(x)v\| \geq \|v\| - \varepsilon n\|v\| \geq \frac{1}{2}\|v\|,$$

as long as $\varepsilon < 1/(2n)$. In particular, it follows that the kernel of the matrix $Df(x)$ is $\{0\}$, and the matrix $Df(x)$ is invertible for all $x \in B(x_0, r)$.

Step 2. Reformulation as a fixed point problem for a contraction mapping. Now, we turn to solving

$$f(x) = y \tag{3.24}$$

for y close to $y_0 = f(x_0)$, with the assumption $Df(x_0) = I_n$. Note that, because of (3.20), the approximate linear equation (3.13) in this case takes the particularly simple form

$$f(x_0) + x - x_0 = y, \tag{3.25}$$

with $y_0 = f(x_0)$, that, naturally, has a unique solution

$$x = x_0 + y - y_0.$$

We assume that $y \in B(y_0, \rho)$ and will later see how small ρ needs to be. In order to turn (3.24) into a contraction mapping question, let us reformulate (3.24) as a fixed point problem by setting

$$F(x) = x - f(x) + y = y - E(x), \tag{3.26}$$

again, with $y \in B(y_0, \rho)$ fixed. Recall that $E(x)$ is defined in (3.21) and satisfies both (3.22) and (3.23). In addition, we have

$$E(x_0) = f(x_0) - x_0 = y_0 - x_0. \tag{3.27}$$

The map F depends on y – we will drop it in the notation but you should keep this in mind. Note that x satisfies (3.24) if and only if it satisfies

$$F(x) = x. \tag{3.28}$$

This is a fixed point problem and we will address it using the contraction mapping principle. Note that for $y = y_0$ we have

$$F(x_0) = y_0 - E(x_0) = y_0 - (y_0 - x_0) = x_0,$$

so that x_0 is a fixed point for F, and for other y we see that

$$F(x_0) = y - E(x_0) = y - y_0 + x_0,$$

is close to x_0 if y is close to y_0. In other words, if y is close to y_0 then x_0 is an approximate fixed point for F. Hence, it is natural to expect that a true fixed point exists near x_0. The following lemma is a key part of the proof.

Lemma 3.8 *There exist $\rho > 0$ and $r_1 > 0$ so that for each $y \in B(y_0, \rho)$ the map F maps the closed ball $\bar{B}(x_0, r_1)$ to itself and is a contraction on $\bar{B}(x_0, r_1)$.*

As for each $y \in B(y_0, \rho)$, the map F is a contraction, the contraction mapping principle will then imply that for each $y \in B(x_0, \rho)$ there exists a unique $x \in B(x_0, r_1)$ such that $F(x) = x$, which means that there is a unique solution to (3.24) in $B(x_0, r_1)$ for each $y \in B(y_0, \rho)$. In other words, we may define the inverse map f^{-1} that maps $B(y_0, \rho)$ to $B(x_0, r_1)$.

Proof of Lemma 3.8. The proof requires us to verify two conditions: first, that F maps the closed ball $\bar{B}(x_0, r_1)$ to itself, and, second, that there exists $q \in (0, 1)$ so that for any $z, w \in \bar{B}(x_0, r_1)$ we have

$$\|F(z) - F(w)\| \leq q\|z - w\|. \tag{3.29}$$

We will start with (3.29). Note that the derivative matrix of $F(x)$ is simply $DF(x) = DE(x)$, and thus the entries of the matrix $DF(x)$ satisfy

$$DF(x_0) = 0 \tag{3.30}$$

and

$$\Big|\frac{\partial F_i}{\partial x_j}\Big| = \Big|\frac{\partial E_i}{\partial x_j}\Big| \le \varepsilon \text{ for all } 1 \le i, j \le n, \tag{3.31}$$

for all $x \in B(x_0, r_1)$, as long as r is chosen so that $r_1 < r$, with r defined as in (3.23), with ε sufficiently small.

To verify that F satisfies (3.29), let us take $z, w \in B(x_0, r)$, with r determined by the condition that (3.23) holds in $B(x_0, r)$, and define

$$g(t) = E(z + t(w - z)),$$

for $0 \le t \le 1$, so that $F(z) = y - g(0)$, and $F(w) = y - g(1)$. The fundamental theorem of calculus implies that

$$F_k(w) - F_k(z) = -\int_0^1 g_k'(t)dt, \quad 1 \le k \le n. \tag{3.32}$$

We compute the derivative $g_k'(t)$ using the chain rule:

$$\frac{dg_k(t)}{dt} = \sum_{j=1}^{n} \frac{\partial E_k(z + t(w - z))}{\partial x_j}(w_j - z_j). \tag{3.33}$$

Now, using (3.31) in (3.33) gives

$$\Big|\frac{dg_k(t)}{dt}\Big| \le \sum_{j=1}^{n} \varepsilon |w_j - z_j| \le n\varepsilon \|w - z\|, \tag{3.34}$$

for all $1 \le k \le n$. Using this in (3.32) gives

$$|F_k(w) - F_k(z| \le \int_0^1 |g_k'(t)|dt \le \int_0^1 n\varepsilon\|w - z\|dt = n\varepsilon\|w - z\|, \tag{3.35}$$

for all $1 \le k \le n$. It follows from (3.35) that

$$\|F(w) - F(z)\| \le \Big(\sum_{k=1}^{n} |F_k(w) - F_k(z|^2\Big)^{1/2} \le (n(n^2\varepsilon^2\|w - z\|^2))^{1/2} = \varepsilon n^{3/2}\|w - z\|. \tag{3.36}$$

Therefore, F satisfies (3.29) with $q = 1/4$ on $\bar{B}(x_0, r)$ if we take $\varepsilon = 1/(4n^{3/2})$ and then take $r_1 > 0$ so small that (3.23) holds in $\bar{B}(x_0, r_1)$. Thus, we have established that

$$\|F(w) - F(z)\| \le \frac{1}{4}\|w - z\|, \quad \text{for all } z, w \in B(x_0, r_1), \tag{3.37}$$

if r_1 is sufficiently small. Note that even though the map F depends on y, this estimate holds for all $y \in B(y_0, \rho)$: this will be important in Step 3 below.

Next, we need to verify that F maps $\bar{B}(x_0, r_1)$ to itself if ρ is sufficiently small and $y \in B(y_0, \rho)$. Note that

$$F(x_0) = y - E(x_0) = x_0 - y_0 + y, \tag{3.38}$$

hence we have

$$\|F(x_0) - x_0\| = \|y - y_0\| \leq \rho < \frac{r_1}{2}, \tag{3.39}$$

provided we take $\rho < r_1/2$, thus, in particular, $F(x_0) \in B(x_0, r_1)$. Next, we take $x \in B(x_0, r_1)$ and write, using the triangle inequality

$$\|F(x) - x_0\| = \|F(x) - F(x_0) + F(x_0) - x_0\| \leq \|F(x) - F(x_0)\| + \|F(x_0) - x_0\|. \tag{3.40}$$

The first term in the right can be estimated using (3.37), and the second using (3.39), to give

$$\|F(x) - x_0\| \leq \|F(x) - F(x_0)\| + \|F(x_0) - x_0\| \leq \frac{1}{4}\|x - x_0\| + \frac{r_1}{2} \leq r_1, \tag{3.41}$$

finishing the proof of Lemma 3.8. □

Exercise 3.9 Use the above argument to show that there exists r_0 so that for any $r < r_0$ the image of the ball $B(x_0, r)$ under f is an open set: it contains a ball $B(y_0, \rho)$ centered around $y_0 = f(x_0)$, with $\rho > 0$ that depends on r.

Step 3: continuity of the inverse map. We now show that the inverse map $g = f^{-1}$ is a continuous map from $B(y_0, \rho)$ to $B(x_0, r_1)$, provided that ρ and r_1 are sufficiently small. Recall that for a given $y \in B(y_0, \rho)$, the point $g(y) \in B(x_0, r_1)$ is the unique fixed point in $B(x_0, r_1)$ of the map

$$F(x) = y - E(x),$$

that depends on y, and that is how $g(y)$ depends on y. To show that $g(y)$ is continuous, we need to show that the fixed point of F depends continuously on y. Let us take $y_1, y_2 \in B(x_0, \rho)$ and consider the corresponding maps

$$F_1(x) = y_1 - E(x), \quad F_2(x) = y_2 - E(x).$$

Then, for any $z, w \in B(x_0, r_1)$, we have

$$F_1(z) - F_2(w) = y_1 - E(z) - (y_2 - E(w)) = E(w) - E(z) + y_1 - y_2, \tag{3.42}$$

hence

$$\|F_1(z) - F_2(w)\| \leq \|E(z) - E(w)\| + \|y_1 - y_2\|. \tag{3.43}$$

Using (3.37), we get

$$\|F_1(z) - F_2(w)\| \leq \frac{1}{4}\|z - w\| + \|y_1 - y_2\|. \tag{3.44}$$

Let us now take $z = g(y_1)$ and $w = g(y_2)$, so that $F_1(z) = z$ and $F_2(w) = w$. It follows from (3.44) that

$$\|z - w\| \leq \frac{1}{4}\|z - w\| + \|y_1 - y_2\|, \tag{3.45}$$

thus

$$\|z - w\| \leq \frac{4}{3}\|y_1 - y_2\|. \tag{3.46}$$

In other words, we have shown that the map g satisfies

$$\|g(y_1) - g(y_2)\| \leq \frac{4}{3}\|y_2 - y_1\|, \text{ for all } y_1, y_2 \in B(y_0, \rho). \tag{3.47}$$

It follows that g is a continuous map on $B(y_0, \rho)$.

Step 4: computing the derivative of the inverse map. Now that we have shown that the inverse map $g = f^{-1} : B(y_0, \rho) \to B(x_0, r_1)$ is well defined and continuous, provided that ρ and r_1 are sufficiently small, it remains to show that g is continuously differentiable and for $y \in B(y_0, \rho)$ we have

$$Dg(y) = [Df(g(y))]^{-1}. \tag{3.48}$$

Note that it actually suffices to prove differentiability and relation (3.48) only for $y = y_0 = f(x_0)$ since x_0 was chosen arbitrarily, as an arbitrary point such that $Df(x_0)$ is an invertible matrix.

In this step of the proof we will no longer assume that $Df(x_0)$ is the $n \times n$ identity matrix I_n, and will denote $A = Df(x_0)$. For a given $\varepsilon > 0$ we can find $\delta > 0$ such that $\delta < r_0$, with r_0 as in Exercise 3.9, and if $\|x - x_0\| < \delta$, then

$$\|f(x) - f(x_0) - A(x - x_0)\| < \varepsilon \|x - x_0\|. \tag{3.49}$$

The result of Exercise 3.9 shows that the image of the ball $B(x_0, \delta)$ under f is an open set, hence it contains a ball $B(y_0, \rho)$. Let us take any $y \in B(y_0, \rho)$, so that $y = f(x)$ with some $x \in B(x_0, \delta)$. Thus we may use (3.49) with $y = f(x)$ and $y_0 = f(x_0)$, and $x = g(y)$, $x_0 = g(y_0)$, to get

$$\|y - y_0 - A(g(y) - g(y_0))\| < \varepsilon \|g(y) - g(y_0)\|. \tag{3.50}$$

Let us write

$$y - y_0 - A(g(y) - g(y_0)) = A[A^{-1}(y - y_0) - (g(y) - g(y_0))], \tag{3.51}$$

and recall that by Lemma 3.5 there exists $\beta > 0$ so that

$$\|Aw\| \geq \beta \|w\| \text{ for all } w \in \mathbb{R}^n. \tag{3.52}$$

Using this in (3.51) gives

$$\begin{aligned} \|y - y_0 - A(g(y) - g(y_0))\| &= \|A[A^{-1}(y - y_0) - (g(y) - g(y_0))]\| \geq \beta \|A^{-1}(y - y_0) - g(y) - g(y_0)\| \\ &= \beta \|g(y) - g(y_0) - A^{-1}(y - y_0)\|. \end{aligned} \tag{3.53}$$

Inserting this into (3.50) gives

$$\|g(y) - g(y_0) - A^{-1}(y - y_0)\| \leq \frac{\varepsilon}{\beta} \|g(y) - g(y_0)\|, \tag{3.54}$$

for all $y \in B(y_0, \rho)$. This is almost what we need to say that $Dg(y_0) = A^{-1}$ except in the right side we have $\|g(y) - g(y_0)\|$ rather than $\|y - y_0\|$ that we need. However, we can bootstrap: (3.54) implies that

$$\|g(y) - g(y_0)\| - \|A^{-1}(y - y_0)\| \leq \frac{\varepsilon}{\beta} \|g(y) - g(y_0)\|, \tag{3.55}$$

so that

$$\Big(1 - \frac{\varepsilon}{\beta}\Big) \|g(y) - g(y_0)\| \leq \|A^{-1}(y - y_0)\|. \tag{3.56}$$

Lemma 3.5 tells us that there exists $\alpha > 0$ so that

$$\|A^{-1}w\| \leq \alpha \|w\| \text{ for all } w \in \mathbb{R}^n. \tag{3.57}$$

Using (3.57) in (3.56) gives us

$$\Big(1 - \frac{\varepsilon}{\beta}\Big) \|g(y) - g(y_0)\| \leq \alpha \|(y - y_0)\|. \tag{3.58}$$

Now, (3.55) and (3.58) together imply that

$$\|g(y)-g(y_0)-A^{-1}(y-y_0)\| \leq \Big(1-\frac{\varepsilon}{\beta}\Big)^{-1}\frac{\alpha}{\beta}\varepsilon\|y-y_0\|, \tag{3.59}$$

for all $y \in B(y_0,\rho)$. It follows that g is differentiable at y_0 and $Dg(y_0) = A^{-1}$.

The final step, continuity of $Dg(y)$ is surprisingly easy: we know that $Dg(y) = [Df^{-1}(g(y))]$ and the matrix $Df(x)$ is invertible at $x = x_0$. It follows that $\det Df(x) \neq 0$ in a ball around x_0. Then, the explicit formula for the inverse of a matrix in terms of its minors implies that the inverse matrix $[Df]^{-1}(x)$ is a continuous function of x, hence its composition $[Df]^{-1}(g(y))$ with a continuous map $g(y)$ is also continuous. This completes the proof of the inverse function theorem. □

3.5 The implicit function theorem

The implicit function theorem starts with a system of n equations for n unknowns $x = (x_1,\dots,x_n)$, parametrized by $y \in \mathbb{R}^m$:

$$\begin{aligned} &G_1(x_1,\dots,x_n,y) = 0,\\ &\dots\dots\\ &G_n(x_1,\dots,x_n,y) = 0. \end{aligned} \tag{3.60}$$

Let us assume that $z = (z_1,\dots,z_n)$ is a solution of this system for some $y_0 \in \mathbb{R}^m$:

$$G(z,y_0) = 0, \tag{3.61}$$

or, in the system form,

$$\begin{aligned} &G_1(z_1,\dots,z_n,y_0) = 0,\\ &\dots\dots\\ &G_n(z_1,\dots,z_n,y_0) = 0. \end{aligned} \tag{3.62}$$

The question is whether if we take y close to y_0, can we find a solution x of (3.60) that is close to z. The system (3.7), addressed by the inverse function theorem, is a special case of this problem, with $y \in \mathbb{R}^n$ and $G(x,y) = G(x) - y$. There is an elegant way to understand the general case via an application of the inverse function theorem. One problem to apply this theorem is that $G(x,y)$ maps $\mathbb{R}^{n+m}$ to $\mathbb{R}^n$ and not to $\mathbb{R}^{n+m}$, and in the inverse function theorem we need the domain of the map and the image of the map to have the same dimension. To fix this, consider the map

$$\tilde{G}(x,y) = (G(x,y),y), \tag{3.63}$$

that maps $\mathbb{R}^{n+m}$ to $\mathbb{R}^{n+m}$. This simply means that we re-write (3.60) by adding to it m equations of the form $y_k = y_k$, and the system (3.60) is equivalent to

$$\tilde{G}(x,y) = (0,y). \tag{3.64}$$

We know from (3.61) that the point (z,y_0) satisfies

$$\tilde{G}(z,y_0) = (0,y_0). \tag{3.65}$$

In addition, the derivative matrix $D\tilde{G}(x,y)$ has the block form

$$D\tilde{G}(x,y) = \begin{pmatrix} D_xG(x,y) & D_yG(x,y)\\ 0 & I_{m\times m}\end{pmatrix}. \tag{3.66}$$

Thus, the matrix $D\tilde{G}(x,y)$ is invertible if and only if the matrix $D_xG(x,y)$ is invertible.

Exercise 3.10 Check this.

Let us assume that the derivative matrix $DG(z, y_0)$ is invertible and take y close to y_0. Now, the inverse function theorem implies that for all y close to y_0 the system (3.64) has a unique solution x close to z – this is the implicit function theorem. Let us formulate it precisely.

Theorem 3.11 *Let a map $G : \mathbb{R}^{n+m} \to \mathbb{R}^n$ satisfy $G(x_0, y_0) = 0$ for some $x_0 \in \mathbb{R}^n$ and $y_0 \in \mathbb{R}^m$. Assume that G is continuously differentiable at (x_0, y_0) and the $n \times n$ matrix $D_x G(x_0, y_0)$ is invertible. Then there exist $r > 0$ and $\rho > 0$ so that for every $y \in B(y_0, \rho)$ there exists a unique $x \in B(x_0, r)$ such that $G(x, y) = 0$.*