2

Solutions of Equations in One Variable

CONTENTS

"nobreak

2.1 Introduction

A problem of great importance in science and engineering is that of determining the roots of an equation in the form

$$f(x) = 0.$$

Definition 4. *A solution x^* of $f(x) = 0$ is a zero of multiplicity m of $f(x)$ if for $x \neq x^*$, we can write*

$$f(x) = (x - x^*)^m g(x),$$

where m is positive integer, and $g(x^*) \neq 0$. When $m = 1$, it is said simple root of $f(x) = 0$ or simple zero of $f(x)$.

Why do we consider the approximations of the roots of $f(x) = 0$? For example, a polynomial equation of the form

$$P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n = 0$$

is called an algebraic equation. As we know, when $n > 5$, it is not possible to solve explicitly for the roots of this equation. Numerical methods will be discussed in this chapter.

There are two steps in finding the roots of the nonlinear equations:

Step 1 Confirm the interval which the roots of the equations belong to;

Step 2 Use numerical methods to obtain the approximation to the roots.

For the first step, there are three methods to confirm the intervals of the roots:

(1) Plot the figure of the function $f(x)$. For example,

$$x^2 - \sin x - 1 = 0,$$

we can plot the curves of $y = x^2$ and $y = 1 + \sin x$ to determine the intervals which can be seen in Fig. 2.1.

(2) Use some theorems in calculus, such as Intermediate Value Theorem.

(3) Set a fixed step to search the interval of the root.

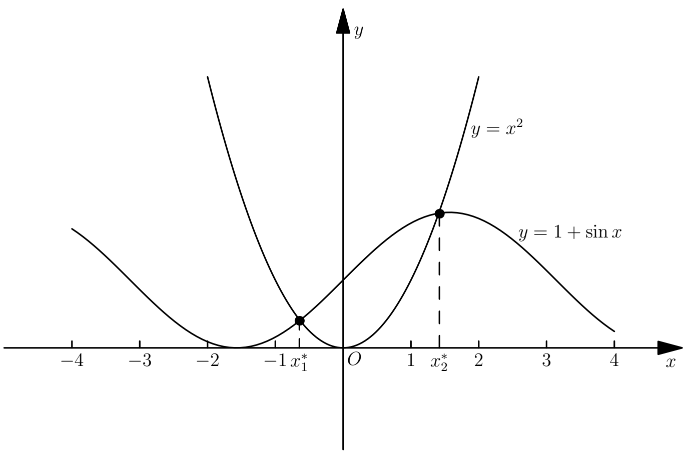

FIGURE 2.1
The plots of $y = x^2$ and $y = 1 + \sin x$

2.2 The Bisection Method

The first technique, based on the Intermediate Value Theorem, is called the **Bisection**, or **Binary-search, method**.

Suppose f is a continuous function defined on the interval $[a,b]$, with $f(a)$ and $f(b)$ of opposite sign. The Intermediate Value Theorem implies that a number x^* exists in (a,b) with $f(x^*)=0$.

Set $a_0=a,\ b_0=b$ and let $x_0=\frac{1}{2}(a_0+b_0)$. In $[a_0,x_0]$ and $[x_0,b_0]$:

1. If $f(x_0)=0$, then $x^*=x_0$ and we are done.
2. If $f(x_0)\neq 0$, there is $f(a_0)f(x_0)<0$ or $f(x_0)f(b_0)<0$
   (a) if $f(a_0)f(x_0)<0$, set $a_1=a_0,\ b_1=x_0$;
   (b) if $f(x_0)f(b_0)<0$, set $a_1=x_0,\ b_1=b_0$.
3. Reapply the process to the interval $[a_1,b_1]$.
4. Generate a sequence

$$[a_0,b_0]\supset[a_1,b_1]\supset[a_2,b_2]\supset\cdots\supset[a_k,b_k]\supset\cdots,$$

and

(a) $b_k-a_k=\frac{1}{2}(b_{k-1}-a_{k-1})=\frac{1}{2^k}(b-a)$;
(b) $f(a_k)f(b_k)<0$.

When b_k-a_k is very small, the number $x_k=\frac{1}{2}(b_k-a_k)$ is the approximation of the root of $f(x)=0$. And it follows that

$$|x^*-x_k|\leq\frac{1}{2}(b_k-a_k)\leq\frac{1}{2^{k+1}}(b-a). \tag{2.1}$$

For given precision ε, we can choose a bigger k with

$$\frac{1}{2^{k+1}}(b-a)\leq\varepsilon,$$

i.e.

$$|x^*-x_k|\leq\varepsilon.$$

Exercise 2.1. *Give equation* $f(x)=x^3+4x^2-10=0$.

1. Show the equation has only one root in $[1,1.5]$.

2. Determine the number of iterations necessary to solve the above equation with accuracy $\frac{1}{2}\times 10^{-2}$ *using* $a=1$ *and* $b=1.5$.

3. Use the Bisection method to determine an approximation to the root to 3 significant figures.

Solution

1. We see that $f(1)=-5$, $f(1.5)=2.375$. When $x\in[1,1.5]$, $f'(x)=3x^2+8x>0$. The Intermediate Value Theorem implies that a simple zero of $f(x)$ in $[1,1.5]$.

TABLE 2.1
Computing results by the Bisection Method

k	a_k (sign of $f(a_k)$)	x_k (sign of $f(x_k)$)	b_k (sign of $f(b_k)$)
0	1(-)	1.25(-)	1.5(+)
1	1.25(-)	1.375(+)	1.5(+)
2	1.25(-)	1.3125(-)	1.375(+)
3	1.3125(-)	1.34375(-)	1.375(+)
4	1.34375(-)	1.359375(-)	1.375(+)
5	1.359375(-)	1.3671875(+)	1.375(+)
6	1.359375(-)	1.36328125(-)	1.3671875(+)

2. Let $a = 1,\ b = 1.5$, and $\varepsilon = \frac{1}{2} \times 10^{-2}$. Since

$$\frac{b-a}{2^{k+1}} \le \varepsilon,$$

we have

$$k \ge \frac{2}{\lg 2} - 1 = 5.64.$$

Hence, 6 iterations will ensure an approximation accurate to within $\frac{1}{2} \times 10^{-2}$.

3. Computing results are listed in Table 2.1. The approximation $x_6 = 1.36328125$ is with 3 significant figures.

2.3 Fixed-Point Iteration

2.3.1 Fixed-Point Iteration

A fixed point for a function is a number at which the value of the function does not change when the function is applied.

Definition 5. *The number p is a fixed point for a given function g if $g(p) = p$.*

In this section we consider the problem of finding solutions to fixed-point problems and the connection between the fixed-point problems and the root-finding problems we wish to solve. Root-finding problems and fixed-point problems are equivalent classes in the following sense:

- Given a root-finding problem $f(x^*) = 0$, we can define functions φ with a fixed point at x^* in a number of ways, for example, as

$$\varphi(x) = x - f(x) \quad \text{or as} \quad \varphi(x) = x + 3f(x)$$

- Conversely, if the function φ has a fixed point at x^*, then the function defined by

$$f(x) = x - \varphi(x)$$

has a zero at x^*.

Although the problems we wish to solve are in the root-finding form, the fixed-point form is easier to analyze, and certain fixed-point choices lead to very powerful root-finding techniques.

We first need to become comfortable with this new type of problem, and to decide when a function has a fixed point and how the fixed points can be approximated to within a specified accuracy.

For any initial approximation $x_0 \in [a, b]$, we can get the iteration

$$x_{k+1} = \varphi(x_k), \quad k = 0, 1, 2, \cdots . \tag{2.2}$$

and generate the sequence $\{x_k\}_{k=0}^{\infty}$. If the sequence $\{x_k\}_{k=0}^{\infty}$ converges to $\tilde{x}$ and $\varphi(x)$ is continuous near $\tilde{x}$, then

$$\tilde{x} = \lim_{k \to \infty} x_k = \lim_{k \to \infty} \varphi(x_{k-1}) = \varphi(\tilde{x}).$$

So we get $f(\tilde{x}) = 0$, i.e.

$$\lim_{k \to \infty} x_k = x^*.$$

The iteration (2.2) is called **fixed-point iteration** or **functional iteration**. $\varphi(x)$ is the iterative function and $\{x_k\}_{k=0}^{\infty}$ is the sequence. If for any $x_0 \in [a, b]$ the sequence is convergent, then the iteration (2.2) is **convergent**. $e_k = x^* - x_k$ is the iteration error. If the sequence does not converge on $[a, b]$ when $x_0 \neq x^*$, the iteration is **divergent**.

Exercise 2.2. *Find the root of the equation $x^3 - x - 1 = 0$ around $x_0 = 1.5$ by fixed-point iteration.*

Mehtod 1 We rewrite the equation as: $x = x^3 - 1$. Let $\varphi_1(x) = x^3 - 1$, and

$$x_{k+1} = x_k^3 - 1, \quad k = 0, 1, 2, \cdots$$

Taking $x_0 = 1.5$, the computing results are list as follows:

k	0	1	2	3	$\cdots$
x_k	1.5	2.375	12.396	1903.779	$\cdots$

The sequence generated by Method 1 does not converge to a finite number.

Method 2 The equation can be written as: $x = \sqrt[3]{x+1}$. Let $\varphi_2(x) = \sqrt[3]{x+1}$, then

$$x_{k+1} = \sqrt[3]{x_k+1}, \quad k = 0, 1, 2, \cdots$$

Also taking $x_0 = 1.5$, the computing results are list in the table:

k	0	1	2	3	$\cdots$	7	8
x_k	1.5	1.35721	1.33086	1.32588	$\cdots$	1.32472	1.32472

We can see that the above iteration is convergent. The procedure is illustrated in Figure 2.2 and Figure 2.3.

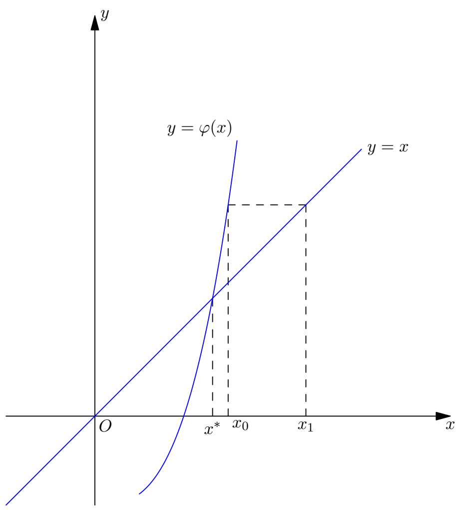

FIGURE 2.2
Method 1

2.3.2 Convergence

Although the various functions we have given are fixed-point problems for the same root-finding problem, they differ vastly as techniques for approximating the solution to the root-finding problem. Their purpose is to illustrate what needs to be answered:

• How can we find a fixed-point problem that produces a sequence that reliably and rapidly converges to a solution to a given root-finding problem?

The following theorem gives us some clues concerning the paths we should pursue and, perhaps more importantly, some we should reject.

Theorem 2. *Suppose* $\varphi \in C^1[a,b]$.

1. $\varphi(x) \in [a,b]$ *for all* $x \in [a,b]$;

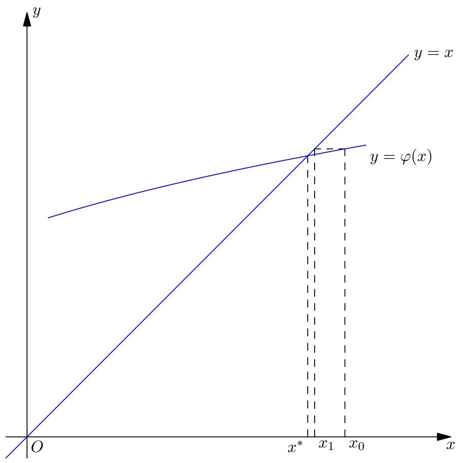

FIGURE 2.3
Method 2

2. There exists positive constant, $|\varphi'(x)| \leq L < 1$ for all $x \in [a,b]$.

Then

1. There exists unique $x^* \in [a,b]$ such that $x^* = \varphi(x^*)$.

2. For any initial data $x_0 \in [a,b]$, the fixed-point iteration (2.2) is convergent and $\lim\limits_{k\to\infty} x_k = x^*$;

3.

$$|x^* - x_k| \leq \frac{L}{1-L}|x_k - x_{k-1}|, \quad k = 1, 2, 3 \cdots ; \tag{2.3}$$

4.

$$|x^* - x_k| \leq \frac{L^k}{1-L}|x_1 - x_0|, \quad k = 1, 2, 3 \cdots ; \tag{2.4}$$

5.

$$\lim_{k\to\infty} \frac{x^* - x_{k+1}}{x^* - x_k} = \varphi'(x^*). \tag{2.5}$$

Proof. (1) Let $g(x) = x - \varphi(x)$ be continuous on $[a,b]$ with

$$g(a) = a - \varphi(a) \leqslant 0, \quad g(b) = b - \varphi(b) \geqslant 0.$$

The Intermediate Value Theorem implies that there exists $x^* \in [a,b]$ for which $g(x^*) = 0$, i.e. $x^* = \varphi(x^*)$.

Suppose x_1^* and x_2^* are both fixed points in $[a,b]$, i.e.

$$x_1^* = \varphi(x_1^*), \quad x_2^* = \varphi(x_2^*).$$

If $x_1^* \neq x_2^*$, then the Mean Value Theorem implies that a number ξ between x_1* and x_2^*.

Thus,

$$\begin{aligned}|x_1^* - x_2^*| &= |\varphi(x_1^*) - \varphi(x_2^*)| \\ &= |\varphi'(\xi)|\,|x_1^* - x_2^*| \\ &\leqslant L\,|x_1^* - x_2^*|\end{aligned}$$

where ξ between x_1^* and x_2^*. Because $L < 1$, it is a contradiction, this contradiction must come from the only supposition $x_1^* \neq x_2^*$. Hence $x_1^* = x_2^*$, and the fixed point in $[a,b]$ is unique.

(2) Since φ maps $[a,b]$ into itself, the sequence $\{x_k\}_{k=0}^{\infty}$ is defined for all $k \geq 0$, and $x_k \in [a,b]$ for all k. Using the fact that $|\varphi'(x)| \leq L$ and the Mean Value Theorem, we have, for each k

$$x^* - x_{k+1} = \varphi(x^*) - \varphi(x_k) = \varphi'(\xi_k)(x^* - x_k), \tag{2.6}$$

where ξ_k between x_k and x^*. Then

$$|x^* - x_{k+1}| \leqslant L\,|x^* - x_k| \quad (k = 0, 1, 2, \cdots). \tag{2.7}$$

Applying this inequality inductively gives

$$|x^* - x_k| \leqslant L\,|x^* - x_{k-1}| \leqslant L^2\,|x^* - x_{k-2}| \leqslant \cdots \leqslant L^k\,|x^* - x_0|.$$

Since $L < 1$, we have

$$\lim_{k\to\infty} x_k = x^*.$$

Hence, $\{x_k\}_{k=0}^{\infty}$ converges to x^*.

(3) From (2.7),

$$\begin{aligned}|x^* - x_k| &= |x^* - x_{k+1} + x_{k+1} - x_k| \leqslant |x^* - x_{k+1}| + |x_{k+1} - x_k| \\ &\leqslant L\,|x^* - x_k| + |x_{k+1} - x_k|,\end{aligned}$$

then

$$|x^* - x_k| \leqslant \frac{1}{1-L}\,|x_{k+1} - x_k|. \tag{2.8}$$

Noticing that

$$\begin{aligned}|x_{k+1} - x_k| &= |\varphi(x_k) - \varphi(x_{k-1})| = |\varphi'(\eta_k)(x_k - x_{k-1})| \\ &\leqslant L\,|x_k - x_{k-1}|,\end{aligned} \tag{2.9}$$

we take (2.9) to (2.8) and obtain

$$|x^* - x_k| \leqslant \frac{L}{1-L}\,|x_k - x_{k-1}| \quad (k = 1, 2, 3, \cdots).$$

(4) From (2.9), we have

$$|x_{k+1} - x_k| \leqslant L^k |x_1 - x_0| .$$

Taking the above equation to the right of (2.8), there is

$$|x^* - x_k| \leqslant \frac{L^k}{1-L} |x_1 - x_0| \quad (k = 1, 2, 3, \cdots).$$

(5) Since (2.6),

$$\frac{x^* - x_{k+1}}{x^* - x_k} = \varphi'(\xi_k) \quad (k = 0, 1, 2, \cdots).$$

For $\lim_{k\to\infty} \xi_k = x^*$, we obtain

$$\lim_{k\to\infty} \frac{x^* - x_{k+1}}{x^* - x_k} = \varphi'(x^*) .$$

□

Both inequalities in the theorem relate the rate at which $\{x_k\}_{k=0}^{\infty}$ converges to the bound L on the first derivative. The rate of convergence depends on the factor L^k. The smaller the value of L, the faster the convergence, which may be very slow if L is close to 1.

Theorem 3. *Suppose $f(x) = 0$ has zeros in $[a,b]$. If $|\varphi'(x)| \geq 1$ for any $x \in [a,b]$, the iteration (2.2) is divergent for any $x_0 \in [a,b]$ and $x_0 \neq x^*$.*

Proof. For $x_0 \in [a,b]$ and $x_0 \neq x^*$, we have

$$|x^* - x_1| = |\varphi(x^*) - \varphi(x_0)| = |\varphi'(\xi_0)(x^* - x_0)| \geqslant |x^* - x_0| > 0.$$

If $x_1 \in [a,b]$, there is

$$\begin{aligned} |x^* - x_2| &= |\varphi(x^*) - \varphi(x_1)| = |\varphi'(\xi_1)(x^* - x_1)| \\ &\geqslant |x^* - x_1| \geqslant |x^* - x_0| . \end{aligned}$$

Applying this inequality inductively gives $x_k \notin [a,b]$, or $|x^* - x_k| \geqslant |x^* - x_0|$. Hence the sequence is divergent.

□

Exercise 2.3. *The three-degree polynomial $f(x) = x^3 + 4x^2 - 10 = 0$ has a root x^* in $[1, 1.5]$:*

1. Attempt to analyze the convergency of the following three iterations.

$$(1)\ x_{k+1} = 10 + x_k - 4x_k^2 - x_k^3; \tag{2.10}$$

$$(2)\ x_{k+1} = \frac{1}{2}\sqrt{10 - x_k^3}; \tag{2.11}$$

$$(3)\ x_{k+1} = \sqrt{\frac{10}{x_k + 4}}. \tag{2.12}$$

2. Find the real root to 4 significant figures by the iteration with higher convergency order.

Solution

1. (1) The iterative function is

$$\varphi(x) = 10 + x - 4x^2 - x^3, \tag{2.13}$$
$$\varphi'(x) = 1 - 8x - 3x^2. \tag{2.14}$$

When $x \in [1, 1.5]$, $|\varphi'(x)| \geq 10 > 1$. The iteration is not convergent.
(2) Let $\varphi(x) = \frac{1}{2}\sqrt{10 - x^3}$. If $x \in [1, 1.5]$,

$$|\varphi'(x)| = \left|\frac{3}{4}\frac{x^2}{\sqrt{10 - x^3}}\right| \nearrow, \quad \Longrightarrow \tag{2.15}$$
$$|\varphi(x)| \leq |\varphi(1.5)| = 0.6556 < 1. \tag{2.16}$$

If $x \in [1, 1.5]$,

$$1 < \varphi(1.5) \leq \varphi(x) \leq \varphi(1) = 1.5,$$

So the iteration is convergent.
(3) The iterative function is

$$\varphi(x) = \sqrt{\frac{10}{x+4}}$$

If $x \in [1, 1.5]$, there is

$$\varphi(x) \in [\varphi(1.5), \varphi(1)] = \left[\sqrt{\frac{10}{1.5+4}}, \sqrt{\frac{10}{1+4}}\right]$$
$$= [1.348, 1.414] \subset [1, 1.5].$$

Since

$$\varphi'(x) = -\frac{1}{2}\sqrt{10}(x+4)^{-\frac{3}{2}} < 0, \quad \varphi''(x) = \frac{3}{4}\sqrt{10}(x+4)^{-\frac{5}{2}} > 0,$$

we have

$$|\varphi'(x)| \leqslant |\varphi'(1)| = \frac{1}{2}\sqrt{10}(1+4)^{-\frac{3}{2}} = \frac{\sqrt{2}}{10} = 0.1414$$

for $x \in [1, 1.5]$. Thus, the iteration is convergent.

2. The convergence of the third iteration is faster than the others. Computing results are listed as follows

k	0	1	2	3	4
x_k	1.25	1.38013	1.36334	1.36547	1.36512

So $x^* = 1.365$.

Exercise 2.4. *Let $x^2 + \ln x - 2 = 0$. Use fixed-point iteration to find solution to 4 significant figures and show the convergence of the iteration.*

Solution

1. Let $f(x) = x^2 + \ln x - 2$. $f(1) = 1 - 2 < 0, f(2) = 4 + \ln 2 - 2 > 0$, $f'(x) = 2x + \frac{1}{x} > 0$. $f(x) = 0$ has only real root $x^* \in (1, 2)$.
2. The fixed-point iteration is :

$$\begin{aligned} &x_{k+1} = \sqrt{2 - \ln x_k}, \quad k = 0, 1 \cdots, \\ &x_0 = 1.3. \end{aligned}$$

Let $\varphi(x) = \sqrt{2 - \ln x}$.

(a) $\forall x \in [1, 2]$, $|\varphi'(x)| = \frac{1}{2x\sqrt{2-\ln x}} \le \frac{1}{2\sqrt{2-\ln 2}} < 1$.

(b) If $x \in [1, 2]$, $1 < \sqrt{2 - \ln 2} \le \varphi(x) \le \sqrt{2} < 2$.

From Th. 2, the iteration is convergent.

The computing results of the fixed-point iteration are as follows:
$x_1 = 1.318194$, $x_2 = 1.312911$, $x_3 = 1.314440$, $x_4 = 1.313997$, $|x_4 - x_3| < \frac{1}{2} \times 10^{-3}$, $x^* \approx 1.313997$.

2.3.3 Order of Convergence

Definition 6. *Suppose $\{x_k\}_{k=0}^{\infty}$ is a sequence that converges to x^*, with $x_k \neq x^*$ for all k, and denote $e_k = x^* - x_k$. If positive constant $p \ge 1$ and non-zero (finite) constant C exists with*

$$\lim_{k \to \infty} \frac{e_{k+1}}{e_k^p} = C,$$

then $\{x_k\}_{k=0}^{\infty}$ converges to x^* of order p, with asymptotic error constant C.

An iterative technique of the form $x_k = \varphi(x_{k-1})$ is said to be of order p if the sequence $\{x_k\}_{k=0}^{\infty}$ converges to the solution $x^* = \varphi(x^*)$ of order p.

In general, a sequence with a high order of convergence converges more rapidly than a sequence with a lower order. The asymptotic constant affects the speed of convergence but not to the extent of the order. Two cases of order are given special attention.

(1) If $p = 1$ and $0 < |C| < 1$, the sequence is linearly convergent;

(2) If $p > 1$, the sequence is superlinear convergent, especially, $p = 2$ it is quadratically convergent.

Exercise 2.5. *Suppose that* $\{e_k\}_{k=0}^{\infty}$ *is linearly convergent to 0 with*

$$\frac{e_{k+1}}{e_k} = \frac{1}{2}(k = 0, 1, 2, \cdots)$$

and that $\{\tilde{e_k}\}_{k=0}^{\infty}$ *is quadratically convergent to 0 with the same asymptotic error constant,*

$$\frac{\tilde{e}_{k+1}}{\tilde{e}_k^2} = \frac{1}{2}(k = 0, 1, 2, \cdots).$$

If $e_0 = \tilde{e}_0 = 1$*, try to confirm the numbers with the two sequence converge to 0 within* $\varepsilon = 10^{-16}$*.*

Solution (1) From $\frac{e_{k+1}}{e_k} = \frac{1}{2}(k = 0, 1, 2, \cdots)$ and $e_0 = 1$, we have

$$e_k = \frac{1}{2}e_{k-1} = \cdots = \frac{1}{2^k}e_0 = \frac{1}{2^k}.$$

Since $|e_k| \leqslant 10^{-16}$, this means $\frac{1}{2^k} \leqslant 10^{-16}$ or $2^k \geqslant 10^{16}$. Solve the above equation, we obtain

$$k \lg 2 \geqslant 16 \Rightarrow k \geqslant \frac{16}{\lg 2} = 53.15.$$

The linear convergent sequence is within 10^{-16} of 0 by the 54 iterations.

(2) For $\frac{\tilde{e}_{k+1}}{\tilde{e}_k^2} = \frac{1}{2}(k = 0, 1, 2, \cdots)$ and $\tilde{e}_0 = 1$, there is

$$\begin{aligned}\tilde{e}_k &= \frac{1}{2}\tilde{e}_{k-1}^2 = \frac{1}{2}(\frac{1}{2}\tilde{e}_{k-2}^2)^2 = (\frac{1}{2})^{1+2}(\tilde{e}_{k-2})^{2^2}\\ &= \cdots = (\frac{1}{2})^{1+2+\cdots+2^{k-1}}\tilde{e}_0^{2^k} = (\frac{1}{2})^{2^k-1}\end{aligned}$$

When $|\tilde{e}_k| \leqslant 10^{-16}$, i.e

$$\left(\frac{1}{2}\right)^{2^k-1} \leqslant 10^{-16} \quad \text{or} \quad 2^{2^k-1} \geqslant 10^{16}.$$

We can get

$$\left(2^k - 1\right)\lg 2 \geqslant 16 \Rightarrow 2^k - 1 \geqslant \frac{16}{\lg 2} = 53.15 \Rightarrow 2^k \geqslant 54.15.$$

Then

$$k \geqslant \frac{\lg 54.15}{\lg 2} = 5.76.$$

At least 6 iterations are needed to ensure this accuracy for the quadratically convergent sequence. In fact,

$$\tilde{e}_0 = 1, \quad \tilde{e}_1 = 0.5, \quad \tilde{e}_2 = 0.125, \quad \tilde{e}_3 = 0.78125 \times 10^{-2},$$

$$\tilde{e}_4 = 0.30517578 \times 10^{-4}, \quad \tilde{e}_5 = 0.46566128 \times 10^{-9},$$

$$\tilde{e}_6 = 0.10842021 \times 10^{-18}.$$

Definition 7. *Let x^* be a solution of the equation $x = \varphi(x)$. Suppose there exists a δ for $x_0 \in S = \{x||x-x^*| \leq \delta\}$, the sequence defined by $x_{k+1} = \varphi(x_k)$, when $k \geq 1$, converges to x^*. We call it is* **local convergent**.

Theorem 4. *Suppose x^* is zero of $x = \varphi(x)$, and there exists a δ for $\varphi(x) \in C^1[x^* - \delta, x^* + \delta]$.*

1. *The sequence of the iteration is local convergent for $|\varphi'(x^*)| < 1$;*
2. *When $|\varphi'(x^*)| > 1$, the sequence of the iteration is not convergent.*

Proof. 1° Suppose $|\varphi'(x^*)| < 1$. Since $|\varphi'(x)|$ is continuous around x^*, taking $\varepsilon = \frac{1}{2}(1 - |\varphi'(x^*)|)$, there exists a smaller number $\delta(\ \delta \leqslant \tilde{\delta})$. For $x \in [x^* - \delta, x^* + \delta]$, we have

$$||\varphi'(x)| - |\varphi'(x^*)|| \leqslant \frac{1}{2}(1 - |\varphi'(x^*)|).$$

From the above equation, there is

$$\varphi'(x)| \leqslant |\varphi'(x^*)| + \frac{1}{2}(1 - |\varphi'(x^*)|) = \frac{1}{2}(1 + |\varphi'(x^*)|) < 1.$$

For the given δ above and all $x \in [x^* - \delta, x^* + \delta]$,

$$|\varphi(x) - x^*| = |\varphi(x) - \varphi(x^*)| = |\varphi'(\xi)(x - x^*)| \leqslant |x - x^*| \leqslant \delta.$$

Hence, the iteration is local convergent from Theorem 5.

2° Suppose $|\varphi'(x^*)| > 1$. Then $|\varphi'(x)| > 1$ for $x \in [x^* - \delta, x^* + \delta]$. Thus, the iteration is divergent.

□

Theorem 5. *Suppose there is a $\delta > 0$ such that $\varphi(x) \in C^p[x^* - \delta, x^* + \delta], p \geq 1$, and*

$$\varphi^{(k)}(x^*) = 0, \quad k = 1, 2, \cdots, p-1, \tag{2.17}$$

$$\varphi^{(p)}(x^*) \neq 0. \tag{2.18}$$

Then the sequence defined by $x_k = \varphi(x_{k-1})$ is local convergent and

$$\lim_{k\to\infty} \frac{x^* - x_{k+1}}{(x^* - x_k)^p} = (-1)^{p-1} \frac{\varphi^{(p)}(x^*)}{p!}. \tag{2.19}$$

If $p = 1$, $|\varphi'(x^*)| < 1$ is required.

Proof. It is clear that the iteration (2.2) is local convergent according to Theorem 4. Applying Taylor's theorem, we have

$$\begin{aligned}
x_{k+1} =&\varphi\left(x_k\right) \\
=&\varphi\left(x^*\right)+\varphi'\left(x^*\right)\left(x_k-x^*\right)+\cdots+\frac{\varphi^{(p-1)}\left(x^*\right)}{(p-1)!}\left(x_k-x^*\right)^{p-1} \\
&+\frac{\varphi^{(p)}\left(x^*+\theta\left(x_k-x^*\right)\right)}{p!}\left(x_k-x^*\right)^p \\
=&x^*+\frac{\varphi^{(p)}\left(x^*+\theta\left(x_k-x^*\right)\right)}{p!}\left(x_k-x^*\right)^p
\end{aligned}$$

where $0<\theta<1$. Then

$$\frac{x_{k+1}-x^*}{\left(x_k-x^*\right)^p}=\frac{\varphi^p\left(x^*+\theta\left(x_k-x^*\right)\right)}{p!}$$

or

$$\frac{x^*-x_{k+1}}{\left(x^*-x_k\right)^p}=(-1)^{p-1}\frac{\varphi^{(p)}\left(x^*+\theta\left(x_k-x^*\right)\right)}{p!}.$$

□

2.3.4 Aitken Method

From Theorem 5, we know that the order of convergence of the fixed-point iterration

$$x_{k+1}=\varphi(x_k), \quad k=0,1,2,\cdots \tag{2.20}$$

is related with the function φ. In general, it is linearly convergent and we try to construct a new iterative method with

(1) the equations $x=\Phi(x)$ and $x=\varphi(x)$ have the same root x^*;

(2) the sequence generated by

$$x_{k+1}=\Phi\left(x_k\right) \quad (k=0,1,2,\cdots)$$

converge more rapidly to x^* than does (2.20).

The technique called **Aitken method** can be used to accelerate the convergence of an iterative method.

Suppose (2.20) is convergent. From Theorem 5, we have

$$\lim_{k\to\infty}\frac{x_{k+1}-x^*}{x_k-x^*}=\varphi'\left(x^*\right).$$

For larger k, there is

$$\frac{x_{k+2}-x^*}{x_{k+1}-x^*}\approx\frac{x_{k+1}-x^*}{x_k-x^*}.$$

The root can be solved as the follows

$$x^* \approx \frac{x_k x_{k+2} - x_{k+1}^2}{x_k - 2x_{k+1} + x_{k+2}}.$$

Taking $x_{k+1} = \varphi(x_k)$ and $x_{k+2} = \varphi(\varphi(x_k))$ to the upper equation, we have

$$x^* \approx \frac{x_k \varphi(\varphi(x_k)) - \varphi^2(x_k)}{x_k - 2\varphi(x_k) + \varphi(\varphi(x_k))}.$$

The term in right hand is considered as a new approximation to x_{k+1}, we obtain a new iterative method

$$x_{k+1} = \Phi(x_k) \quad (k = 0, 1, 2, \cdots), \tag{2.21}$$

where

$$\Phi(x) = \frac{x\varphi(\varphi(x)) - \varphi^2(x)}{x - 2\varphi(x) + \varphi(\varphi(x))}.$$

Theorem 6. *Suppose $x = \varphi(x)$ has the root x^*, and second-order continuous around x^*. If the iterative method* (2.20) *is linearly convergent, the method* (2.21) *is of quadratic convergence.*

Proof. As we know

$$\varphi'(x^*) \neq 0, \quad |\varphi'(x^*)| < 1.$$

From Taylor's theorem, we have

$$\varphi(x^* + h) = \varphi(x^*) + h\varphi'(x^*) + \frac{1}{2}h^2\varphi''(x^* + \theta h)$$

where $0 < \theta < 1$ and h is smaller. Denote

$$A = \varphi'(x^*), \quad B(h) = \frac{1}{2}\varphi''(x^* + \theta h).$$

We have

$$\varphi(x^* + h) = x^* + Ah + B(h)h^2$$

and

$$A \neq 0, \quad |A| < 1, \quad B(0) \neq 0.$$

Let

$$\delta = Ah + B(h)h^2.$$

Then

$$\varphi(x^* + h) = x^* + \delta.$$

From the above four equations, we have

$$\begin{aligned}
\Phi(x^*+h) &= \frac{(x^*+h)\,\varphi(\varphi(x^*+h)) - \varphi^2(x^*+h)}{(x^*+h) - 2\varphi(x^*+h) + \varphi(\varphi(x^*+h))} \\
&= \frac{(x^*+h)\,\varphi(x^*+\delta) - (x^*+\delta)^2}{(x^*+h) - 2(x^*+\delta) + \varphi(x^*+\delta)} \\
&= \frac{(x^*+h)\left(x^*+A\delta+B(\delta)\delta^2\right) - (x^*+\delta)^2}{(x^*+h) - 2(x^*+\delta) + (x^*+A\delta+B(\delta)\delta^2)} \\
&= \frac{\left[h+A\delta+B(\delta)\delta^2-2\delta\right]x^* + h\left[A\delta+B(\delta)\delta^2\right] - \delta^2}{h-2\delta+A\delta+B(\delta)\delta^2} \\
&= x^* + \frac{h\left[A\delta+B(\delta)\delta^2\right]-\delta^2}{h+(A-2)\delta+B(\delta)\delta^2} \\
&= x^* + h^2\frac{A^2B(\delta)-AB(h)+O(h)}{(1-A)^2+O(h)}.
\end{aligned}$$

So

$$\begin{aligned}
&\lim_{h\to x^*}\Phi(x) = \lim_{h\to 0}\Phi(x^*+h) = x^* \equiv \Phi(x^*) \\
&\lim_{x\to x^*}\Phi'(x) = \lim_{h\to 0}\frac{\Phi(x^*+h)-\Phi(x^*)}{h} \\
&= \lim_{k\to 0}\left[h\frac{A^2B(\delta)-AB(h)+O(h)}{(1-A)^2+O(h)}\right] = 0.
\end{aligned}$$

The iterative method (2.21) is local convergent.

$$\begin{aligned}
\lim_{k\to\infty}\frac{x^*-x_{k+1}}{(x^*-x_k)^2} &= \lim_{k\to\infty}\frac{x^*-\Phi(x_k)}{(x^*-x_k)^2} = \lim_{k\to\infty}\frac{x^*-\Phi(x^*+(x_k-x^*))}{(x^*-x_k)^2} \\
&= \lim_{h\to 0}\frac{x^*-\Phi(x^*+h)}{h^2} = \lim_{h\to 0}\frac{-A^2B(\delta)+AB(h)+O(h)}{(1-A)^2+O(h)} \\
&= \frac{AB(0)}{1-A} = \frac{1}{2}\frac{\varphi'(x^*)\,\varphi''(x^*)}{1-\varphi'(x^*)} \neq 0.
\end{aligned}$$

From Theorem 5, the iterative method (2.21) is of quadratic convergence. □

If the iteration (2.2.20) is $p(\geqslant 2)$ order of convergence and $\varphi(x)$ is $(p+1)$nd-order continuous around x^*, then the method (2.21) is $(2p-1)$ order of convergence.

Exercise 2.6. *Find the root of the equation* $3x-\sin x-\cos x=0$ *around 0.5 by Aitken method and the approximation is of 4 significant figures.*

Solution We rewrite the equation as

$$x = \frac{\sqrt{2}}{3}\sin\left(x+\frac{\pi}{4}\right).$$

Then $\varphi(x) = \frac{\sqrt{2}}{3}\sin\left(x+\frac{\pi}{4}\right)$ With $\varphi'(x) = \frac{\sqrt{2}}{3}\cos\left(x+\frac{\pi}{4}\right)$, we have

$$|\varphi'(0.5)| = 0.132719.$$

The iteration

$$\begin{cases} x_{k+1} = \varphi\left(x_k\right) \quad (k = 0, 1, 2, \cdots); \\ x_0 = 0.5 \end{cases}$$

is of linear convergence. The Aitken method is

$$\begin{cases} x_{k+1} = \frac{x_k\varphi(\varphi(x_k)) - [\varphi(x_k)]^2}{\varphi(\varphi(x_k)) - 2\varphi(x_k) + x_k} \quad (k = 0, 1, 2, \cdots); \\ x_0 = 0.5 \end{cases}$$

Then $x_1 = 0.444354$, $x_2 = 0.444236$. The approximation of the root is $x* \approx 0.4442$.

2.4 Newton's Method

Newton's (or the **Newton-Raphson**) method is one of the most powerful and well-known numerical methods for solving a root-finding problem.

Consider the following equation

$$f(x) = 0. \tag{2.22}$$

Suppose x_k is known, and consider the first Taylor polynomial for $f(x)$ expanded about x_k

$$f(x) \approx f(x_k) + f'(x_k)(x - x_k).$$

Then

$$f(x_k) + f'(x_k)(x - x_k) = 0$$

is taken as the approximation to (2.22) by assuming that since $|x^* - x_k|$ is small. Solving for x^* gives

$$x^* \approx x_k - \frac{f(x_k)}{f'(x_k)}. \tag{2.23}$$

This set the stage for Newton's method, which starts with an initial approximation x_0 and generates the sequence x_k by

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, \cdots. \tag{2.24}$$

The equation

$$y = f\left(x_k\right) + f'\left(x_k\right)\left(x - x_k\right) \tag{2.25}$$

is tangent equation through point $(x_k, f(x_k))$. The idea of Newton's method (2.24) is to replace a general equation $f(x) = 0$ by a simpler function, i.e. tangent equation, and do the required computation exactly on the simpler function. Figure 2.4 shows the geometric of the Newton's method.

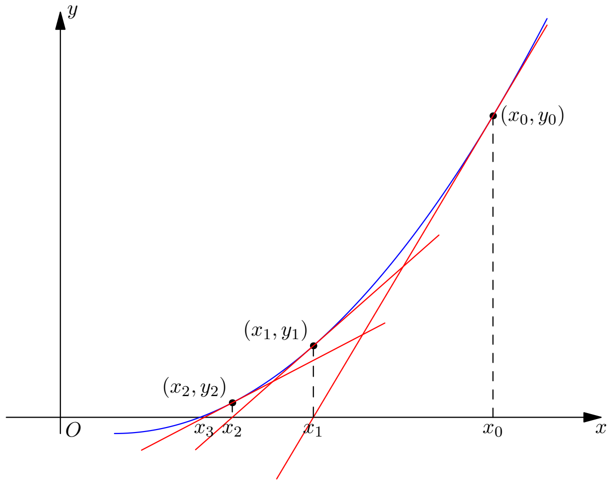

FIGURE 2.4
The geometric of the Newton's method

2.4.1 Local Convergence of Newton's method.

Let $\varphi(x) = x - \frac{f(x)}{f'(x)}$. Suppose x^* is a root of multiplicity m of $f(x) = 0$, then $f(x) = (x - x^*)^m g(x), g(x^*) \neq 0$. Hence

$$
\begin{aligned}
f'(x) &= (x - x^*)^{m-1}(mg(x) + (x - x^*)g'(x)), \\
\varphi(x) &= x - \frac{(x - x^*)g(x)}{mg(x) + (x - x^*)g'(x)}, \\
\varphi'(x^*) &= \lim_{x \to x^*} \frac{\varphi(x) - \varphi(x^*)}{x - x^*} \\
&= \lim_{x \to x^*} \frac{1}{x - x^*}\left(x - \frac{(x - x^*)g(x)}{mg(x) + (x - x^*)g'(x)} - x^*\right) \\
&= 1 - \frac{1}{m}.
\end{aligned}
$$

We conclude that if f is enough smooth

- If $m = 1$, the sequence by Newton's method converges at least quadratically to x^*.
- If $m(m \geq 2)$, the sequence by Newton's method converges linearly to x^*.

Exercise 2.7. *Use Newton's method to find all real solutions to 3 significant figures to the problem*

$$e^x + x - 3 = 0.$$

Solution Let $f(x) = e^x + x - 3$. $f(0) = 1 + 0 - 2 < 0, f(1) = e + 1 - 3 > 0,$

TABLE 2.2
Computing results by Newton's method

k	x_k
0	5.000 000
1	4.682 017
2	4.572 805
3	4.560 161
4	4.560 00

$f'(x) = e^x + 1 > 0$, so there exists only one real root of $f(x) = 0$ in $(0, 1)$ i.e. $x^* \in (0, 1)$.

The Newton's Method is

$$x_{k+1} = x_k - \frac{e^{x_k} + x_k - 3}{e^{x_k} + 1}, \quad k = 0, 1 \cdots,$$
$$x_0 = 0.5.$$

We can obtain $x_1 = 0.8214$, $x_2 = 0.7924$, $x_3 = 0.7921$, $|x_3 - x_2| < \frac{1}{2} \times 10^{-3}$, $x^* \approx 0.7921$.

Exercise 2.8. *Find the roots of the equation*

$$f(x) = (x - 1.56)^3(x - 4.56) = 0$$

by Newton's method.

Solution $x_1^* = 4.56$ is the simple root of $f(x) = 0$ and $x_2^* = 1.56$ is multiple root of multiplicity 3 of $f(x) = 0$. Take initial data $x_0 = 5.000000$ and $x_0 = 2.000000$ to find the roots by Newton's method correspondingly. The computing results are list in Table 2.2 and 2.3. From the table, we can see that the rapid convergence of the Newton's method in the case of a simple zero.

2.4.2 Multiple Roots

Suppose x^* is the root of multiplicity m of $f(x) = 0$.

(1) If m is known, modified Newton's method is

$$x_{k+1} = x_k - m\frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, \cdots.$$

Exercise 2.9. *Find the multiple root of*

$$f(x) = (x - 1.56)^3(x - 4.56) = 0$$

by modified Newton's method.

TABLE 2.3
Computing results by Newton's method

k	x_k	k	x_k
0	2.000000	10	1.567042
1	1.844420	11	1.564692
2	1.246184	12	1.563128
3	1.682723	13	1.562085
4	1.641225	14	1.561390
5	1.613896	15	1.560926
6	1.595821	16	1.560617
7	1.583832	17	1.560412
8	1.575867	18	1.560274
9	1.570569	19	1.560183

Solution The modified Newton's method is

$$x_{k+1} = x_k - 3\frac{f\left(x_k\right)}{f'\left(x_k\right)} \quad (k = 0, 1, 2, \cdots).$$

The sequence generates starting with the initial data $x_0 = 2$ and the computing results are list in the following table.

k	0	1	2	3
x_k	2.000000	1.533260	1.559921	1.560000

(2) If m is unknown and define $u(x) = \frac{f(x)}{f'(x)}$. For x^* is a zero of f of multiplicity m with $f(x) = (x - x^*)^m g(x)$, then

$$\begin{aligned} u(x) &= \frac{(x - x^*)^m g(x)}{m(x - x^*)^{m-1} g(x) + (x - x^*)^m g'(x)} \\ &= (x - x^*)\frac{g(x)}{mg(x) + (x - x^*)g'(x)} \end{aligned}$$

also has a zero at x^*. However, $g(x^*) \neq 0$, so

$$\frac{g(x^*)}{mg(x^*) + (x^* - x^*)g'(x^*)} = \frac{1}{m} \neq 0$$

and x^* is a simple zero of $u(x)$. Newton's method can then be applied to $u(x)$ to give

$$\varphi(x) = x - \frac{u(x)}{u'(x)} = x - \frac{f(x)/f'(x)}{\left\{[f'(x)]^2 - [f(x)]\,[f''(x)]\right\} / [f'(x)]^2}$$

which simplifies to

$$\varphi(x) = x - \frac{f(x)f'(x)}{[f'(x)]^2 - f(x)f''(x)}. \tag{2.26}$$

If φ has the required continuity conditions, functional iteration applied to φ will be quadratically convergent regardless of the multiplicity of the zero of f. Theoretically, the only drawback to this method is the additional calculation of $f''(x)$ and the more laborious procedure of calculating the iterates. In practice, however, multiple roots can cause serious round-off problems because the denominator of (2.26) consists of the difference of two numbers that are both close to 0 .

2.4.3 The Secant Method

Newton's method is an extremely powerful technique, but it has a major weakness: the need to know the value of the derivative of f at each approximation. Frequently, $f'(x)$ is far more difficult and needs more arithmetic operations to calculate than $f(x)$.

To circumvent the problem of the derivative evaluation in Newton's method, we introduce a slight variation. By definition,

$$f'\left(x_k\right)=\lim_{x \rightarrow x_k} \frac{f(x)-f\left(x_k\right)}{x-x_k}$$

If x_{k-1} is close to x_k, then

$$f'\left(x_k\right) \approx \frac{f\left(x_{k-1}\right)-f\left(x_k\right)}{x_{k-1}-x_k}=\frac{f\left(x_k\right)-f\left(x_{k-1}\right)}{x_k-x_{k-1}}$$

Using this approximation for $f'\left(x_k\right)$ in Newton's formula gives

$$x_{k+1}=x_k-\frac{f(x_k)}{f(x_k)-f(x_{k-1})}(x_k-x_{k-1}), \quad k=1,2, \cdots .$$

This technique is called the **Secant method** Starting with the two initial approximations x_0 and x_1, the approximation x_2 is the x -intercept of the line joining $\left(x_0, f\left(x_0\right)\right)$ and $\left(x_1, f\left(x_1\right)\right)$. The approximation x_3 is the x -intercept of the line joining $\left(x_1, f\left(x_1\right)\right)$ and $\left(x_2, f\left(x_2\right)\right)$, and so on. Note that only one function evaluation is needed per step for the Secant method after x_2 has been determined. In contrast, each step of Newton's method requires an evaluation of both the function and its derivative.

2.5 Zeros of Polynomials

“nobreak

2.6 Excercise

1. Use three fixed-point iterations to find the positive root of $x^3 - 5x - 3 = 0$. Show the iterations are convergent or divergent. Find the solutions of 4 significant figures by the rapid convergent method.
2. Use Newton's method to determine all the solutions of 5 significant figures for $f(x) = \mathrm{e}^x - 3x^2 = 0$.
3. Use Newton's method to find an approximation to $\sqrt{a}$ through solving $f(x) = x^n - a = 0$ or $f(x) = 1 - \frac{a}{x^n} = 0$. Try to compute $\lim_{k\to\infty} \left(\sqrt[n]{a} - x_{k+1}\right) / \left(\sqrt[n]{a} - x_k\right)^2$.
4. Determine the multiplicity of $x^* = 0$ of $f(x) = \mathrm{e}^{2x} - 1 - 2x - 2x^2$. If $x_0 = 0.5$, find the approximation to x^* accurate within $|\, f\,(x_k)\,| \leqslant 10^{-3}$ by Newton's method and modified Newton's method.
5. Given equation $f(x) = x^3/3 - x = 0$, it is easy to know it has three roots $x_1^* = -\sqrt{3}, x_2^* = 0, x_3^* = \sqrt{3}$.

   (1) Try to find the $\delta > 0$ as large as possible such that the sequence converges to x_2^* for any $x_0 \in (-\delta, \delta)$;

   (2) If the initial data $x_0 \in (-\infty, -1), (-1, -\delta), (-\delta, \delta), (\delta, 1)$ or (1, $+\infty$, which root does the sequence converge?

   (3) What do you understand?