CAMBRIDGE

Cambridge University Press
978-0-521-31897-6 - Metric Spaces: Iteration and Application
Victor Bryant
Excerpt
More information

1

Sequences by iteration

1.1 Where are we heading?

Imagine a table top covered with a thin layer of dust. We give it a flick with a duster, but all the dust settles back on the table top. Suppose that after being rearranged every two specks of dust are closer together than they were before. Then the remarkable thing is that there is one and only one speck of dust which is back in the same place that it started in.

That is an example of a set X and a function $f\colon X \to X$. The property that f brings the points of X closer together is enough to ensure that X has one and only one point which is not moved by f; i.e. a unique $x \in X$ with $x = f(x)$. These so-called 'fixed points' of functions are invaluable in solving equations. For example, the fixed points of the function f given by $f(x) = \frac{1}{3}(x^3 - 5x^2 + 1)$ are those x for which $x = \frac{1}{3}(x^3 - 5x^2 + 1)$; i.e. they are precisely the roots of the equation $x^3 - 5x^2 - 3x + 1 = 0$. In fact, it turns out that this real function f has just one fixed point and that it is easy to find with a calculator; so the unique real root of the cubic equation will be easily found.

In this book our search will be for a large collection of situations X and functions $f\colon X \to X$ which have unique fixed points. The study of such situations provides an interesting piece of mathematics, but the real fascination will lie in the wide range of problems which they will enable us to solve.

1.2 Sequences of numbers by iteration

To find a real root of the cubic equation

$$x^3 + x^2 - 5x - 3 = 0$$

rewrite the equation as

$$x = \tfrac{1}{5}(x^3 + x^2 - 3)$$

or $x = f(x)$, where $f(x) = \frac{1}{5}(x^3 + x^2 - 3)$. Start with a first guess at a root, 0 for example. Apply f to this first guess to give $f(0) = -0.6$. Apply f to this answer to give $f(-0.6) = -0.5712$. Continue in this way to give a list or sequence of answers:

$$\begin{array}{l} 0 \\ \downarrow \\ f(0) = -0.6 \\ \quad\downarrow \\ \quad f(-0.6) = -0.5712 \\ \qquad\downarrow \\ \qquad f(-0.5712) = -0.5720191 \\ \qquad\quad\downarrow \\ \qquad\quad f(-0.5720191) = -0.5719924 \\ \qquad\qquad\downarrow \\ \qquad\qquad f(-0.5719924) = -0.5719933 \\ \qquad\qquad\quad\downarrow \\ \qquad\qquad\quad f(-0.5719933) = -0.5719933 \\ \qquad\qquad\qquad\downarrow \\ \qquad\qquad\qquad\vdots \end{array}$$

This is the best my calculator can manage, but to this accuracy at least there is no point in continuing as the answer -0.5719933 will keep repeating itself. So (to within my calculator's accuracy)

$$f(-0.5719933) = -0.5719933.$$

In other words, we have found an x with $x = f(x)$ – called a 'fixed point' of f. By our above remarks this x will be (again to within my calculator's accuracy) a root of the cubic equation $x^3 + x^2 - 5x - 3 = 0$. Let us check:

$$(-0.5719933)^3 + (-0.5719933)^2 - 5(-0.5719933) - 3 = -0.0000000081\ldots,$$

which, of course, is very nearly zero as expected. So having rewritten our equation in the form $x = f(x)$ the solution of the equation was a mechanical process requiring no further thought. But it must be remembered that this process is only finding an approximate root: the sequence which we produce only seems to come to an end because of the limited accuracy of our calculators. And in all cases we ought to substitute our answers back into the equations to check their validity.

Exercise 1 Find a root, to four decimal places, of the cubic equation $x^3 + 2x^2 - 7x - 1 = 0$ by rearranging it into the form $x = f(x) = \frac{1}{7}(x^3 + 2x^2 - 1)$ and making a first guess between -3 and 1. Reapplying f (or 'iterating' with f) should give

Cambridge University Press
978-0-521-31897-6 - Metric Spaces: Iteration and Application
Victor Bryant
Excerpt
More information

you a sequence settling down (or 'converging') to a root of the equation.

(Calculators are obviously needed! A fairly economical way of calculating $f(x)$, having entered x into the display, is by the following steps:

STO	+	2	=	×	RCL	x^2	−	1	=	÷	7	=
store in memory					recall from memory							and repeat

Of course, a programmable machine can be used to great advantage.)

In Exercise 1 you should have found that one root of the equation $x^3+2x^2-7x-1=0$ is approximately -0.1378: you could check by substituting this value back into the cubic that it does give a value very near zero.

Now let us examine exactly what was happening in the above examples: we wanted to solve the equation $x=f(x)$ and so we made a first guess x_1 at a root and calculated all future guesses by reapplying f to give

$$x_2=f(x_1),\ x_3=f(x_2),\ \ldots,\ x_{n+1}=f(x_n),\ \ldots$$

We then looked at the limit of the sequence

$$x_1, x_2, x_3, \ldots, x_n, \ldots \to x$$

and noted that it had the required property that $x=f(x)$. Of course, as logical people, we must ask some crucial questions about this process:

Will the sequence which we produce by iterating with f definitely converge?

If it does converge will the limit definitely be a root of $x=f(x)$?

If the answers to both these questions were 'yes' and any function led to a sequence which converged to a root of the required equation, then it would be such a wonderful method that practically all other ways of solving equations would be redundant. Let us answer the first question by means of an exercise.

Exercise 2 Repeat the previous exercise by rearranging the cubic equation $x^3+2x^2-7x-1=0$ into the form $x=f(x)=\frac{1}{7}(x^3+2x^2-1)$. Try to find a root of the cubic by making a first guess of 2 or more and iterating with f.

Exercise 3 Rearrange $x^3+2x^2-7x-1=0$ into the form $x=f(x)=\sqrt{\{\frac{1}{2}(-x^3+7x+1)\}}$. Try iterating with this f for various first guesses. See if you can find a sequence in this way which converges to a root of the cubic.

([STO] [x^2] [+/−] [+] [7] [=] [×] [RCL] [+] [1] [=] [÷] [2] [=] [$\sqrt{x}$] and repeat.)

In Exercise 2 you should have got a sequence which very soon became out of hand – clearly not settling down to a root of the cubic. In Exercise 3 if you started with an x_1 between 0 and 2 then you should have got a sequence converging to around 1.91909 – which is very close indeed to a root of the cubic. But for some starting values (such as $x_1=3$) you cannot even proceed because $f(3)$ is not defined. So, to summarise, some functions and some starting points will give sequences converging to the required roots, but often a function will lead to a sequence which does not converge. The principles of numerical analysis are beyond our scope here but it is worth mentioning the Newton–Raphson method of iteration, designed as a systematic way of deciding which function to use for iteration: it often leads to a successful iterative process. We content ourselves with one exercise on this.

Exercise 4 To solve the cubic $2x^3+3x^2+6x+1=0$ rearrange it to $6x^3+6x^2+6x=4x^3+3x^2-1$ and hence $x=f(x)=(4x^3+3x^2-1)/6(x^2+x+1)$. Confirm that

$$f(x)=x-\frac{\text{the original cubic}}{\text{the derivative of the cubic}}$$

(this is how, in general, the f is chosen in the Newton–Raphson method). Choose any x_1, iterate with f and hence find to four decimal places the unique real root of $2x^3+3x^2+6x+1=0$.

The function in Exercise 4 always leads to a convergent sequence, regardless of the first guess x_1. One of the aims of this book is to find a large class of functions which are equally dependable.

Although we have been able to find functions for which the iterative process failed to produce convergent sequences, so far whenever a convergent sequence has been produced it *has* led us to a root. So the answer to our second question

If the sequence produced by iteration with f converges will the limit definitely be a root of $x=f(x)$?

seems much more likely to have a favourable answer. Imagine, for example, that we are trying to find an x for which $x=f(x)$, where f is a cubic in x, and we produce a convergent sequence.

$$x_1, x_2=f(x_1), x_3=f(x_2), \ldots, x_{n+1}=f(x_n), \ldots \to x.$$

Then $f(x_n)$ is simply a cubic in x_n and so as the x_ns get closer to x so do the $f(x_n)$s get closer to $f(x)$; i.e.

$$\begin{array}{cccccl} f(x_1), & f(x_2), & f(x_3), & \ldots, & f(x_n), & \ldots \to f(x) \\ \| & \| & \| & & \| & \\ x_2, & x_3, & x_4, & \ldots, & x_{n+1}, & \ldots (\to x!!!) \end{array}$$

But, as you can see, the sequence of $f(x_n)$s is simply part of the sequence of x_ns, which converge to x. Therefore $x=f(x)$ and the limit x does satisfy the equation.

Most of the functions which we encounter will be 'continuous'; i.e. if x and y are sufficiently close, then $f(x)$ and $f(y)$ are close. If f is of this type and an iterative process leads to a convergent sequence, then the limit of that sequence must be a root of $x=f(x)$, as we now see.

Theorem 1.1 Let X be a subset of the real numbers $\mathbb{R}$, let $f\colon X \to X$ be continuous and let $x_1 \in X$. Then if the sequence

$$x_1, x_2=f(x_1), x_3=f(x_2), \ldots, x_{n+1}=f(x_n), \ldots$$

converges to $x \in X$ it follows that $x=f(x)$.

Proof Let $x_2=f(x_1)$, $x_3=f(x_2)$, etc. and assume that

$$x_1, x_2, x_3, \ldots, x_n, \ldots \to x$$

as stated. Then the x_ns get very close to x and so (by continuity) the $f(x_n)$s get very close to $f(x)$; i.e.

$$\left.\begin{array}{cccccl} f(x_1), & f(x_2), & f(x_3), & \ldots, & f(x_n), & \ldots \to f(x) \\ \| & \| & \| & & \| & \\ x_2, & x_3, & x_4, & & x_{n+1}, & \ldots \to x \end{array}\right\} \therefore\ x=f(x). \qquad \square$$

The above result is only a very special case of one which we shall meet later (where we will also discuss the necessary conditions on the domain X of f).

For those interested enough to check that the condition of continuity of f is actually needed in the above theorem we will consider a slightly more unusual example (the reader not interested in this esoteric point can turn immediately to the goat example below – or indeed to the beginning of Section 1.3 if he has had enough of solving equations by producing real sequences). Let $[x]$ denote the 'largest integer not exceeding x'. For example $[3.142]=3$. This is one of the simpler functions which is not continuous, for it is possible for x to be very close to y (for example 1.99 and 2) without $[x]$ being close to $[y]$. Now let us try to solve the equation

$$x=f(x)=[x]+1-\tfrac{1}{2}([x]+1-x)^2$$

by an iterative process starting with $x_1=1$. This gives

$$x_2=f(x_1)=[1]+1-\tfrac{1}{2}([1]+1-1)^2=2-\tfrac{1}{2}=1.5,$$
$$x_3=f(x_2)=[1.5]+1-\tfrac{1}{2}([1.5]+1-1.5)^2=2-\tfrac{1}{2}(0.5)^2=1.875 \quad \text{etc.}$$

which leads to the sequence

$$1, 1.5, 1.875, 1.9921875, 1.9999695, \ldots$$

which is clearly converging to 2. So

$$\begin{array}{llll} f(1), & f(1.5), & f(1.875), & f(1.9921875), \ldots \\ \| & \| & \| & \| \\ 1.5, & 1.875, & 1.9921875, & 1.9999695, \ldots \end{array}$$

also converges to 2. But

$$f(1), f(1.5), f(1.875), \ldots \not\to f(2)$$

since $f(2)=2.5$. So the limit does not satisfy $x=f(x)$ and the continuity of f *is* needed in the above theorem.

To conclude our series of numerical examples we will solve a very famous problem concerning a goat: the complicated function with which we will iterate begins to show the power of this method. The problem is that a farmer owns a circular plot of land (of unit diameter, say) which is covered in grass. He wants to tie his goat by a long rope, one end attached to a point on the circumference of the field and the other end attached to the goat. How long should the rope be in order that the goat can eat exactly half the grass?

Exercise 5 Show (or take my word for it) that the shaded area in Figure 1.1 opposite is

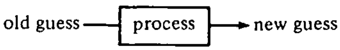
$$\frac{1}{2}\sin^{-1}x-\frac{x}{2}(1-x^2)^{1/2}+x^2\cos^{-1}x.$$

Hence find an equation in x if the shaded area is equal to half of the area of the larger circle. By rearranging your equation into the form $x=f(x)$ for some f and by choosing a suitable first guess for x_1, obtain a sequence which converges to the required value of x.

Note that the equation you get in Exercise 5 may have other roots besides the one relevant to the goat problem, so make a quick mental check that the limit of your sequence is the right sort of size. The rearrangement which I used was obtained by finding an expression for $\sin^{-1} x$ and hence for x: the iterative process with a starting value of just over $\frac{1}{2}$ then produced a convergent sequence (with limit of approximately 0.57935).

1.3 Iterations in a different world

The general principle behind our iterative techniques in the previous section was



where the process can be repeated to give a sequence of improving guesses. In all the examples so far the process has been dealing with numbers, but now let us illustrate a process which deals with something different.

Note before proceeding that the list of numbers

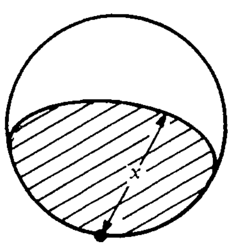
3, 2, 1, 1, 0, 0, 0

is an inventory of itself, in the sense that

3	2	1	1	0	0	0
‖	‖	‖	‖	‖	‖	‖
no. of 0s in the list	no. of 1s in the list	no. of 2s in the list	no. of 3s in the list	no. of 4s in the list	no. of 5s in the list	no. of 6s in the list.

We are going to find a self-counting list of ten numbers.

Fig. 1.1



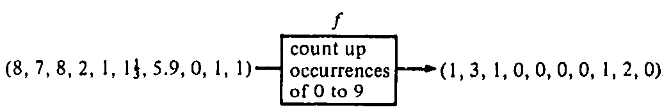
Field of diameter 1
rope of length x

So let $\mathbb{R}^{10}$ be the set of points with ten real coordinates (or 'lists of ten real numbers'). We shall find a member $\mathbf{x}$ of $\mathbb{R}^{10}$ with

$$\mathbf{x}=(x_0, \quad x_1, \quad x_2, \quad x_3, \quad x_4, \quad x_5, \quad x_6, \quad x_7, \quad x_8, \quad x_9)$$

x_0	x_1	x_2
||	||	
no. of 0s in the list	no. of 1s in the list	etc.

In order to employ an iterative technique we will use the process whereby given any member of $\mathbb{R}^{10}$ we count up the occurrences of the numbers 0 to 9 as its coordinates. For example,



Repeatedly applying the process f to give a sequence of points of $\mathbb{R}^{10}$ yields, for the above starting point,

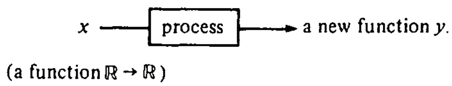
$$\begin{aligned}
\mathbf{x}_1 &= (8, 7, 8, 2, 1, 1\tfrac{1}{3}, 5.9, 0, 1, 1),\\
\mathbf{x}_2 &= f(\mathbf{x}_1) = (1, 3, 1, 0, 0, 0, 0, 1, 2, 0),\\
\mathbf{x}_3 &= f(\mathbf{x}_2) = (5, 3, 1, 1, 0, 0, 0, 0, 0, 0),\\
\mathbf{x}_4 &= f(\mathbf{x}_3) = (6, 2, 0, 1, 0, 1, 0, 0, 0, 0),\\
\mathbf{x}_5 &= f(\mathbf{x}_4) = (6, 2, 1, 0, 0, 0, 1, 0, 0, 0),\\
\mathbf{x}_6 &= f(\mathbf{x}_5) = (6, 2, 1, 0, 0, 0, 1, 0, 0, 0).
\end{aligned}$$

So already the process f is repeating itself and we have actually reached a point

$$\mathbf{x} = (6, 2, 1, 0, 0, 0, 1, 0, 0, 0)$$

for which $f(\mathbf{x}) = \mathbf{x}$; i.e. $\mathbf{x}$ is a self-counting list. Try this process for yourself with different starting points (and with lists of different lengths).

The above example illustrated that our iterations need not only deal with single numbers. Our next few examples will concern iteration with functions and this will have far-reaching consequences. To start with, imagine that we have a process which takes in some function x (e.g. $x(t) = t^2$ for $t \in \mathbb{R}$) and, after processing, gives out a new function y (e.g. $y(t) = t^2 + 1$ for $t \in \mathbb{R}$). In general



Since y is a function, to define it we must stipulate what $y(t)$ is in terms of x's values which we know all about. For example, our process could start with x and produce y with, for each $t \in \mathbb{R}$,

$$y(t) = 1 + \int_0^t u^2 x(u)\, du.$$

So if x is the zero function, then $y(t) = 1$ for all t; and if $x(t) = \sin(t^3)$, then

$$\begin{aligned} y(t) &= 1 + \int_0^t u^2 \sin(u^3)\, du = 1 + [-\tfrac{1}{3}\cos(u^3)]_0^t \\ &= \tfrac{4}{3} - \tfrac{1}{3}\cos(t^3). \end{aligned}$$

This process can be applied to any continuous function (so that the integral is defined) to produce another continuous function. So if X is the set of continuous functions (from $\mathbb{R}$ to $\mathbb{R}$, say) then the process takes an $x \in X$ and creates a $y \in X$: the process is simply a function $f: X \to X$. This is no different in principle from the examples in the previous section except that the elements of X are themselves functions. To summarise, the above process is a function $f: X \to X$ where $x \in X$ is taken to $y = f(x)$ given by

$$(f(x))(t) = y(t) = 1 + \int_0^t u^2 x(u)\, du.$$

Exercise 6 Let x be the function given by $x(t) = 1$. Evaluate the function $f(x)$; i.e. find $(f(x))(t)$ in terms of t, where f is as stated above.

In Exercise 6 $f(x)$ turned out to be the function given by $(f(x))(t) = 1 + \frac{1}{3}t^3$.

Exercise 7 Let x_1 be the function given by $x_1(t) = 1$. Let $f(x_1) = x_2$ where f is as above. Then x_2 is the function given by $x_2(t) = 1 + \frac{1}{3}t^3$. Let $f(x_2) = x_3$: evaluate $x_3(t)$. Let $f(x_3) = x_4$: evaluate $x_4(t)$.

We learnt in the previous section of the possible interest of reapplying a function and investigating the behaviour of the resulting sequence. If we consider our new f defined as above and start with the function

$x_1(t)=1$, then reapplying f gives the sequence

$$x_1(t)=1,\ x_2(t)=1+\frac{t^3}{3},\ x_3(t)=1+\frac{t^3}{3}+\frac{t^6}{9\times 2!},$$
$$x_4(t)=1+\frac{t^3}{3}+\frac{t^6}{9\times 2!}+\frac{t^9}{27\times 3!},\ldots$$

As yet, we have no formal way of testing whether a sequence of functions settles down, but you may have already spotted a pattern in the above sequence and guessed that it is settling down to the function given by

$$x(t)=1+\frac{t^3}{3}+\frac{t^6}{3^2\times 2!}+\frac{t^9}{3^3\times n!}+\cdots+\frac{t^{3n}}{3^n\times n!}+\cdots=e^{t^3/3}.$$

Our work in the previous section might lead us to suspect that this function is a fixed point of f; i.e. $f(x)=x$ or

$$x(t)=(f(x))(t)=1+\int_0^t u^2x(u)\,\mathrm{d}u \quad (\text{all } t\in\mathbb{R}).$$

And indeed this is the case, as can be checked by integration.

So it seems that we have solved the equation in x,

$$x(t)=1+\int_0^t u^2x(u)\,\mathrm{d}u \quad (t\in\mathbb{R}),$$

by starting with any function x_1 and reapplying f to give a sequence of functions: their limit was the required root.

At the moment much of this is informal: we do not really know what it means to say that a sequence of functions converges. (Nor have we seen the significance of equations like the integral equation above.) It seems that the general principle of iteration with a function f might be useful in many situations apart from solving numerical equations. But we must develop the theory for coping with functions $f\colon X\to X$ and for determining in the most general situations whether the sequence

$$x_1,\quad x_2=f(x_1),\quad x_3=f(x_2),\quad \ldots$$

converges. Before doing so, we conclude this section with one further ‘functional’ example of the above type.

> ***Exercise 8*** Let X be the set of continuous real functions defined on the open interval $]-1,1[$. (We shall always use the notation $]a,b[$ for the open interval $\{x\in\mathbb{R}: a<x<b\}$.) For $x\in X$ let $f(x)\in X$ be defined by