Contraction mappings

Definition 1. Suppose that V is a normed vector space, and that $T : V \to V$. T is called a *contraction mapping* (or *contractive mapping*, or *contraction*) if there exists a constant k with $0 \le k < 1$, such that

$$\|T(x) - T(y)\| \le k\|x - y\| \quad \text{for all } x, y \in V. \tag{1}$$

EXAMPLES

1. Suppose that $f : \mathbb{R} \to \mathbb{R}$ and $|f'(x)| \le k < 1$ for all $x \in \mathbb{R}$. By the Mean Value Theorem, for any $x, y \in \mathbb{R}$, there exists a c between x and y such that $f(x) - f(y) = f'(c)(x - y)$. Thus $|f(x) - f(y)| \le k|x - y|$, and f is a contraction.

2. Suppose that $T : \mathbb{R}^n \to \mathbb{R}^n$ is given by $T(x) = Ax + b$, where A is an $n \times n$ matrix with $\|A\| = k < 1$. Then

   $$\|T(x) - T(y)\| = \|A(x - y)\| \le \|A\|\|x - y\| = k\|x - y\|,$$

   and T is a contraction.

Definition 2. Suppose that $T : V \to V$. Then $x \in V$ is a *fixed point* of T iff $T(x) = x$.

Definition 3. A normed vector space V is called a *Banach space* if it is complete (i.e. Cauchy sequences converge).

Theorem 1 (The Contraction Mapping Theorem)**.** *Suppose that V is a Banach space, and that $T : V \to V$ is a contraction. Then T has a unique fixed point x in V. Moreover, if x_0 is an arbitrary point in V, and the sequence $\{x_n\}$ is defined by*

$$x_{n+1} = T(x_n), \quad n = 0, 1, 2, \ldots, \tag{2}$$

then $x_n \to x$.

Proof. Let $x_0 \in V$, and define the sequence $\{x_n\}$ as above. We will show that $\{x_n\}$ is Cauchy. We have,

$$\begin{aligned}
\|x_2 - x_1\| = \|T(x_1) - T(x_0)\| &\le k\|x_1 - x_0\| \\
\|x_3 - x_2\| = \|T(x_2) - T(x_1)\| &\le k\|x_2 - x_1\| \le k^2\|x_1 - x_0\| \\
&\vdots \\
\|x_{n+1} - x_n\| \le k^n\|x_1 - x_0\|, &\quad n = 1, 2, \ldots.
\end{aligned}$$

Now, if $m > n$, then

$$
\begin{aligned}
\|x_m - x_n\| &= \|x_m - x_{m-1} + x_{m-1} + \cdots + x_{n+1} - x_n\| \\
&\le \|x_m - x_{m-1}\| + \|x_{m-1} - x_{m-2}\| + \cdots + \|x_{n+1} - x_n\| \\
&\le \left(\sum_{j=n}^{m-1} k^j \right) \|x_1 - x_0\| \\
&\le \left(\sum_{j=n}^{\infty} k^j \right) \|x_1 - x_0\| \\
&= \frac{k^n}{1-k} \|x_1 - x_0\| .
\end{aligned}
$$

It follows that the sequence $\{x_n\}$ is Cauchy, and therefore convergent, since V is complete. Thus $x_n \to x \in V$.

To show that x is a unique fixed point for T, note that

$$\|x_{n+1} - T(x)\| = \|T(x_n) - T(x)\| \le k\|x_n - x\|,$$

and therefore, $x_{n+1} \to T(x)$. But since $\{x_{n+1}\}$ also converges to x, $T(x) = x$. Thus, x is a fixed point of T.

Finally, suppose that $y \in V$ is also a fixed point of T. Then

$$\|x - y\| = \|T(x) - T(y)\| \le k\|x - y\|,$$

with $k < 1$. But this is impossible unless $y = x$. Thus, x is a unique fixed point of T. ■

Often in applications we only know that a mapping is a contraction on some subset of V. For instance, Newton's method is a contraction mapping on a neighborhood of a root of f. One must begin close enough in order for the sequence of approximations to converge. The following theorem and corollary are thus often more useful. The proof of the theorem is similar to that of Theorem 1.

Theorem 2 (Contraction Mapping on a Subset)**.** *Suppose that V is a Banach space, and that S is a closed subset of V. Suppose that $T : S \to S$ is a contraction on S, i.e. that $\|T(x) - T(y)\| \le k\|x - y\|$ for all $x, y \in S$, with $k < 1$. Then T has a unique fixed point, x, in S. Moreover, if x_0 is an arbitrary point in S, and the sequence $\{x_n\}$ is defined by $x_{n+1} = T(x_n)$ for $n = 0, 1, 2, \ldots,$ then $x_n \to x$.*

Corollary 1. *Let $T(x)$ be continuously differentiable and suppose $T(x_*) = x_*$. If $|T'(x_*)| < 1$ then the iteration $x_{n+1} = T(x_n)$ converges to x_* if x_0 is close enough to x_*.*

Proof. $T(x) - T(y) = T'(z)(x - y)$ by the mean value theorem. Since T' is continuous, near x_* $T'(z) < 1$ for x, y close enough to x_*, so T is a contraction near x_*. It remains to be shown that T maps a neighborhood of x_* into itself. Let $S = [x_* - \epsilon,\, x_* + \epsilon]$ where ϵ is small enough so that T is a contraction on S. Then if $x \in S$, $|T(x) - x_*| = |T(x) - T(x_*)| < |x - x_*| \le \epsilon$, so $T(x) \in S$. Thus $T : S \to S$. ■

EXAMPLES

1. *Newton's Method.* Let $f : \mathbb{R} \to \mathbb{R}$. Suppose $f(x_*) = 0$, f is twice continuously differentiable in a neighborhood of x_*, and $f'(x_*) \neq 0$. Define
$$T(x) = x - \frac{f(x)}{f'(x)}.$$
Notice that the fixed points of T are exactly the solutions of $f(x) = 0$. And we have
$$T'(x) = 1 - \frac{[f'(x)]^2 - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2}.$$
Thus $T'(x_*) = 0$, so by Corollary 1, if x_0 is close enough to x_*, the iteration
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$
will converge to x_*.

2. (Without Newton's method.) To find the solution of
$$2x - \cos x = 0,$$
first write the equation as
$$x = \frac{1}{2}\cos x =: T(x).$$
Since $|T'(x)| = \left|\frac{1}{2}\sin x\right| \leq 1/2$ for all $x \in \mathbb{R}$, $T : \mathbb{R} \to \mathbb{R}$ is a contraction. Thus T has a unique fixed point x, which is also the solution of the original equation. Moreover, for any starting point $x_0 \in \mathbb{R}$, if $x_n = T(x_{n-1}) = \frac{1}{2}\cos(x_{n-1})$, then $x_n \to x$.

   Taking, for example, $x_0 = 0$, then
$$x_1 = \frac{1}{2}\cos 0 = .5,\ x_2 = .43879128,\ x_3 = .452632921,\ x_4 = .449649376,\ x_{14} = x_{15} = .450183611.$$

3. To solve the system of equations
$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix},$$
rewrite the system as
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & -1/2 \\ -1/2 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 3/2 \\ 3/2 \end{pmatrix} =: Ax + b.$$
Since $\|A\| = 1/2$, the mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x) = Ax + b$ is a contraction. Thus, T has a unique fixed point $x^* \in \mathbb{R}^2$, which is therefore the unique solution of the original equation. For any starting point x_0, the iteration $x_n = T(x_{n-1}) = Ax_{n-1} + b$ will converge to x^*.

   Taking, for example, $x_0 = (0, 0)^T$. Then
$$x_1 = \begin{pmatrix} 3/2 \\ 3/2 \end{pmatrix},\ x_2 = \begin{pmatrix} 3/4 \\ 3/4 \end{pmatrix},\ x_3 = \begin{pmatrix} 9/8 \\ 9/8 \end{pmatrix},\ x_4 = \begin{pmatrix} 15/16 \\ 15/16 \end{pmatrix},\ \ldots.$$
Note that the unique solution of the original equation is $x^* = (1, 1)^T$.

Error analysis for iteration of functions of a single variable

Let $T(x)$ be a function of a real (or complex) variable x. Suppose T has a fixed point x_*. Define the usual iteration by $x_{n+1} = T(x_n)$. Let

$$e_n = x_n - x_*$$

be the error on the nth iterate. Suppose the derivatives of T exist. Then

$$\begin{aligned} e_{n+1} &= x_{n+1} - x_* = T(x_n) - T(x_*) = T'(z_n)(x_n - x_*) \\ &= T'(z_n)e_n \end{aligned}$$

where z_n is between x_* and x_n. Thus, the iteration is at least linear. If $T'(x_*) = 0$ then $T'(z_n) \to 0$, so the iteration will be superlinear. This pleasant situation can be improved even further if more derivatives are zero. To see this, let q be the smallest integer such that the q-th derivative is nonzero:

$$T^{(k)}(x_*) = 0, \quad \text{for} \quad 1 \le k < q, \quad \text{and} \quad T^{(q)}(x_*) \ne 0.$$

Then, expanding T around x_*, we have

$$\begin{aligned} e_{n+1} &= x_{n+1} - x_* \\ &= T(x_n) - T(x_*) \\ &= T(x_* + e_n) - T(x_*) \\ &= T(x_*) + e_n T'(x_*) + \frac{e_n^2}{2} T''(x_*) + \cdots + \frac{e_n^{q-1}}{(q-1)!} T^{(q-1)}(x_*) + \frac{e_n^q}{q!} T^{(q)}(z_q) - T(x_*) \\ &= \left(\frac{T^{(q)}(z_q)}{q!} \right) e_n^q \end{aligned}$$

Therefore, the order of convergence of the iteration is q. For instance, in Newton's method, $T'(x_*) = 0$, and the convergence is quadratic.

Application: Existence and uniqueness of solutions to differential equations

One question that arises when confronted with a differential equation

$$x' = f(x), \qquad x(0) = x_0, \tag{3}$$

is whether a solution even exists, and if it does, is it unique? This is important, for example, if you are trying to approximate the solution numerically. Equation (3) denotes a system of equations, actually, since we take $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ to be a function from $\mathbb{R}^n$ to itself. That is,

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ f_2(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{pmatrix}.$$

So the equation (3) can be written as

$$\begin{aligned} x_1'(t) &= f_1(x_1, x_2, \dots, x_n) \\ x_2'(t) &= f_2(x_1, x_2, \dots, x_n) \\ &\vdots \\ x_n'(t) &= f_n(x_1, x_2, \dots, x_n) \end{aligned}$$

...but (3) is much more pleasant. There x_0 is a vector of initial conditions.

Existence: Not every equation has a solution. For consider

$$x' = \begin{cases} -1 & \text{if } x > 0 \\ 1 & \text{if } x \le 0 \end{cases}$$

This equation does not have a solution if $x(0) = 0$ since if it did it would have $x' = 1$ and so would be increasing but as soon as x is the slightest bit greater than 0 it is supposed to be decreasing. This is impossible.

Uniqueness: Not every equation has a unique solution even if it has a solution. For consider

$$x' = \frac{3}{2} x^{1/3}, \qquad x(0) = 0. \tag{4}$$

This equation has the trivial solution, as well as others. You can verify that both of the following are solutions of the initial value problem:

$$x(t) = 0 \qquad \text{and} \qquad x(t) = t^{3/2}.$$

The problem in (4) is that the RHS of the equation $f(x) = \frac{3}{2} x^{1/3}$ is not differentiable at $x = 0$. Note that

$$f'(x) = \frac{1}{2} x^{-2/3},$$

which blows up near $x = 0$. This makes it so the equation (4) can behave "badly" for initial conditions at this point. Fortunately this circumstance is the only way that a solution can fail to have a solution. We have a theorem that guarantees that as long as the RHS of a differential equation has a continuous first derivative, then it has a unique solution. We may as well state it...

Theorem 3 (Existence and Uniqueness Theorem for Ordinary Differential Equations)**.** *Consider the differential equation*

$$x' = f(x), \qquad x'(0) = x_0. \tag{5}$$

Then, if f is continuously differentiable at x_0, then there is an $\alpha > 0$ such that there is a unique solution to (5) *for* $0 \le t \le \alpha$.

Note that the theorem only says that we can find a solution for some time $0 \le t \le \alpha$. It doesn't guarantee a solution for all time. Let us prove the theorem in the particular one dimensional case. The extension to the general n-dimensional case is straightforward. The main idea is the *Picard iteration.* So, engage!

Sketch of proof: For x_0 given we define the following functions recursively

$$\begin{aligned} u_0(t) &= x_0 \\ u_1(t) &= x_0 + \int_0^t f\left(u_0(\tau)\right) d\tau \\ &\vdots \\ u_k(t) &= x_0 + \int_0^t f\left(u_{k-1}(\tau)\right) d\tau \end{aligned}$$

In other words,

$$u_{k+1}(t) = T\left(u_k(t)\right), \quad \text{where} \quad T\left(u(t)\right) = x_0 + \int_0^t f\left(u(\tau)\right) d\tau \tag{6}$$

These iterations are called *Picard iterations.* Note that if $T(u(t)) = u(t)$ then $u'(t) = f(u(t))$, so u is a solution of (5). In other words, solutions of (5) are fixed points of T. So by the contraction mapping theorem, it suffices to show that T is a contraction for small t. Recall that, by the Mean Value Theorem, if u and v are continuous then there is a continuous w such that

$$f(u(\tau)) - f(v(\tau)) = f'(w(\tau))\left[u(\tau) - v(\tau)\right].$$

We thus consider

$$\begin{aligned} T(u(t)) - T(v(t)) &= \int_0^t f(u(\tau))d\tau - \int_0^t f(v(\tau))d\tau \\ &= \int_0^t f'(w(\tau))\left[u(\tau) - v(\tau)\right] d\tau \\ \Rightarrow \quad |T(u(t)) - T(v(t))| &= \left| \int_0^t f'(w(\tau))\left[u(\tau) - v(\tau)\right] d\tau \right| \\ &\le \int_0^t \left|f'(w(\tau))\right| \left|u(\tau) - v(\tau)\right| d\tau \\ &\le t \left[\max_{0 \le \tau \le t} |f'(w(\tau))|\right] \max_{0 \le \tau \le t} |u(\tau) - v(\tau)| \\ &\le t \left[\max_{0 \le \tau \le \alpha} |f'(w(\tau))|\right] \max_{0 \le \tau \le \alpha} |u(\tau) - v(\tau)| \end{aligned}$$

Thus, if t is small enough,

$$\|T(u) - T(v)\|_\infty \leq k \, \|u - v\|_\infty \, ,$$

where $k < 1$ and $\|\cdot\|_\infty$ is the max norm on $C[0, \alpha]$. Thus T is a contraction for small t. Since $C[0, t]$ with the $\|\cdot\|_\infty$ norm is complete, it follows that the iterates converge to the unique solution of the differential equation.

NOTES

1. Actually, we have only shown that the iterations converge to a continuous function. But if the function is to be a solution of a differential equation then it must, of course, be differentiable. This extra step is found in most textbooks on differential equations.

2. There is nothing special about $t = 0$. We can easily modify the theorem and proof for $x'(t_0) = x_0$ and solutions guaranteed in $t_0 \leq t \leq t_0 + \alpha$.

3. The theorem only guarantees existence and uniqueness *locally.* That is, for some small time the solution may exist, but this does not mean that we can extend the solution to all time. Consider, for example, the equation

   $$x' = x^2.$$

   The RHS of this equation is as nicely behaved as one might hope for. However, the solution can be obtained by separation of variables, and we see that

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

   where $C = 1/x_0 + t_0$ for $x(t_0) = x_0$. The solution exists for any initial condition at any time, but it blows up at $t = C$, and hence cannot be extended past this time.

4. The generalization of Theorem 3 for

   $$x' = f(x), \qquad x(0) = x_0, \qquad f : \mathbb{R}^n \to \mathbb{R}^n, \quad x_0 \in \mathbb{R}^n,$$

   states that there is a unique solution for small t if f is C^1, which means that f must have continuous partial derivatives near 0. Then the Picard iterations in eqn (6) involve integrations of several variables.

EXAMPLE Consider the equation

$$x'' + x = 0, \qquad x(0) = 1,\ x'(0) = 0.$$

This equation has the unique solution $x(t) = \cos(t)$. It can be written in matrix form as

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix}, \qquad \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

That is, $x' = f(x)$, where $f(x) = [x_2,\ -x_1]^T$, which is obviously continuously differentiable. The Picard iterations are

$$\begin{aligned}
u_0 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
u_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) d\tau \ = \ \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} 0 \\ -1 \end{pmatrix} d\tau \\
&= \begin{pmatrix} 1 \\ -t \end{pmatrix} \\
u_2 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t f\left(\begin{pmatrix} 1 \\ -\tau \end{pmatrix}\right) d\tau \ = \ \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} -\tau \\ -1 \end{pmatrix} d\tau \\
&= \begin{pmatrix} 1 - \frac{1}{2}t^2 \\ -t \end{pmatrix} \\
u_3 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t f\left(\begin{pmatrix} 1 - \tau^2/2 \\ -\tau \end{pmatrix}\right) d\tau \ = \ \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} -\tau \\ -1 + \tau^2/2 \end{pmatrix} d\tau \\
&= \begin{pmatrix} 1 - \frac{1}{2}t^2 \\ -t + \frac{1}{6}t^3 \end{pmatrix} \\
u_4 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t f\left(\begin{pmatrix} 1 - \tau^2/2 \\ -\tau + \tau^3/6 \end{pmatrix}\right) d\tau \ = \ \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} -\tau + \tau^3/6 \\ -1 + \tau^2/2 \end{pmatrix} d\tau \\
&= \begin{pmatrix} 1 - \frac{1}{2}t^2 + \frac{1}{24}t^4 \\ -t + \frac{1}{6}t^3 \end{pmatrix}
\end{aligned}$$

Continuing in this manner we find that the first component of u_n converges to $x(t)$, so

$$x(t) = 1 - \frac{1}{2!}t^2 + \frac{1}{4!}t^4 - \frac{1}{6!}t^6 + \cdots$$

This is just the Taylor series expansion of $\cos t$.

Exercises

1. In Theorem 2, one of the conditions is that S is closed. Explain why this condition is necessary.

2. Prove the following.

   Theorem Let V be a normed linear space, S a subset of V, and $T : S \to S$. Suppose that T has a fixed point $x_* \in S$, and that $\|T(x_*) - T(y)\| \leq k \cdot \|x_* - y\|$ for all $y \in S$, where $k < 1$. Then x_* is the unique fixed point of T in S. Moreover, if x_0 is an arbitrary point in S, and the sequence $\{x_n\}$ is defined by $x_{n+1} = T(x_n)$ for $n = 0, 1, 2, \ldots$, then $x_n \to x_*$.

3. Use the contraction mapping theorem to find the solution to the following equation in (e, ∞).

$$\log x - \frac{x}{3} = 0.$$

4. *Halley's method* is another iterative method for solving $f(x) = 0$. The Halley iteration formula for finding a simple zero is $x_{n+1} = g(x_n)$, where

$$g(x) = x - \frac{2f(x)f'(x)}{2[f'(x)]^2 - f(x)f''(x)}$$

   (a) Show that g is a contraction near a zero of f.

   (b) Find the order of convergence of Halley's method under suitable conditions. How does this compare to Newton's method?

5. Show that $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(x) = Ax + b$, where $A \in M_n(\mathbb{R}),\ b \in \mathbb{R}^n$, is a contraction if $\rho(A) < 1$.

6. Explain how to use the Contraction Mapping Theorem to devise an iterative scheme to solve the linear system of equations $Ax = b$.

7. Show that the differential equation

$$x' = \frac{3}{2}x^{1/3}, \quad x(0) = x_0$$

   has a unique solution for all initial conditions $x_0 > 0$. What is it? Write out the first few Picard iterations.

8. Calculate the Picard iterations for the differential equation

$$x'(t) = x(t), \qquad x(0) = 1,$$

   and thereby find the Taylor series for e^t.