4

Simultaneous nonlinear equations

4.1 Introduction

In Chapter 1 we discussed iterative methods for the solution of a single nonlinear equation of the form $f(x) = 0$ where f is a continuous real-valued function of a single real variable. In Chapters 2 and 3, on the other hand, we were concerned with direct (as opposed to iterative) methods for systems of linear equations. The purpose of the present chapter is to extend the techniques developed in Chapter 1 to systems of simultaneous nonlinear equations for functions of several real variables. We shall concentrate on two methods: the generalisation of simple iteration, usually referred to as simultaneous iteration, and Newton's method.

Given that $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$, as in Chapters 2 and 3 we denote by $\|\boldsymbol{x}\|_\infty$ the ∞-norm of $\boldsymbol{x}$ defined by

$$\|\boldsymbol{x}\|_\infty = \max_{i=1}^{n} |x_i| \,.$$

Throughout the chapter, $\mathbb{R}^n$ will be thought of as a linear space equipped with the ∞-norm; with only minor alterations all of our results can be restated in the p-norm with $p \in [1, \infty)$ on replacing $\|\cdot\|_\infty$ by $\|\cdot\|_p$ throughout. We begin with some basic definitions which involve the concept of *open ball* defined in Section 2.7.

Let $\boldsymbol{\xi} \in \mathbb{R}^n$; the open ball in $\mathbb{R}^n$ (with respect to the ∞-norm) of radius $\varepsilon > 0$ and centre $\boldsymbol{\xi}$ is defined as the set

$$B_\varepsilon(\boldsymbol{\xi}) = \{\boldsymbol{x} \in \mathbb{R}^n \colon \|\boldsymbol{x} - \boldsymbol{\xi}\|_\infty < \varepsilon\} \,.$$

A set $D \subset \mathbb{R}^n$ is said to be an **open set** in $\mathbb{R}^n$ if for every $\boldsymbol{\xi} \in D$ there exists $\varepsilon = \varepsilon(\boldsymbol{\xi}) > 0$ such that $B_\varepsilon(\boldsymbol{\xi}) \subset D$ (see Figure 4.1). For example, any open ball in $\mathbb{R}^n$ is an open set in $\mathbb{R}^n$. Given $\boldsymbol{\xi} \in \mathbb{R}^n$, any open set

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

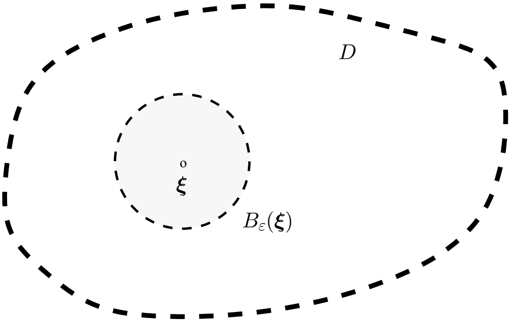

Fig. 4.1. Open set D: for each $\boldsymbol{\xi} \in D$ there exists $\varepsilon = \varepsilon(\boldsymbol{\xi})$ such that the open ball $B_\varepsilon(\boldsymbol{\xi})$ of radius ε and centre $\boldsymbol{\xi}$ is contained in D.

$N(\boldsymbol{\xi}) \subset \mathbb{R}^n$ containing $\boldsymbol{\xi}$ will be called a **neighbourhood** of $\boldsymbol{\xi}$; thus, any open set in $\mathbb{R}^n$ is a neighbourhood of each of its elements.

A set $D \subset \mathbb{R}^n$ is said to be a **closed set** in $\mathbb{R}^n$ if its complement $\mathbb{R}^n \setminus D$ is an open set in $\mathbb{R}^n$. For example, the closed ball of radius $\varepsilon > 0$ and centre $\boldsymbol{\xi}$, defined by

$$\bar{B}_\varepsilon(\boldsymbol{\xi}) = \{\boldsymbol{x} \in \mathbb{R}^n \colon \|\boldsymbol{x} - \boldsymbol{\xi}\|_\infty \le \varepsilon\},$$

is a closed set in $\mathbb{R}^n$.

A sequence $(\boldsymbol{x}^{(k)}) \subset \mathbb{R}^n$ is called a **Cauchy sequence** in $\mathbb{R}^n$ if for any $\varepsilon > 0$ there exists a positive integer $k_0 = k_0(\varepsilon)$ such that

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(m)}\|_\infty < \varepsilon \qquad \forall\, k, m \ge k_0 .$$

We shall make use of the fact that $\mathbb{R}^n$ is **complete**: that is, if $(\boldsymbol{x}^{(k)})$ is a Cauchy sequence in $\mathbb{R}^n$, then there exists $\boldsymbol{\xi}$ in $\mathbb{R}^n$ such that $(\boldsymbol{x}^{(k)})$ converges to $\boldsymbol{\xi}$; *i.e.*,

$$\lim_{k\to\infty} \|\boldsymbol{x}^{(k)} - \boldsymbol{\xi}\|_\infty = 0 . \tag{4.1}$$

For the sake of brevity, we shall write $\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{\xi}$ instead of (4.1).

Lemma 4.1 *Suppose that D is a nonempty closed subset of $\mathbb{R}^n$ and $(\boldsymbol{x}^{(k)}) \subset D$ is a Cauchy sequence in $\mathbb{R}^n$. Then, $\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{\xi}$ exists and $\boldsymbol{\xi} \in D$.*

Proof As $(\boldsymbol{x}^{(k)})$ is a Cauchy sequence in $\mathbb{R}^n$, there exists $\boldsymbol{\xi} \in \mathbb{R}^n$ such that $\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{\xi}$. It remains to prove that $\boldsymbol{\xi} \in D$. Suppose, otherwise, that $\boldsymbol{\xi}$ belongs to the open set $\mathbb{R}^n \setminus D$. Then, there exists

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

$\varepsilon > 0$ such that $B_\varepsilon(\boldsymbol{\xi}) \subset \mathbb{R}^n \setminus D$. As $(\boldsymbol{x}^{(k)}) \subset D$, no member of the sequence $(\boldsymbol{x}^{(k)})$ can enter $B_\varepsilon(\boldsymbol{\xi})$. This, however, contradicts the fact that $(\boldsymbol{x}^{(k)})$ converges to $\boldsymbol{\xi}$. The contradiction implies that $\boldsymbol{\xi} \in D$. □

Suppose that D is a nonempty subset of $\mathbb{R}^n$ and $\boldsymbol{f}\colon D(\subset \mathbb{R}^n) \to \mathbb{R}^n$ is a function defined on D. Given that $\boldsymbol{\xi} \in D$, we shall say that $\boldsymbol{f}$ is continuous at $\boldsymbol{\xi}$ if for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that, for every $\boldsymbol{x} \in B_\delta(\boldsymbol{\xi}) \cap D$,

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{\xi})\|_\infty < \varepsilon\,.$$

When a function $\boldsymbol{f}$, defined on the set D, is continuous at each point of D, it is said to be a **continuous function** on D.

Lemma 4.2 *Let D be a nonempty subset of $\mathbb{R}^n$ and $\boldsymbol{f}\colon D(\subset \mathbb{R}^n) \to \mathbb{R}^n$ a function, defined and continuous on D. If $(\boldsymbol{x}^{(k)}) \subset D$ converges in $\mathbb{R}^n$ to $\boldsymbol{\xi} \in D$, then $\lim_{k\to\infty} \boldsymbol{f}(\boldsymbol{x}^{(k)}) = \boldsymbol{f}(\boldsymbol{\xi})$.*

Proof Due to the continuity of $\boldsymbol{f}$ at $\boldsymbol{\xi} \in D$, given $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that if $\|\boldsymbol{x} - \boldsymbol{\xi}\|_\infty < \delta$ for some $\boldsymbol{x} \in D$, then

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{\xi})\|_\infty < \varepsilon\,. \tag{4.2}$$

Further, as $(\boldsymbol{x}^{(k)})$ converges to $\boldsymbol{\xi}$, there exists $k_0 = k_0(\delta) = k_0(\delta(\varepsilon))$ such that

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{\xi}\|_\infty < \delta \qquad \forall\, k \ge k_0\,.$$

Hence, taking $\boldsymbol{x} = \boldsymbol{x}^{(k)}$ in (4.2), we deduce that for each $\varepsilon > 0$ there exists k_0 such that

$$\|\boldsymbol{f}(\boldsymbol{x}^{(k)}) - \boldsymbol{f}(\boldsymbol{\xi})\|_\infty < \varepsilon \qquad \forall\, k \ge k_0\,,$$

which means that $\lim_{k\to\infty} \boldsymbol{f}(\boldsymbol{x}^{(k)}) = \boldsymbol{f}(\boldsymbol{\xi})$. □

After this brief preparation, we are ready to embark on the development of numerical algorithms for the solution of systems of simultaneous nonlinear equations.

4.2 Simultaneous iteration

Let D be a nonempty closed subset of $\mathbb{R}^n$ and $\boldsymbol{f}\colon D(\subset \mathbb{R}^n) \to \mathbb{R}^n$ a continuous function defined on D. We shall be concerned with the problem of finding $\boldsymbol{\xi} \in D$ such that $\boldsymbol{f}(\boldsymbol{\xi}) = \boldsymbol{0}$. If such $\boldsymbol{\xi}$ exists, it is

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

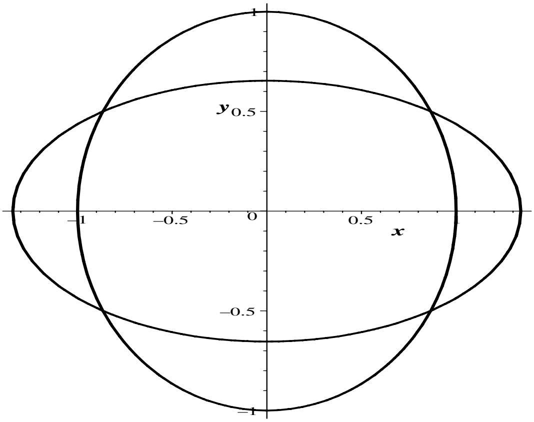

Fig. 4.2. Graphs of the curves $x_1^2 + x_2^2 - 1 = 0$ and $5x_1^2 + 21x_2^2 - 9 = 0$.

called a **solution** to the equation $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ (in D). When written in componentwise form, $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ becomes

$$f_i(x_1, \dots, x_n) = 0\,, \quad i = 1, \dots, n\,,$$

a system of n simultaneous nonlinear equations for n unknowns, where $f_1, \dots, f_n$ are the components of $\boldsymbol{f}$.

Example 4.1 *Consider the system of two simultaneous nonlinear equations in two unknowns, x_1 and x_2, defined by*

$$\begin{aligned} x_1^2 + x_2^2 - 1 &= 0\,, \\ 5x_1^2 + 21x_2^2 - 9 &= 0\,. \end{aligned}$$

Here $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$ and $\boldsymbol{f} = (f_1, f_2)^{\mathrm{T}}$ with

$$\begin{aligned} f_1(x_1, x_2) &= x_1^2 + x_2^2 - 1\,, \\ f_2(x_1, x_2) &= 5x_1^2 + 21x_2^2 - 9\,. \end{aligned}$$

The equation $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ has four solutions:

$$\begin{aligned} \boldsymbol{\xi}_1 &= (-\sqrt{3}/2, 1/2)^{\mathrm{T}}\,, & \boldsymbol{\xi}_2 &= (\sqrt{3}/2, 1/2)^{\mathrm{T}}\,, \\ \boldsymbol{\xi}_3 &= (-\sqrt{3}/2, -1/2)^{\mathrm{T}}\,, & \boldsymbol{\xi}_4 &= (\sqrt{3}/2, -1/2)^{\mathrm{T}}\,. \end{aligned}$$

The curves $f_1(x_1, x_2) = 0$ and $f_2(x_1, x_2) = 0$ are depicted in Figure 4.2. The four solutions correspond to the four points of intersection of the two curves in the figure.

Example 4.2 *Let us suppose that $A \in \mathbb{R}^{n\times n}$ and $\boldsymbol{b} \in \mathbb{R}^n$. On letting $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{b} - A\boldsymbol{x}$ we deduce that the problem of solving the system of simultaneous linear equations considered in Chapters 2 and 3 can be restated in the form: find $\boldsymbol{x} \in \mathbb{R}^n$ such that $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$.*

Let us assume that we have transformed the equation $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$ into an equivalent form $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x}$, where $\boldsymbol{g}\colon \mathbb{R}^n \to \mathbb{R}^n$ is a continuous function, defined on the closed subset $D \subset \mathbb{R}^n$, such that $\boldsymbol{g}(D) \subset D$. For example, one can choose $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x} - \alpha\boldsymbol{f}(\boldsymbol{x})$, with $\alpha \in \mathbb{R}$ a suitable parameter. By 'equivalent' we mean that $\boldsymbol{\xi} \in D$ satisfies $\boldsymbol{f}(\boldsymbol{\xi}) = \mathbf{0}$ if, and only if, $\boldsymbol{g}(\boldsymbol{\xi}) = \boldsymbol{\xi}$. Any $\boldsymbol{\xi} \in D$ such that $\boldsymbol{g}(\boldsymbol{\xi}) = \boldsymbol{\xi}$ is called a **fixed point** of the function $\boldsymbol{g}$ in D. Thus the problem of finding a solution $\boldsymbol{\xi} \in D$ to the equation $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$ has been converted into one of finding a fixed point in D of the function $\boldsymbol{g}$. We embark on the latter task by considering the natural extension to $\mathbb{R}^n$ of the simple iteration discussed in Section 1.2 for the solution of the scalar nonlinear equation $g(x) = x$.

Definition 4.1 *Suppose that $\boldsymbol{g}\colon \mathbb{R}^n \to \mathbb{R}^n$ is a function, defined and continuous on a closed subset D of $\mathbb{R}^n$, such that $\boldsymbol{g}(D) \subset D$. Given that $\boldsymbol{x}_0 \in D$, the recursion defined by*

$$\boxed{\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)}), \qquad k = 0, 1, 2, \dots, \qquad (4.3)}$$

is called a **simultaneous iteration**. *For $n = 1$ the recursion (4.3) is just the simple iteration considered in (1.3).*

Note that here we use the superscript k as the sequence index; following the convention adopted in Chapters 2 and 3, we reserve subscripts for labelling the entries of vectors. Thus $x_i^{(k)}$ is entry i of the vector $\boldsymbol{x}^{(k)}$, the kth member of the sequence $(\boldsymbol{x}^{(k)})$. The motivation behind the definition of the simultaneous iteration (4.3) is, of course, our hope that, under suitable conditions on $\boldsymbol{g}$ and D, the sequence $(\boldsymbol{x}^{(k)})$ will converge to a fixed point $\boldsymbol{\xi}$ of $\boldsymbol{g}$.

Two remarks are in order at this point. First, it is easy to show that if a sequence of vectors $(\boldsymbol{x}^{(k)})$ converges in $\mathbb{R}^n$ to $\boldsymbol{\xi}$ in the norm $\|\cdot\|_\infty$, then it also converges to this same limit in the norm $\|\cdot\|_p$ for any $p \in [1, \infty)$. To see this, note that

$$\|\boldsymbol{w}\|_\infty \le \|\boldsymbol{w}\|_p \le n^{1/p}\|\boldsymbol{w}\|_\infty \qquad \forall\, \boldsymbol{w} \in \mathbb{R}^n, \qquad (4.4)$$

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

for $1 \le p < \infty$, and take $\boldsymbol{w} = \boldsymbol{x}^{(k)} - \boldsymbol{\xi}$ to deduce that, as $k \to \infty$, convergence in the ∞-norm implies convergence in the p-norm for any $p \in [1, \infty)$, and *vice versa.* Thus, in this sense, the choice of norm on $\mathbb{R}^n$ is irrelevant. Second, the assumption that D is a closed set is crucial in our discussion. If D is not closed, $\boldsymbol{g}$: $D \to D$ need not have a fixed point in D, even if $\boldsymbol{x}^{(k)} \in D$ for all $k \ge 0$ and $(\boldsymbol{x}^{(k)})$ converges in $\mathbb{R}^n$. We verify this claim through a simple example.

Example 4.3 *Suppose that D is the open unit disc in $\mathbb{R}^2$ in the ∞-norm, which is just the square defined by $-1 < x_1 < 1$, $-1 < x_2 < 1$. Consider the simultaneous iteration defined by (4.3), where $\boldsymbol{x}^{(0)} = \boldsymbol{0} \in D$, and*

$$\boldsymbol{g}(\boldsymbol{x}) = \tfrac{1}{2}(\boldsymbol{x} + \boldsymbol{u})\,, \quad \boldsymbol{u} = (1,1)^{\mathrm{T}}\,.$$

If $\|\boldsymbol{x}\|_\infty < 1$ it is easy to see that $\|\boldsymbol{g}(\boldsymbol{x})\|_\infty < 1$; hence, starting the iteration $\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)})$ from $\boldsymbol{x}^{(0)} = \boldsymbol{0}$, it follows that $\boldsymbol{x}^{(k)} \in D$ for all $k \ge 0$. The definition of $\boldsymbol{g}$ implies at once that

$$\boldsymbol{x}^{(k+1)} - \boldsymbol{u} = \tfrac{1}{2}(\boldsymbol{x}^{(k)} - \boldsymbol{u})\,,$$

and therefore

$$\|\boldsymbol{x}^{(k+1)} - \boldsymbol{u}\|_\infty = \tfrac{1}{2}\|\boldsymbol{x}^{(k)} - \boldsymbol{u}\|_\infty = \cdots = \left(\tfrac{1}{2}\right)^{k+1}\|\boldsymbol{x}^{(0)} - \boldsymbol{u}\|_\infty = \left(\tfrac{1}{2}\right)^{k+1}\,,$$

from which it is obvious that the sequence $(\boldsymbol{x}^{(k)})$ converges in $\mathbb{R}^2$ to the limit $\boldsymbol{u}$. However, $\boldsymbol{u} \notin D$, since $\boldsymbol{u}$ lies on the unit circle in the ∞-norm that represents the boundary of the open set D. ◇

Up to now we have been assuming that the function $\boldsymbol{g}$: $\mathbb{R}^n \to \mathbb{R}^n$ is defined and continuous on a closed subset D of $\mathbb{R}^n$. In order to ensure that $\boldsymbol{g}$ has a (unique) fixed point in D, we strengthen our hypotheses on the function $\boldsymbol{g}$.

Definition 4.2 *Suppose that $\boldsymbol{g}$: $\mathbb{R}^n \to \mathbb{R}^n$ is defined on a closed subset D of $\mathbb{R}^n$. If there exists a positive constant L such that,*

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\|_\infty \;\le\; L\,\|\boldsymbol{x} - \boldsymbol{y}\|_\infty \tag{4.5}$$

for all $\boldsymbol{x}$ and $\boldsymbol{y}$ in D, then we say that $\boldsymbol{g}$ satisfies a **Lipschitz condition** *on D in the ∞-norm. The number L is called a* **Lipschitz constant** *for $\boldsymbol{g}$ in the ∞-norm. In particular, if $L \in (0,1)$, then $\boldsymbol{g}$ is said to be a* **contraction** *on D in the ∞-norm.*

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

Any function $\boldsymbol{g}$ that satisfies a Lipschitz condition on a set D is continuous on D. For let $\boldsymbol{x}_0 \in D$ and $\varepsilon > 0$; then, on defining $\delta = \varepsilon/L$, we deduce from (4.5) that if $\|\boldsymbol{x} - \boldsymbol{x}_0\|_\infty < \delta$ for some $\boldsymbol{x} \in D$, then

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{x}_0)\|_\infty \ \leq \ L\,\|\boldsymbol{x} - \boldsymbol{x}_0\|_\infty < \varepsilon\,.$$

It follows from (4.4) that if $\boldsymbol{g}$ satisfies a Lipschitz condition on D in the ∞-norm then it also does so in the p-norm for any $p \in [1, \infty)$, and *vice versa.* However, in general, the size of the constant L may depend on the choice of norm. Specifically, if $\boldsymbol{g}$ is a contraction on a set D in the ∞-norm (*i.e.*, (4.5) holds with $L < 1$), then $\boldsymbol{g}$ *need not* be a contraction in the p-norm, unless $L < n^{-1/p}$. (See Exercise 1.) Conversely, if $\boldsymbol{g}$ is a contraction on D in the p-norm for some $p \in [1, \infty)$, it does *not* follow that $\boldsymbol{g}$ is a contraction on D in the ∞-norm.

For example, suppose that $\boldsymbol{g}\colon \mathbb{R}^2 \to \mathbb{R}^2$ is the linear function defined by $\boldsymbol{g}(\boldsymbol{x}) = A\boldsymbol{x}$, where A is the 2×2 matrix

$$A = \begin{pmatrix} 3/4 & 1/3 \\ 0 & 3/4 \end{pmatrix}.$$

This function $\boldsymbol{g}$ satisfies a Lipschitz condition on $\mathbb{R}^2$ in $\|\cdot\|_p$ for any $p \in [1, \infty]$, and if L is a Lipschitz constant for $\boldsymbol{g}$ in the p-norm, then $L \geq \|A\|_p$, in the subordinate matrix norm. It is easy to see that $\|A\|_1 = \|A\|_\infty = 13/12$, and a small calculation gives $\|A\|_2 = 0.935$ to three decimal digits. Hence the function $\boldsymbol{g}$ is a contraction in the 2-norm, but not in the 1- or ∞-norm.

Our next result is a direct generalisation of Theorem 1.3 formulated in Chapter 1.

Theorem 4.1 (**Contraction Mapping Theorem**) *Suppose that D is a closed subset of $\mathbb{R}^n$, $\boldsymbol{g}\colon \mathbb{R}^n \to \mathbb{R}^n$ is defined on D, and $\boldsymbol{g}(D) \subset D$. Suppose further that $\boldsymbol{g}$ is a contraction on D in the ∞-norm. Then, $\boldsymbol{g}$ has a unique fixed point $\boldsymbol{\xi}$ in D, and the sequence $(\boldsymbol{x}^{(k)})$ defined by (4.3) converges to $\boldsymbol{\xi}$ for any starting value $\boldsymbol{x}^{(0)} \in D$.*

Proof Assuming that $\boldsymbol{g}$ has a fixed point $\boldsymbol{\xi}$ in D, the *uniqueness* of the fixed point is easy to show: for suppose that $\boldsymbol{\eta}$ is also a fixed point of $\boldsymbol{g}$ in D. Then, by (4.5),

$$\|\boldsymbol{\xi} - \boldsymbol{\eta}\|_\infty = \|\boldsymbol{g}(\boldsymbol{\xi}) - \boldsymbol{g}(\boldsymbol{\eta})\|_\infty \leq L\|\boldsymbol{\xi} - \boldsymbol{\eta}\|_\infty\,,$$

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

i.e., $(1-L)\|\boldsymbol{\xi}-\boldsymbol{\eta}\|_\infty \le 0$. Since $L \in (0,1)$, and $\|\cdot\|_\infty$ is a norm, it follows that $\boldsymbol{\xi}-\boldsymbol{\eta}=\mathbf{0}$, and hence $\boldsymbol{\xi}=\boldsymbol{\eta}$. Consequently, if $\boldsymbol{g}$ has a fixed point in D, then this is the unique fixed point of $\boldsymbol{g}$ in D.

Now, still *assuming* that $\boldsymbol{g}$ possesses a fixed point $\boldsymbol{\xi} \in D$, we shall show that the sequence $(\boldsymbol{x}^{(k)})$ defined by (4.3) converges to $\boldsymbol{\xi}$ for any starting value $\boldsymbol{x}^{(0)} \in D$. By repeating the argument from Chapter 1 which led to (1.10), with the absolute value sign $|\cdot|$ replaced by $\|\cdot\|_\infty$ throughout, we find that

$$\|\boldsymbol{x}^{(k)}-\boldsymbol{\xi}\|_\infty \le L^k \frac{1}{1-L}\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}\|_\infty .$$

As $L \in (0,1)$, we deduce that $\lim_{k\to\infty} L^k = 0$, and hence,

$$\lim_{k\to\infty}\|\boldsymbol{x}^{(k)}-\boldsymbol{\xi}\|_\infty = 0\,,$$

showing that the sequence $(\boldsymbol{x}^{(k)})$ defined by (4.3) converges to $\boldsymbol{\xi}$ for any starting value $\boldsymbol{x}^{(0)} \in D$. In particular, if $\varepsilon > 0$, then letting

$$k_0 = k_0(\varepsilon) = \left[\frac{\ln\|\,x_1 - x_0\,\|_\infty - \ln(\varepsilon(1-L))}{\ln(1/L)}\right] + 1\,, \tag{4.6}$$

we find that

$$L^k \frac{1}{1-L}\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}\|_\infty \le \varepsilon$$

for all $k \ge k_0(\varepsilon)$, and therefore

$$\|\boldsymbol{x}^{(k)}-\boldsymbol{\xi}\|_\infty \le \varepsilon\,, \tag{4.7}$$

for all $k \ge k_0(\varepsilon)$, as in Chapter 1. A brief comment on the notation: in (4.6), $[x]$ denotes the integer part of the real number x; *i.e.*, $[x]$ is the largest integer such that $[x] \le x$ – just as in Theorem 1.4.

In order to complete the proof of the theorem, it remains to show the *existence* of a fixed point $\boldsymbol{\xi} \in D$ for $\boldsymbol{g}$. In contrast with the proof of existence of a fixed point for a real-valued function of a single real variable presented in Chapter 1, here we cannot rely on the Intermediate Value Theorem (unless, of course, $n = 1$), so we shall develop a different argument. The essence of this will be to show that $(\boldsymbol{x}^{(k)}) \subset D$ is a Cauchy sequence in $\mathbb{R}^n$; for then we can apply Lemmas 4.1 and 4.2 to deduce that the sequence converges to a fixed point $\boldsymbol{\xi}$ of the function $\boldsymbol{g}$.

Let us begin by noting that since $\boldsymbol{g}(D) \subset D$, if $\boldsymbol{x}^{(0)}$ belongs to D, then $\boldsymbol{x}^{(k)} = \boldsymbol{g}(\boldsymbol{x}^{(k-1)}) \in D$ for all $k \ge 1$. Further, since $\boldsymbol{g}$ is a contraction on D in the ∞-norm, we have that

$$\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^{(k-1)}\|_\infty = \|\boldsymbol{g}(\boldsymbol{x}^{(k-1)})-\boldsymbol{g}(\boldsymbol{x}^{(k-2)})\|_\infty \le L\|\boldsymbol{x}^{(k-1)}-\boldsymbol{x}^{(k-2)}\|_\infty$$

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

for all $k \geq 2$. We then deduce by induction that

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\|_\infty \leq L^{k-1}\|\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}\|_\infty\,, \qquad k \geq 1\,. \tag{4.8}$$

Suppose that m and k are positive integers and $m \geq k+1$. Then, by repeated application of the triangle inequality in the ∞-norm and using (4.8), we have that

$$\begin{aligned}
\|\boldsymbol{x}^{(m)} - \boldsymbol{x}^{(k)}\|_\infty &= \|(\boldsymbol{x}^{(m)} - \boldsymbol{x}^{(m-1)}) + \cdots + (\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)})\|_\infty \\
&\leq \|\boldsymbol{x}^{(m)} - \boldsymbol{x}^{(m-1)}\|_\infty + \cdots + \|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\|_\infty \\
&\leq (L^{m-1} + \cdots + L^k)\|\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}\|_\infty \\
&= L^k(L^{m-k-1} + \cdots + 1)\|\boldsymbol{x}^{(1)} - \boldsymbol{x}^{0)}\|_\infty \\
&\leq L^k\frac{1}{1-L}\|\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}\|_\infty\,,
\end{aligned} \tag{4.9}$$

where, in the transition to the last line, we made use of the fact that the geometric series $1 + L + L^2 + \cdots$, with $L \in (0,1)$, sums to $1/(1-L)$.

As $\lim_{k\to\infty} L^k = 0$, it follows from (4.9) that $(\boldsymbol{x}^{(k)})$ is a Cauchy sequence in $\mathbb{R}^n$; that is, for each $\varepsilon > 0$ there exists $k_0 = k_0(\varepsilon)$ (defined by (4.6) above) such that

$$\|\boldsymbol{x}^{(m)} - \boldsymbol{x}^{(k)}\|_\infty < \varepsilon \qquad \forall\, m, k \geq k_0 = k_0(\varepsilon)\,. \tag{4.10}$$

Any Cauchy sequence in $\mathbb{R}^n$ is convergent in $\mathbb{R}^n$; consequently, there exists $\boldsymbol{\xi} \in \mathbb{R}^n$ such that $\boldsymbol{\xi} = \lim_{k\to\infty} \boldsymbol{x}^{(k)}$. Further, since $\boldsymbol{g}$ satisfies a Lipschitz condition on D, the discussion in the paragraph following Definition 4.2 shows that $\boldsymbol{g}$ is continuous on D. Hence, by Lemma 4.2,

$$\boldsymbol{\xi} = \lim_{k\to\infty} \boldsymbol{x}^{(k+1)} = \lim_{k\to\infty} \boldsymbol{g}(\boldsymbol{x}^{(k)}) = \boldsymbol{g}\left(\lim_{k\to\infty} \boldsymbol{x}^{(k)}\right) = \boldsymbol{g}(\boldsymbol{\xi})\,,$$

which proves that $\boldsymbol{\xi}$ is a fixed point of $\boldsymbol{g}$.

It remains to show that $\boldsymbol{\xi} \in D$. This follows from Lemma 4.1 since $(\boldsymbol{x}^{(k)}) \subset D$, $\boldsymbol{\xi} = \lim_{k\to\infty} \boldsymbol{x}^{(k)}$ and D is closed. □

As a byproduct of the proof, we deduce from (4.7) that, given a positive tolerance ε, one can compute an approximation $\boldsymbol{x}^{(k)}$ to the unknown solution $\boldsymbol{\xi}$ using (4.3) in no more than $k_0 = k_0(\varepsilon)$ iterations so that the approximation error $\boldsymbol{\xi} - \boldsymbol{x}^{(k)}$, measured in the ∞-norm, is less than ε; the integer $k_0(\varepsilon)$ is defined by (4.6).

The next theorem relates the constant L from the Lipschitz condition (4.5) to the partial derivatives of $\boldsymbol{g}$, giving a more practically useful sufficient condition for convergence.

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

Definition 4.3 *Let* $\boldsymbol{g} = (g_1, \dots, g_n)^{\mathrm{T}} \colon \mathbb{R}^n \to \mathbb{R}^n$ *be a function defined and continuous in an (open) neighbourhood* $N(\boldsymbol{\xi})$ *of* $\boldsymbol{\xi} \in \mathbb{R}^n$*. Suppose further that the first partial derivatives* $\frac{\partial g_i}{\partial x_j}$*,* $j = 1, \dots, n$*, of* g_i *exist at* $\boldsymbol{\xi}$ *for* $i = 1, \dots, n$*. The* **Jacobian matrix** $J_g(\boldsymbol{\xi})$ *of* $\boldsymbol{g}$ *at* $\boldsymbol{\xi}$ *is the* $n \times n$ *matrix with elements*

$$J_g(\boldsymbol{\xi})_{ij} = \frac{\partial g_i}{\partial x_j}(\boldsymbol{\xi}), \qquad i, j = 1, \dots, n.$$

Theorem 4.2 *Suppose that* $\boldsymbol{g} = (g_1, \dots, g_n)^{\mathrm{T}} \colon \mathbb{R}^n \to \mathbb{R}^n$ *is defined and continuous on a closed set* $D \subset \mathbb{R}^n$*. Let* $\boldsymbol{\xi} \in D$ *be a fixed point of* $\boldsymbol{g}$*, and suppose that the first partial derivatives* $\frac{\partial g_i}{\partial x_j}$*,* $j = 1, \dots, n$*, of* g_i*,* $i = 1, \dots, n$*, are defined and continuous in some (open) neighbourhood* $N(\boldsymbol{\xi}) \subset D$ *of* $\boldsymbol{\xi}$*, with*

$$\|J_g(\boldsymbol{\xi})\|_\infty < 1.$$

Then, there exists $\varepsilon > 0$ *such that* $\boldsymbol{g}(\bar{B}_\varepsilon(\boldsymbol{\xi})) \subset \bar{B}_\varepsilon(\boldsymbol{\xi})$*, and the sequence defined by (4.3) converges to* $\boldsymbol{\xi}$ *for all* $\boldsymbol{x}^{(0)} \in \bar{B}_\varepsilon(\boldsymbol{\xi})$*.*

Proof The proof is a natural extension of that of Theorem 1.5. We write $K = \|J_g(\boldsymbol{\xi})\|_\infty$. Since the partial derivatives $\frac{\partial g_i}{\partial x_j}$, $i, j = 1, \dots, n$, are continuous in the neighbourhood $N(\boldsymbol{\xi})$ of $\boldsymbol{\xi}$, we can find a closed ball $\bar{B}_\varepsilon(\boldsymbol{\xi}) \subset N(\boldsymbol{\xi}) \subset D$ of radius ε and centre $\boldsymbol{\xi}$ such that

$$\|J_g(\boldsymbol{z})\|_\infty \le \tfrac{1}{2}(K+1) < 1 \qquad \forall\, \boldsymbol{z} \in \bar{B}_\varepsilon(\boldsymbol{\xi}). \tag{4.11}$$

Now, suppose that $\boldsymbol{x}$ and $\boldsymbol{y}$ are both in $\bar{B}_\varepsilon(\boldsymbol{\xi})$ and, for $i \in \{1, \dots, n\}$ fixed, define the function $t \mapsto \varphi_i(t)$ of the single variable $t \in [0,1]$ by

$$\varphi_i(t) = g_i(t\boldsymbol{x} + (1-t)\boldsymbol{y});$$

thus, $\varphi_i(0) = g_i(\boldsymbol{y})$ and $\varphi_i(1) = g_i(\boldsymbol{x})$. The function $t \mapsto \varphi_i(t)$ has a continuous derivative in t on the interval $[0,1]$; thus, by the Mean Value Theorem (Theorem A.3), there exists $\eta \in (0,1)$ such that

$$g_i(\boldsymbol{x}) - g_i(\boldsymbol{y}) = \varphi_i(1) - \varphi_i(0) = \varphi_i'(\eta)(1-0) = \varphi_i'(\eta).$$

This means that

$$g_i(\boldsymbol{x}) - g_i(\boldsymbol{y}) = \sum_{j=1}^{n} (x_j - y_j) \frac{\partial g_i}{\partial x_j}(\eta\boldsymbol{x} + (1-\eta)\boldsymbol{y}) \tag{4.12}$$

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

for $i = 1, \ldots, n$. Now $|x_j - y_j| \leq \|\boldsymbol{x} - \boldsymbol{y}\|_\infty$ for all $j \in \{1, \ldots, n\}$, and so (4.12) gives

$$\begin{aligned} |g_i(\boldsymbol{x}) - g_i(\boldsymbol{y})| &\leq \|\boldsymbol{x} - \boldsymbol{y}\|_\infty \sum_{j=1}^{n} \left| \frac{\partial g_i}{\partial x_j}(\eta \boldsymbol{x} + (1-\eta)\boldsymbol{y}) \right| \\ &\leq \|\boldsymbol{x} - \boldsymbol{y}\|_\infty \, \|J_g(\eta \boldsymbol{x} + (1-\eta)\boldsymbol{y})\|_\infty \,, \end{aligned}$$

for all $i = 1, \ldots, n$. Consequently, for any $\boldsymbol{x}, \boldsymbol{y} \in \bar{B}_\varepsilon(\boldsymbol{\xi})$,

$$\begin{aligned} \|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\|_\infty &\leq \max_{t \in [0,1]} \|J_g(t\boldsymbol{x} + (1-t)\boldsymbol{y})\|_\infty \|\boldsymbol{x} - \boldsymbol{y}\|_\infty \\ &\leq \tfrac{1}{2}(1 + K)\|\boldsymbol{x} - \boldsymbol{y}\|_\infty \,, \end{aligned} \qquad (4.13)$$

due to (4.11), given that $t\boldsymbol{x} + (1-t)\boldsymbol{y} \in \bar{B}_\varepsilon(\boldsymbol{\xi})$ for all $t \in [0, 1]$. It follows that $\boldsymbol{g}$ satisfies a Lipschitz condition (4.5), in the ∞-norm, on the closed ball $\bar{B}_\varepsilon(\boldsymbol{\xi})$ with $L = \frac{1}{2}(1 + K) < 1$. Furthermore, on selecting $\boldsymbol{y} = \boldsymbol{\xi}$ in (4.13) we get that

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{\xi}\|_\infty = \|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{\xi})\|_\infty < \|\boldsymbol{x} - \boldsymbol{\xi}\|_\infty \leq \varepsilon$$

for all $\boldsymbol{x} \in \bar{B}_\varepsilon(\boldsymbol{\xi})$. Hence, $\boldsymbol{g}(\bar{B}_\varepsilon(\boldsymbol{\xi})) \subset \bar{B}_\varepsilon(\boldsymbol{\xi})$. The convergence of the iteration (4.3) to $\boldsymbol{\xi}$, for an arbitrary starting value $\boldsymbol{x}^{(0)} \in \bar{B}_\varepsilon(\boldsymbol{\xi})$, now follows from Theorem 4.1. □

We close this section with an example which illustrates the application of the method of simultaneous iteration to the solution of a system of nonlinear equations.

Example 4.4 *Let us consider, as in Example 4.1, the system of two simultaneous nonlinear equations in the unknowns x_1 and x_2, defined by*

$$\begin{aligned} x_1^2 + x_2^2 - 1 &= 0\,, \\ 5x_1^2 + 21x_2^2 - 9 &= 0\,. \end{aligned}$$

Here $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$ and $\boldsymbol{f} = (f_1, f_2)^{\mathrm{T}}$ with

$$\begin{aligned} f_1(x_1, x_2) &= x_1^2 + x_2^2 - 1\,, \\ f_2(x_1, x_2) &= 5x_1^2 + 21x_2^2 - 9\,. \end{aligned}$$

Let us suppose that we need to find the solution of the system $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ in the first quadrant of the (x_1, x_2)-coordinate system.

Of course, the example is a little artificial, since we already know from Example 4.1 that $\boldsymbol{\xi}_2 = (\sqrt{3}/2, 1/2)^{\mathrm{T}}$ is the required solution. In what follows, however, we proceed as if we knew nothing about the location

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

of $\boldsymbol{\xi}_2$. Our aim here is to illustrate the construction of the function $\boldsymbol{g}$ from $\boldsymbol{f}$ and the verification of the hypotheses of Theorem 4.1.

Let us rewrite the two equations as

$$x_1 = \left(1 - x_2^2\right)^{1/2}, \qquad x_2 = \frac{1}{\sqrt{21}}\left(9 - 5x_1^2\right)^{1/2},$$

and define $g_1(x_1, x_2)$ and $g_2(x_1, x_2)$ as the right-hand sides of these, respectively. We consider the simultaneous iteration

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)}), \qquad k = 0, 1, 2, \ldots, \tag{4.14}$$

with suitably chosen $\boldsymbol{x}^{(0)}$ and $\boldsymbol{g} = (g_1, g_2)^{\mathrm{T}}$.

Our first task is to find a closed subset D of $\mathbb{R}^2$ containing the required solution, such that $\boldsymbol{g}$ satisfies the hypotheses of Theorem 4.1 on D. In order to ensure that $\boldsymbol{x} \mapsto \boldsymbol{g}(\boldsymbol{x})$ is real-valued and continuous, and that the partial derivatives of g_1 and g_2 are continuous at $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}} \in \mathbb{R}^2$, we demand that $|x_2| < 1$ and $|x_1| < 3/(\sqrt{5})$. In fact, since we are looking for a solution in the first quadrant, it is natural to suppose that $x_1 \geq 0$, $x_2 \geq 0$. Hence we let $M = \{\boldsymbol{x} \in \mathbb{R}^2 \colon 0 \leq x_1 < 3/\sqrt{5},\ \ 0 \leq x_2 < 1\}$, and we seek D as a suitable closed subset of M.

For $\boldsymbol{x} \in M$, let

$$J_g(\boldsymbol{x}) = \begin{pmatrix} \partial g_1/\partial x_1 & \partial g_1/\partial x_2 \\ \partial g_2/\partial x_1 & \partial g_2/\partial x_2 \end{pmatrix}.$$

Clearly,

$$\frac{\partial g_1}{\partial x_1} = 0, \qquad \frac{\partial g_1}{\partial x_2} = -x_2\left(1 - x_2^2\right)^{-1/2},$$
$$\frac{\partial g_2}{\partial x_1} = -\frac{5}{\sqrt{21}}\,x_1\left(9 - 5x_1^2\right)^{-1/2}, \qquad \frac{\partial g_2}{\partial x_2} = 0,$$

so we conclude that, for any $\boldsymbol{x} \in M$,

$$\|J_g(\boldsymbol{x})\|_\infty = \max\left(x_2\left(1 - x_2^2\right)^{-1/2}, \frac{5}{\sqrt{21}}\,x_1\left(9 - 5x_1^2\right)^{-1/2}\right).$$

In particular, we have $\|J_g(\boldsymbol{x})\|_\infty < 1$ provided that

$$x_2^2 < 1 - x_2^2 \quad \text{and} \quad 25x_1^2 < 21(9 - 5x_1^2),$$

that is, when $x_2^2 < 1/2$ and $x_1^2 < 189/130$. These conditions are clearly satisfied if, for example, $0 \leq x_1 \leq 1$ and $0 \leq x_2 \leq 3/5$. If we now define $D = [0, 1] \times [0, 3/5]$, then, analogously as in (4.13), we have that

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\|_\infty \leq \max_{t \in [0,1]} \|J_g(t\boldsymbol{x} + (1 - t)\boldsymbol{y})\|_\infty\, \|\boldsymbol{x} - \boldsymbol{y}\|_\infty$$

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

for all $\boldsymbol{x}$ and $\boldsymbol{y}$ in D. Therefore, also,

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\|_\infty \leq L\|\boldsymbol{x} - \boldsymbol{y}\|_\infty$$

with

$$L = \max_{\boldsymbol{z}\in D} \|J_g(\boldsymbol{z})\|_\infty < 1\,. \tag{4.15}$$

With our choice of D, (4.15) holds with $L = \max\{0.75, 0.55\} = 0.75 < 1$. Furthermore, it is easy to check that $\boldsymbol{g}(D) \subset D$. Thus we deduce from Theorem 4.1 that $\boldsymbol{g}$ has a unique fixed point in D – we call this fixed point $\boldsymbol{\xi}_2$, for the sake of consistency with the notation in Example 4.1; moreover, the sequence $(\boldsymbol{x}^{(k)})$ defined by (4.14) converges to $\boldsymbol{\xi}_2$.

After all these preparations you are now probably curious to see what the successive iterates look like: Table 4.1 gives a flavour of the behaviour of the sequence $(\boldsymbol{x}^{(k)})$, with the starting value chosen as $\boldsymbol{x}^{(0)} = (0.5, 0.3)^{\mathrm{T}}$. You can see that after 15 iterations the first 5 decimal digits have settled to their correct values.[1]

4.3 Relaxation and Newton's method

We now go on to apply the ideas developed in the previous section to the construction of an iteration which converges to a solution of the equation $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$, where $\boldsymbol{f}\colon \mathbb{R}^n \to \mathbb{R}^n$. One way of constructing such a sequence is by relaxation.

Definition 4.4 *The recursion*

$$\boxed{\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \lambda \boldsymbol{f}(\boldsymbol{x}^{(k)})\,, \qquad k = 0, 1, 2, \ldots\,, \qquad (4.16)}$$

where $\boldsymbol{x}_0 \in \mathbb{R}^n$ is given and where $\lambda \neq 0$ is a constant, is called **simultaneous relaxation**.

Suppose that the sequence $(\boldsymbol{x}^{(k)})$ converges to a limit $\boldsymbol{\xi} \in \mathbb{R}^n$ and $\boldsymbol{f}$ is continuous in a neighbourhood of $\boldsymbol{\xi}$; then, on passing to the limit $k \to \infty$ in (4.16), we deduce that $\boldsymbol{\xi}$ is a solution of the equation $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$.

Simultaneous relaxation is evidently a simultaneous iteration defined by taking $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x} - \lambda \boldsymbol{f}(\boldsymbol{x})$.

[1] You may wish to contemplate the following question: *how many iterations should be performed to ensure that all 15 digits have settled to their correct values?* Use inequality (4.6) to get an idea of the (maximum) amount of work involved!

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

Table 4.1. *The first 15 iterates in the sequence* $\boldsymbol{x}^{(k)} = (x_1^{(k)}, x_2^{(k)})^{\mathrm{T}}$ *defined by (4.14), with starting value* $(0.5, 0.3)^{\mathrm{T}}$*. The exact solution is* $\boldsymbol{\xi}_2 = (\sqrt{3}/2, 1/2)^{\mathrm{T}} = (0.866025403784439, 0.500000000000000)^{\mathrm{T}}$ *to 15 decimal digits.*

k	$x_1^{(k)}$	$x_2(k)$
0	0.500000000000000	0.300000000000000
1	0.953939197667987	0.607492896293956
2	0.794325110362489	0.460331145598201
3	0.887747281827575	0.527583804908580
4	0.849502989281489	0.490845908224662
5	0.871246402792635	0.506703790432366
6	0.862120217116774	0.497835722000956
7	0.867271349636195	0.501604267098156
8	0.865097196405654	0.499485546313646
9	0.866322220091208	0.500382434879534
10	0.865804492286815	0.499877559050176
11	0.866096083560039	0.500091082450647
12	0.865972810920378	0.499970850112656
13	0.866042232825645	0.500021687802653
14	0.866012881963649	0.499993059704778
15	0.866029410728674	0.500005163847862

Theorem 4.3 *Suppose that* $\boldsymbol{f}(\boldsymbol{\xi}) = \boldsymbol{0}$*, and that all the first partial derivatives of* $\boldsymbol{f} = (f_1, \ldots, f_n)^{\mathrm{T}}$ *are defined and continuous in some (open) neighbourhood of* $\boldsymbol{\xi}$*, and satisfy a condition of strict diagonal dominance at* $\boldsymbol{\xi}$*; i.e.,*

$$\frac{\partial f_i}{\partial x_i}(\boldsymbol{\xi}) > \sum_{\substack{j=1 \\ j \neq i}}^{n} \left| \frac{\partial f_i}{\partial x_j}(\boldsymbol{\xi}) \right|, \quad i = 1, 2, \ldots, n\,. \tag{4.17}$$

Then, there exist $\varepsilon > 0$ *and a positive constant* λ *such that the relaxation iteration (4.16) converges to* $\boldsymbol{\xi}$ *for any* $\boldsymbol{x}_0$ *in the closed ball* $\bar{B}_\varepsilon(\boldsymbol{\xi})$ *of radius* ε*, centre* $\boldsymbol{\xi}$*.*

Proof The elements of the Jacobian matrix $J_g(\boldsymbol{\xi}) = (\gamma_{ij}) \in \mathbb{R}^{n \times n}$ of the function $\boldsymbol{x} \mapsto \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x} - \lambda \boldsymbol{f}(\boldsymbol{x})$ at $\boldsymbol{x} = \boldsymbol{\xi}$ are

$$\gamma_{ii}(\boldsymbol{\xi}) = 1 - \lambda \frac{\partial f_i}{\partial x_i}(\boldsymbol{\xi})\,, \quad \gamma_{ij}(\boldsymbol{\xi}) = -\lambda \frac{\partial f_i}{\partial x_j}(\boldsymbol{\xi})\,, \quad j \neq i\,, \quad i, j \in \{1, \ldots, n\}\,.$$

We now define

$$m = \max_{i=1}^{n} \frac{\partial f_i}{\partial x_i}(\boldsymbol{\xi})$$

and then choose $\lambda = 1/m$. Under hypothesis (4.17), $m > 0$ and therefore $\lambda > 0$. This choice of λ ensures that all the diagonal elements $\gamma_{ii}(\boldsymbol{\xi})$, $i = 1, \ldots, n$, of $J_g(\boldsymbol{\xi})$ are nonnegative. Moreover, for any $i \in \{1, \ldots, n\}$,

$$\sum_{j=1}^{n} |\gamma_{ij}(\boldsymbol{\xi})| \quad = \quad 1 - \lambda \frac{\partial f_i}{\partial x_i}(\boldsymbol{\xi}) + \lambda \sum_{\substack{j=1 \\ j \neq i}}^{n} \left| \frac{\partial f_i}{\partial x_j}(\boldsymbol{\xi}) \right| < 1\,,$$

by condition (4.17); consequently, $\|J_g(\boldsymbol{\xi})\|_\infty < 1$. As $\boldsymbol{\xi}$ is a fixed point of $\boldsymbol{g}$, it follows from Theorem 4.2 that there exists $\varepsilon > 0$ such that the iteration (4.16) converges to $\boldsymbol{\xi}$ for all $\boldsymbol{x}^{(0)} \in \bar{B}_\varepsilon(\boldsymbol{\xi})$. □

The condition of strict diagonal dominance will only be satisfied in a small class of problems (although this class does contain some examples of practical importance). More generally it will be necessary to replace the scalar λ by a nonsingular constant matrix Λ, giving a more general relaxation iteration

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \Lambda \boldsymbol{f}(\boldsymbol{x}^{(k)})\,, \qquad k = 0, 1, 2, \ldots\,.$$

This may be interpreted as trying to solve the new system of equations $\Lambda \boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$. The Jacobian matrix of this system is ΛJ_f, where J_f is the Jacobian matrix of $\boldsymbol{f}$. It is now possible to select the matrix Λ so that $\Lambda J_f(\boldsymbol{\xi})$ has the property of strict diagonal dominance. In principle, this can obviously be done by choosing $\Lambda = [J_f(\boldsymbol{\xi})]^{-1}$, the inverse of the Jacobian matrix of $\boldsymbol{f}$ evaluated at the solution $\boldsymbol{\xi}$. The Jacobian matrix of the new system is then the identity matrix, which clearly satisfies the diagonal dominance condition. However, this choice is not possible in practice, since of course the solution $\boldsymbol{\xi}$ is unknown. If we allow the matrix Λ to be a function of $\boldsymbol{x}$, instead of being constant, the argument above suggests taking

$$\Lambda = [J_f(\boldsymbol{x}^{(k)})]^{-1}\,,$$

leading to Newton's method for a system of equations.

Definition 4.5 *The recursion defined by*

$$\boxed{\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - [J_f(\boldsymbol{x}^{(k)})]^{-1} \boldsymbol{f}(\boldsymbol{x}^{(k)})\,, \qquad k = 0, 1, 2, \ldots\,, \qquad (4.18)}$$

where $\boldsymbol{x}_0 \in \mathbb{R}^n$, is called **Newton's method** *(or Newton iteration) for*

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

the system of equations $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$. It is implicitly assumed that the matrix $J_f(\boldsymbol{x}^{(k)})$ exists and is nonsingular for each $k = 0, 1, 2, \ldots$.

The next theorem is concerned with the convergence of Newton's method. As in the scalar case, for a starting value $\boldsymbol{x}^{(0)}$ that is sufficiently close to the solution $\boldsymbol{\xi}$ of $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$, Newton's method converges quadratically. The precise definition of quadratic convergence is given below: it resembles Definition 1.7 of Chapter 1.

Definition 4.6 *Suppose that $(\boldsymbol{x}^{(k)})$ is a convergent sequence in $\mathbb{R}^n$ and $\boldsymbol{\xi} = \lim_{k\to\infty} \boldsymbol{x}^{(k)}$. We say that $(\boldsymbol{x}^{(k)})$ converges to $\boldsymbol{\xi}$* **with at least order** *$q > 1$, if there exist a sequence (ε_k) of positive real numbers converging to 0, and $\mu > 0$, such that*

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{\xi}\|_\infty \le \varepsilon_k\,, \quad k = 0, 1, 2, \ldots\,, \qquad \textit{and} \qquad \lim_{k\to\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k^q} = \mu\,. \tag{4.19}$$

If (4.19) holds with $\varepsilon_k = \|\boldsymbol{x}^{(k)} - \boldsymbol{\xi}\|_\infty$, $k = 0, 1, 2, \ldots$, then the sequence $(\boldsymbol{x}^{(k)})$ is said to converge to $\boldsymbol{\xi}$ **with order** *q. In particular, if $q = 2$, then we say that the sequence $(\boldsymbol{x}^{(k)})$ converges to $\boldsymbol{\xi}$* **quadratically**.

Again, due to (4.4), if a sequence $(\boldsymbol{x}^{(k)})$ converges quadratically in the ∞-norm, then it also does so in the p-norm for any $p \in [1, \infty)$, though the constant μ may be different.

Theorem 4.4 *Suppose that $\boldsymbol{f}(\boldsymbol{\xi}) = \mathbf{0}$, that in some (open) neighbourhood $N(\boldsymbol{\xi})$ of $\boldsymbol{\xi}$, where $\boldsymbol{f}$ is defined and continuous, all the second-order partial derivatives of $\boldsymbol{f}$ are defined and continuous, and that the Jacobian matrix $J_f(\boldsymbol{\xi})$ of $\boldsymbol{f}$ at the point $\boldsymbol{\xi}$ is nonsingular. Then, the sequence $(\boldsymbol{x}^{(k)})$ defined by Newton's method (4.18) converges to the solution $\boldsymbol{\xi}$ provided that $\boldsymbol{x}^{(0)}$ is sufficiently close to $\boldsymbol{\xi}$; the convergence of the sequence $(\boldsymbol{x}^{(k)})$ to $\boldsymbol{\xi}$ is at least quadratic.*

Proof Let us begin by writing Newton's method as a simultaneous iteration $\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)})$, $k = 0, 1, 2, \ldots$, as in (4.3), with $\boldsymbol{x}_0$ given and

$$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x} - [J_f(\boldsymbol{x})]^{-1} \boldsymbol{f}(\boldsymbol{x})\,.$$

The idea of the proof is to verify that the function $\boldsymbol{g}$ satisfies all the conditions of Theorem 4.2 in a certain closed ball centred at $\boldsymbol{\xi}$, the fixed point of $\boldsymbol{g}$, and thus deduce that the sequence $(\boldsymbol{x}^{(k)})$ converges to $\boldsymbol{\xi}$.

As the function $\boldsymbol{x} \mapsto \det J_f(\boldsymbol{x})$ is continuous in $N(\boldsymbol{\xi})$ and $\det J_f(\boldsymbol{\xi}) \neq 0$, there exists $\varepsilon > 0$ such that $\det J_f(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in \bar{B}_\varepsilon(\boldsymbol{\xi}) \subset N(\boldsymbol{\xi})$.

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

Further, as the entries of $[J_f(\boldsymbol{x})]^{-1}$ depend continuously on the entries of $J_f(\boldsymbol{x})$ and since the entries of $J_f(\,\cdot\,)$ are continuous functions of $\boldsymbol{x}$ in $N(\boldsymbol{\xi})$, we deduce that $\boldsymbol{x} \mapsto [J_f(\boldsymbol{x})]^{-1}\boldsymbol{f}(\boldsymbol{x})$ is a continuous function on $\bar{B}_\varepsilon(\boldsymbol{\xi})$; therefore,

$$\boldsymbol{x} \mapsto \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x} - [J_f(\boldsymbol{x})]^{-1}\boldsymbol{f}(\boldsymbol{x})$$

is also a continuous function on $\bar{B}_\varepsilon(\boldsymbol{\xi})$. For later reference, we note that $\boldsymbol{x} \mapsto \|[J_f(\boldsymbol{x})]^{-1}\|_\infty$, too, is a continuous function on $\bar{B}_\varepsilon(\boldsymbol{\xi})$, and therefore it is a bounded function on $\bar{B}_\varepsilon(\boldsymbol{\xi})$; we define

$$C = \max_{\boldsymbol{x} \in \bar{B}_\varepsilon(\boldsymbol{\xi})} \|[J_f(\boldsymbol{x})]^{-1}\|_\infty .$$

Now, $\boldsymbol{\xi}$ is a fixed point of $\boldsymbol{g}$ and, by the hypotheses of the theorem, the entries of the Jacobian matrix J_g of $\boldsymbol{g}$ are continuous functions of $\boldsymbol{x}$ on $\bar{B}_\varepsilon(\boldsymbol{\xi})$. Furthermore, it is easy to check that all the elements of the Jacobian matrix $J_g(\boldsymbol{x})$ of $\boldsymbol{g}$ vanish at $\boldsymbol{x} = \boldsymbol{\xi}$; see Exercise 6. Hence, $\|J_g(\boldsymbol{\xi})\|_\infty = 0 < 1$, trivially. Thus we have shown that $\boldsymbol{g}\colon \mathbb{R}^n \to \mathbb{R}^n$ satisfies all the conditions of Theorem 4.2 on the closed set $D = \bar{B}_\varepsilon(\boldsymbol{\xi})$, and the convergence of the sequence $(\boldsymbol{x}^{(k)})$ to $\boldsymbol{\xi}$, as $k \to \infty$, follows.

To show that convergence is at least quadratic, we write the iteration in the form

$$J_f(\boldsymbol{x}^{(k)})\,[\boldsymbol{x}^{(k+1)} - \boldsymbol{\xi}] = J_f(\boldsymbol{x}^{(k)})\,[\boldsymbol{x}^{(k)} - \boldsymbol{\xi}] - \boldsymbol{f}(\boldsymbol{x}^{(k)})\,. \tag{4.20}$$

Taylor's Theorem for a function of n variables, Theorem A.7 (including only the first-order terms), implies that, when $\boldsymbol{x}^{(k)} \in \bar{B}_\varepsilon(\boldsymbol{\xi})$,

$$\mathbf{0} = \boldsymbol{f}(\boldsymbol{\xi}) = \boldsymbol{f}(\boldsymbol{x}^{(k)}) + J_f(\boldsymbol{x}^{(k)})[\boldsymbol{\xi} - \boldsymbol{x}^{(k)}] + \mathbf{E}_f\,, \tag{4.21}$$

where

$$\|\mathbf{E}_f\|_\infty \le \tfrac{1}{2} n^2 A_f \|\boldsymbol{\xi} - \boldsymbol{x}^{(k)}\|_\infty^2\,, \tag{4.22}$$

and

$$A_f = \max_{1 \le i,j,l \le n} \max_{\boldsymbol{x} \in \bar{B}_\varepsilon(\boldsymbol{\xi})} \left| \frac{\partial^2 f_i}{\partial x_j \partial x_l}(\boldsymbol{x}) \right|$$

is a bound on all the second-order partial derivatives of $\boldsymbol{f}$ on $\bar{B}_\varepsilon(\boldsymbol{\xi})$. The factor n^2 in (4.22) stems from the fact that, for each $i \in \{1, \dots, n\}$, f_i is a function of n variables and therefore it has n^2 second-order partial derivatives – each bounded by A_f over $\bar{B}_\varepsilon(\boldsymbol{\xi})$. From (4.21) and (4.20) we see that

$$\boldsymbol{x}^{(k+1)} - \boldsymbol{\xi} = [J_f(\boldsymbol{x}^{(k)})]^{-1}\mathbf{E}_f\,,$$

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

and so

$$\|\boldsymbol{x}^{(k+1)} - \boldsymbol{\xi}\|_\infty \le \tfrac{1}{2} n^2 A_f\, C \|\boldsymbol{x}^{(k)} - \boldsymbol{\xi}\|_\infty^2 \,.$$

On writing $M = \frac{1}{2} n^2 A_f C$, we then deduce by induction that

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{\xi}\|_\infty \le \frac{1}{M} \left(M \|\boldsymbol{x}^{(0)} - \boldsymbol{\xi}\|_\infty \right)^{2^k}, \qquad k = 0, 1, 2, \dots \,.$$

Suppose that $\boldsymbol{x}^{(0)} \in \bar{B}_\varepsilon(\boldsymbol{\xi})$ where $\varepsilon \le \frac{1}{2} \min\{1, 1/M\}$. Then,

$$M \|\boldsymbol{x}^{(0)} - \boldsymbol{\xi}\|_\infty \le \frac{1}{2}\,, \qquad k = 0, 1, 2, \dots \,,$$

and hence

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{\xi}\|_\infty \le \frac{1}{M} \left(\frac{1}{2} \right)^{2^k}$$

This implies that convergence is at least quadratic (on choosing $\varepsilon_k = M^{-1} 2^{-2^k}$ and $q = 2$ in Definition 4.6). □

Newton's method is defined in (4.18) by using the inverse of the Jacobian matrix. As we saw in Chapter 2 it is more efficient to avoid inverting a matrix, if possible. In practice the method is therefore implemented by writing (4.18) in the form

$$J_f(\boldsymbol{x}^{(k)})[\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}] = -\boldsymbol{f}(\boldsymbol{x}^{(k)})\,. \tag{4.23}$$

Given the vector $\boldsymbol{x}^{(k)}$, we calculate $\boldsymbol{f}(\boldsymbol{x}^{(k)})$ and the Jacobian matrix $J_f(\boldsymbol{x}^{(k)}) \in \mathbb{R}^{n \times n}$, and then solve the system of linear equations (4.23) by Gaussian elimination; this gives the increment vector $\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}$, which is added to $\boldsymbol{x}^{(k)}$ to obtain the new iterate $\boldsymbol{x}^{(k+1)}$.

Example 4.5 *We close this section with an example which illustrates the application of Newton's method. Consider the simultaneous nonlinear equations*

$$\begin{aligned} f_1(x, y, z) &\equiv x^2 + y^2 + z^2 - 1 = 0\,, \\ f_2(x, y, z) &\equiv 2x^2 + y^2 - 4z = 0\,, \\ f_3(x, y, z) &\equiv 3x^2 - 4y + z^2 = 0\,. \end{aligned}$$

Letting $\boldsymbol{f} = (f_1, f_2, f_3)^{\mathrm{T}}$ *and* $\boldsymbol{x} = (x, y, z)^{\mathrm{T}}$, *the aim of the exercise is to determine the solution to the equation* $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$ *contained in the first octant* $\{(x, y, z) \in \mathbb{R}^3\colon x > 0\,,\ y > 0\,,\ z > 0\}$ *in* $\mathbb{R}^3$.

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

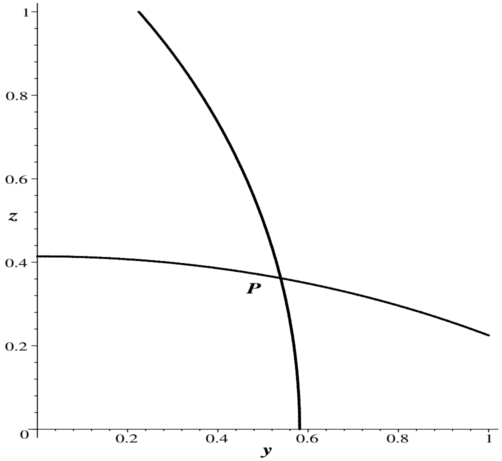

Fig. 4.3. Example 4.5: Projections onto the (y, z)-plane of the intersection-curves of the surfaces $f_1(x, y, z) = 0$ and $f_2(x, y, z) = 0$, and $f_1(x, y, z) = 0$ and $f_3(x, y, z) = 0$. The two curves intersect at the point P whose two coordinates are the y- and z-coordinates of $\boldsymbol{\xi}$, the solution of the system $f_1(x, y, z) = 0$, $f_2(x, y, z) = 0$, $f_3(x, y, z) = 0$.

Note that the Jacobian matrix of $\boldsymbol{f}$ at $\boldsymbol{x} \in \mathbb{R}^3$ is

$$J_f(\boldsymbol{x}) = \begin{pmatrix} 2x & 2y & 2z \\ 4x & 2y & -4 \\ 6x & -4 & 2z \end{pmatrix} .$$

Since the first equation represents a sphere of radius 1 centred at $(0, 0, 0)$, and the second and third equations describe elliptic paraboloids whose axes are aligned with the coordinate semi-axes $(0, 0, z)$, $z \geq 0$, and $(0, y, 0)$, $y \geq 0$, respectively, the point of intersection of the three surfaces belongs to $[0, 1]^3$. Let us denote this point by $\boldsymbol{\xi}$. In order to select a suitable starting value $\boldsymbol{x}^{(0)}$ for the iteration, we observe that the intersection of the first and the second surface is a curve whose projection onto the (y, z)-plane has the equation $y^2 + 2z^2 + 4z = 2$, while the intersection of the first and the third surface is a curve whose projection onto the (y, z)-plane has the equation $3y^2 + 4y + 2z^2 = 3$. The two curves are shown in Figure 4.3; the point P where the curves intersect has the same y- and z-coordinates as $\boldsymbol{\xi}$. The x-coordinate of $\boldsymbol{\xi}$ can be obtained from the first equation in terms of the y- and z-coordinates of P *via* $x = +(1 - y^2 - z^2)^{1/2}$. As the two coordinates of

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

P are, very roughly, $y \approx 0.5$ and $z \approx 0.5$, it is reasonable to choose as starting value for the Newton iteration the point $\boldsymbol{x}^{(0)} = (0.5, 0.5, 0.5)^{\mathrm{T}}$.

Thus, $\boldsymbol{f}(\boldsymbol{x}^{(0)}) = (-0.25, -1.25, -1.00)^{\mathrm{T}}$ and

$$J_f(\boldsymbol{x}^{(0)}) = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -4 \\ 3 & -4 & 1 \end{pmatrix} .$$

On solving the system of linear equations

$$J_f(\boldsymbol{x}^{(0)}) \left(\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}\right) = -\boldsymbol{f}(\boldsymbol{x}^{(0)})$$

for $\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}$, we find that $\boldsymbol{x}^{(1)} = (0.875, 0.500, 0.375)^{\mathrm{T}}$. Similarly,

$$\begin{aligned} \boldsymbol{x}^{(2)} &= (0.78981, 0.49662, 0.36993)^{\mathrm{T}} , \\ \boldsymbol{x}^{(3)} &= (0.78521, 0.49662, 0.36992)^{\mathrm{T}} . \end{aligned}$$

As $\boldsymbol{f}(\boldsymbol{x}^{(3)}) = 10^{-5}(1, 4, 5)^{\mathrm{T}}$, the vector $\boldsymbol{x}^{(3)}$ can be thought of as a satisfactory approximation to the required solution $\boldsymbol{\xi}$; after rounding to four decimal digits, we have that

$$x = 0.7852 , \qquad y = 0.4966 , \qquad z = 0.3699 .$$

◇

4.4 Global convergence

Much of the discussion of the global convergence of Newton's method for a single equation in Section 1.7 applies, with obvious changes, in the case of several variables. If the system has several solutions, $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, …, we can define the corresponding sets $S_1, S_2, \ldots$ in $\mathbb{R}^n$ so that S_j comprises those starting points from which Newton's method converges to $\boldsymbol{\xi}_j$. As before, the sets S_j, $j = 1, 2, \ldots$, have the property that any point on the boundary of one of the sets is also on the boundary of the others. The difference now is that for systems of equations in $\mathbb{R}^n$, $n \geq 2$, these sets can be much more complicated than in the case of a single equation on the real line $\mathbb{R}^1 = \mathbb{R}$.

To illustrate this point for $n = 2$, we return to our earlier example problem, Example 1.7 from Chapter 1, but now extend it to complex variables, so we require to solve $\mathrm{e}^z - z - 2 = 0$ for the complex number $z = x + \imath y$. Separating this equation into real and imaginary parts we obtain a system of two nonlinear equations for the unknowns $x_1 = x$ and $x_2 = y$. The system has the two real solutions which we found in

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

Chapter 1, and also an infinite number of complex solutions. It is easy to see from the periodic character of $e^{\imath y}$ that the equation has a solution near $w_m = (2m + \frac{1}{2})\imath\pi$, $\imath = \sqrt{-1}$, for integer values of m; a better estimate is given in Exercise 9. It is a good deal more difficult to prove that there are no other solutions.

The behaviour of Newton's method for this problem may be illustrated by showing a picture of the complex plane, with the sets S_j depicted in different colours. In our example we cannot, of course, show more than a small number of the solutions, and cannot use an infinite number of colours. We have therefore coloured the sets with six colours cyclically, so that, for example, the sets $S_1, S_7, S_{13}, \dots$ have the same colour. The background colour, white, represents the set S_1 of points from which the iteration converges to the real negative root. It includes most of the negative half-plane. Successive pictures in the series from Figure 4.5 to Figure 4.9 show a magnified view of a small region of the previous picture, the region being outlined in black. In Figure 4.4 the black crosses mark the positions of solutions of $f(z) = 0$. The pictures show in a striking way the fractal behaviour of the boundary of a set. Figure 4.9 is very similar to Figure 4.5; the former is a magnified view of a small part of Figure 4.5, with a magnification of about 50000 in each direction. The same sort of behaviour is repeated when the picture is magnified indefinitely.

4.5 Notes

For an introduction to the topology of $\mathbb{R}^n$, including the definitions of open set, closed set, continuity, convergence and Cauchy sequence, the reader is referred to any standard textbook on the subject; see, *e.g.*,

- W. Rudin, *Principles of Mathematical Analysis*, Third Edition, International Series in Pure and Applied Mathematics, McGraw–Hill, New York, Auckland, Düsseldorf, 1976,
- S.A. Douglass, *Introduction to Mathematical Analysis*, Addison–Wesley, Reading, MA, 1996.

Our first remark concerns the Contraction Mapping Theorem, Theorem 4.1, which is a direct generalisation of Theorem 1.3 from Chapter 1. Comparing the proofs of Theorems 1.3 and 4.1, we see that the proof of Theorem 1.3 is much simpler. This is not accidental: in the case of a single equation $x = g(x)$, involving a real-valued function g of a single real variable x, the existence of a fixed point follows directly from

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

Theorem 1.2, Brouwer's Fixed Point Theorem on a bounded closed interval of the real line. On the other hand, for the simultaneous system of equations $\boldsymbol{x} = \boldsymbol{g}(\boldsymbol{x})$ in $\mathbb{R}^n$ considered in Theorem 4.1 we had to invoke the completeness of $\mathbb{R}^n$ (i.e., the property that every Cauchy sequence in $\mathbb{R}^n$ is a convergent sequence) to show the existence of a fixed point. An alternative, shorter proof of Theorem 4.1 could have been devised by applying Brouwer's Fixed Point Theorem in $\mathbb{R}^n$.

Theorem 4.5 (**Brouwer's Fixed Point Theorem**) *Let us assume that D is a nonempty, closed, bounded and convex subset of $\mathbb{R}^n$. Suppose further that $\boldsymbol{g}\colon \mathbb{R}^n \mapsto \mathbb{R}^n$ is a continuous function defined on D such that $\boldsymbol{g}(D) \subset D$. Then, there exists $\boldsymbol{\xi} \in D$ such that $\boldsymbol{g}(\boldsymbol{\xi}) = \boldsymbol{\xi}$.*

A set $D \subset \mathbb{R}^n$ is said to be convex if, whenever $\boldsymbol{x}$ and $\boldsymbol{y}$ belong to D, also

$$\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y} \in D \qquad \forall\,\theta \in [0,1]\,.$$

For example, any nonempty interval of the real line $\mathbb{R}^1 = \mathbb{R}$ is a convex set, as is a nonempty (open or closed) ball in $\mathbb{R}^n$, $n \geq 2$. Unfortunately, when $n \geq 2$ the proof of Theorem 4.5 is nontrivial and is well beyond the scope of this book.[1]

Benoit Mandelbrot (1924–) has been largely responsible for the present interest in fractal geometry and its connections with iterative methods. Mandelbrot highlighted in his book

- B. Mandelbrot, *Fractals: Form, Chance, and Dimension*, W.H. Freeman, San Francisco, 1977,

and, more fully, in

- B. Mandelbrot, *The Fractal Geometry of Nature*, W.H. Freeman, New York, 1983,

the omnipresence of fractals both in mathematics and elsewhere in nature. In relation with the subject of this chapter, we note that the **Mandelbrot set** is a connected set of points in the complex plane defined as follows. Choose a point z_0 in the complex plane, and consider the iteration $z_{n+1} = z_n^2 + z_0$, $n = 0, 1, 2, \ldots$. If the sequence $z_0, z_1, z_2, \ldots$ remains within a distance of 2 from the origin for ever, then the point z_0

[1] For a proof of Theorem 4.5 in the case when D is a closed ball in $\mathbb{R}^n$, see John W. Milnor, *Topology from the Differentiable Viewpoint*, Princeton Landmarks in Mathematics, 1997.

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

is said to be in the Mandelbrot set. If the sequence diverges from the origin, then the point z_0 is not in the set.

A standard reference for theoretical results concerning the convergence of Newton's method in complete normed linear spaces is

- L.V. Kantorovich and G.P. Akilov, *Functional Analysis*, Second edition, Pergamon Press, Oxford, New York, 1982.

A further significant book in the area of iterative solution of systems of nonlinear equations is the text by

- J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Reprint of the 1970 original, Classics in Applied Mathematics, 30, SIAM, Philadelphia, 2000.

It gives a comprehensive treatment of the numerical solution of n nonlinear equations in n unknowns, covering asymptotic convergence results for a number of algorithms, including Newton's method, as well as existence theorems for solutions of nonlinear equations based on the use of topological degree theory and Brouwer's Fixed Point Theorem.

Exercises

4.1 Suppose that the function $\boldsymbol{g}$ is a contraction in the ∞-norm, as in (4.5). Use the fact that

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\|_p \le n^{1/p}\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\|_\infty$$

to show that $\boldsymbol{g}$ is a contraction in the p-norm if $L < n^{-1/p}$.

4.2 Show that the simultaneous equations $\boldsymbol{f}(x_1, x_2) = \mathbf{0}$, where $\boldsymbol{f} = (f_1, f_2)^{\mathrm{T}}$, with

$$f_1(x_1, x_2) = x_1^2 + x_2^2 - 25\,, \qquad f_2(x_1, x_2) = x_1 - 7x_2 - 25\,,$$

have two solutions, one of which is $x_1 = 4$, $x_2 = -3$, and find the other. Show that the function $\boldsymbol{f}$ does not satisfy the conditions of Theorem 4.3 at either of these solutions, but that if the sign of f_2 is changed the conditions are satisfied at one solution, and that if $\boldsymbol{f}$ is replaced by $\boldsymbol{f}^* = (f_2 - f_1, -f_2)^{\mathrm{T}}$, then the conditions are satisfied at the other. In each case, give a value of the relaxation parameter λ which will lead to convergence.

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

4.3 The complex-valued function $z \mapsto g(z)$ of the complex variable z is holomorphic in a convex region Ω containing the point ζ, at which $g(\zeta) = \zeta$. By applying the Mean Value Theorem (Theorem A.3) to the function φ of the real variable t defined by $\varphi(t) = g((1-t)u + tv)$ show that if u and v lie in Ω, then there is a complex number η in Ω such that

$$g(u) - g(v) = (u - v)g'(\eta)\,.$$

Hence show that if $|g'(\zeta)| < 1$, then the complex iteration defined by $z_{k+1} = g(z_k)$, $k = 0, 1, 2, \ldots$, converges to ζ provided that z_0 is sufficiently close to ζ.

4.4 Suppose that in Exercise 3 the real and imaginary parts of g are u and v, so that $g(x + \imath y) = u(x, y) + \imath v(x, y)$, $\imath = \sqrt{-1}$. Show that the iteration defined by $\boldsymbol{x}^{(k+1)} = \boldsymbol{g}^*(\boldsymbol{x}^{(k)})$, $k = 0, 1, 2, \ldots$, where $\boldsymbol{g}^*(\boldsymbol{x}) = (u(x_1, x_2), v(x_1, x_2))^{\mathrm{T}}$, generates the real and imaginary parts of the sequence defined in Exercise 3. Compare the condition for convergence given in that exercise with the sufficient condition given by Theorem 4.2.

4.5 Verify that the iteration $\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)})$, $k = 0, 1, 2, \ldots$, where $\boldsymbol{g} = (g_1, g_2)^{\mathrm{T}}$ and g_1 and g_2 are functions of two variables defined by

$$g_1(x_1, x_2) = \tfrac{1}{3}(x_1^2 - x_2^2 + 3)\,, \quad g_2(x_1, x_2) = \tfrac{1}{3}(2x_1x_2 + 1)\,,$$

has the fixed point $\boldsymbol{x} = (1, 1)^{\mathrm{T}}$. Show that the function $\boldsymbol{g}$ does not satisfy the conditions of Theorem 4.3. By applying the results of Exercises 3 and 4 to the complex function g defined by

$$g(z) = \tfrac{1}{3}(z^2 + 3 + \imath)\,, \qquad z \in \mathbb{C}\,, \quad \imath = \sqrt{-1}\,,$$

show that the iteration, nevertheless, converges.

4.6 Suppose that all the second-order partial derivatives of the function $\boldsymbol{f}\colon \mathbb{R}^n \to \mathbb{R}^n$ are defined and continuous in a neighbourhood of the point $\boldsymbol{\xi}$ in $\mathbb{R}^n$, at which $\boldsymbol{f}(\boldsymbol{\xi}) = \boldsymbol{0}$. Assume also that the Jacobian matrix, $J_f(\boldsymbol{x})$, of $\boldsymbol{f}$ is nonsingular at $\boldsymbol{x} = \boldsymbol{\xi}$, and denote its inverse by $K(\boldsymbol{x})$ at all $\boldsymbol{x}$ for which it exists. Defining the Newton iteration by $\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)})$, $k = 0, 1, 2, \ldots$, with $\boldsymbol{x}_0$ given, where $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x} - K(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})$, show that the (i, j)-entry

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

of the Jacobian matrix $J_g(\boldsymbol{x}) \in \mathbb{R}^{n\times n}$ of $\boldsymbol{g}$ is

$$\delta_{ij} - \sum_{r=1}^{k} \frac{\partial K_{ir}}{\partial x_j} f_r - \sum_{r=1}^{k} K_{ir} J_{rj}\,, \qquad i,j = 1,\dots,n\,,$$

where J_{rj} is the (r,j)-entry of $J_f(\boldsymbol{x})$. Deduce that all the elements of this matrix vanish at the point $\boldsymbol{\xi}$.

4.7 The vector function $\boldsymbol{x} \mapsto \boldsymbol{f}(\boldsymbol{x})$ of two variables is defined by

$$f_1(x_1,x_2) = x_1^2 + x_2^2 - 2\,, \qquad f_2(x_1,x_2) = x_1 - x_2\,.$$

Verify that the equation $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$ has two solutions, $x_1 = x_2 = 1$ and $x_1 = x_2 = -1$. Show that one iteration of Newton's method for the solution of this system gives $\boldsymbol{x}^{(1)} = (x_1^{(1)}, x_2^{(1)})^{\mathrm{T}}$, with

$$x_1^{(1)} = x_2^{(1)} = \frac{\left(x_1^{(0)}\right)^2 + \left(x_2^{(0)}\right)^2 + 2}{2\left(x_1^{(0)} + x_2^{(0)}\right)}\,.$$

Deduce that the iteration converges to $(1,1)^{\mathrm{T}}$ if $x_1^{(0)} + x_2^{(0)}$ is positive, and, if $x_1^{(0)} + x_2^{(0)}$ is negative, the iteration converges to the other solution. Verify that convergence is quadratic.

4.8 Suppose that $\boldsymbol{\xi} = \lim_{k\to\infty} \boldsymbol{x}^{(k)}$ in $\mathbb{R}^n$. Following Definition 1.4, explain what is meant by saying that *the sequence* $(\boldsymbol{x}^{(k)})$ *converges to* $\boldsymbol{\xi}$ *linearly, with asymptotic rate* $-\log_{10}\mu$, where $0 < \mu < 1$.

Given the vector function $\boldsymbol{x} \mapsto \boldsymbol{f}(\boldsymbol{x})$ of two real variables x_1 and x_2 defined by

$$f_1(x_1,x_2) = x_1^2 + x_2^2 - 2\,, \qquad f_2(x_1,x_2) = x_1 + x_2 - 2\,,$$

show that $\boldsymbol{f}(\boldsymbol{\xi}) = \mathbf{0}$ when $\boldsymbol{\xi} = (1,1)^{\mathrm{T}}$. Suppose that $x_1^{(0)} \neq x_2^{(0)}$; show that one iteration of Newton's method for the solution of $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$ with starting value $\boldsymbol{x}^{(0)} = (x_1^{(0)}, x_2^{(0)})^{\mathrm{T}}$ then gives $\boldsymbol{x}^{(1)} = (x_1^{(1)}, x_2^{(1)})^{\mathrm{T}}$ such that $x_1^{(1)} + x_2^{(1)} = 2$. Determine $\boldsymbol{x}^{(1)}$ when

$$x_1^{(0)} = 1 + \alpha\,,\ x_2^{(0)} = 1 - \alpha\,,$$

where $\alpha \neq 0$. Assuming that $x_1^{(0)} \neq x_2^{(0)}$, deduce that Newton's method converges linearly to $(1,1)^{\mathrm{T}}$, with asymptotic rate of convergence $\log_{10} 2$. Why is the convergence not quadratic?

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

4.9 Suppose that the equation $\mathrm{e}^z = z + 2$, $z \in \mathbb{C}$, has a solution

$$z = (2m + \tfrac{1}{2})\imath\pi + \ln[(2m + \tfrac{1}{2})\pi] + \eta\,,$$

where m is a positive integer and $\imath = \sqrt{-1}$. Show that

$$\eta = \ln[1 - \imath(\ln(2m + \tfrac{1}{2})\pi + \eta + 2)/(2m + \tfrac{1}{2}\pi)]$$

and deduce that $\eta = \mathcal{O}(\ln\, m/m)$ for large m.
(Note that $|\ln(1 + \imath t)| < |t|$ for all $t \in \mathbb{R} \setminus \{0\}$.)

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

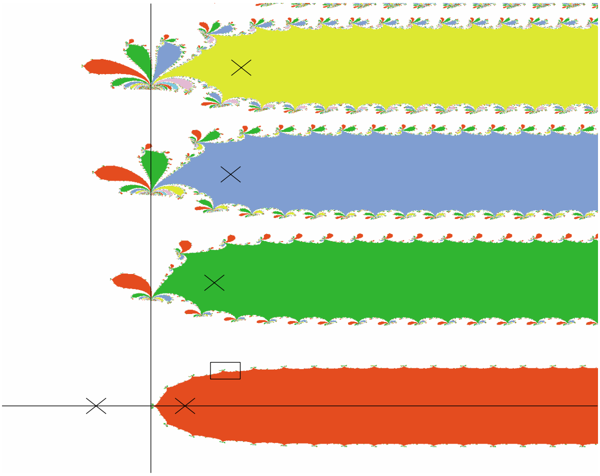

Fig. 4.4. The sets S_k in the region $-5 \le x \le 15$, $-4 \le y \le 24$ of the complex plane.

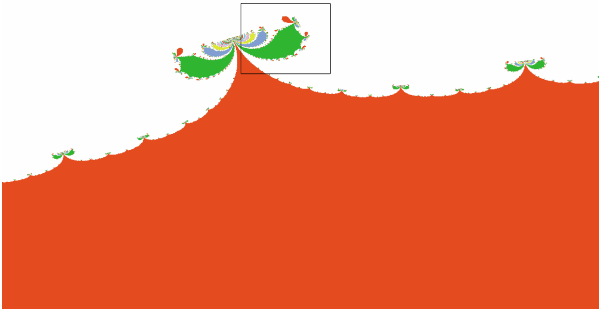

Fig. 4.5. The sets S_k in the region $2 \le x \le 3$, $1.6 \le y \le 2.6$ of the complex plane.

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

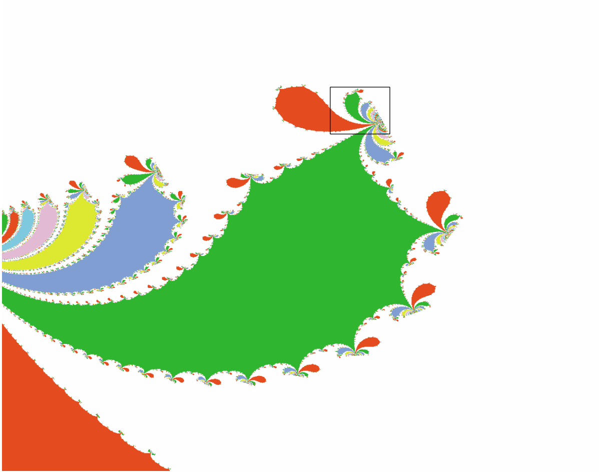

Fig. 4.6. The sets S_k in the region $2.4 \leq x \leq 2.55$, $2.1 \leq y \leq 2.25$ of the complex plane.

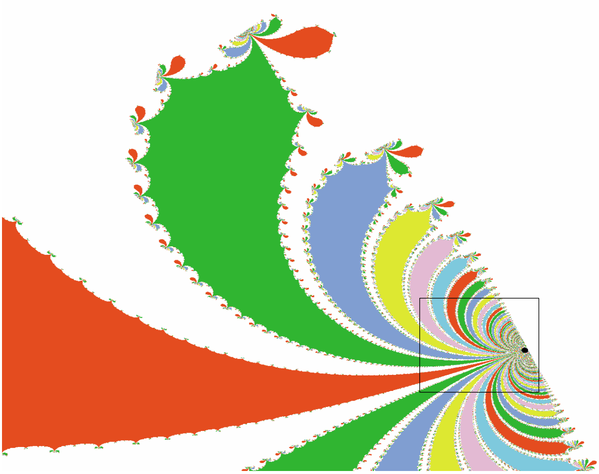

Fig. 4.7. The sets S_k in the region $2.4825 \leq x \leq 2.4975$, $2.2075 \leq y \leq 2.2225$ of the complex plane.

https://doi.org/10.1017/CBO9780511801181.005 Published online by Cambridge University Press

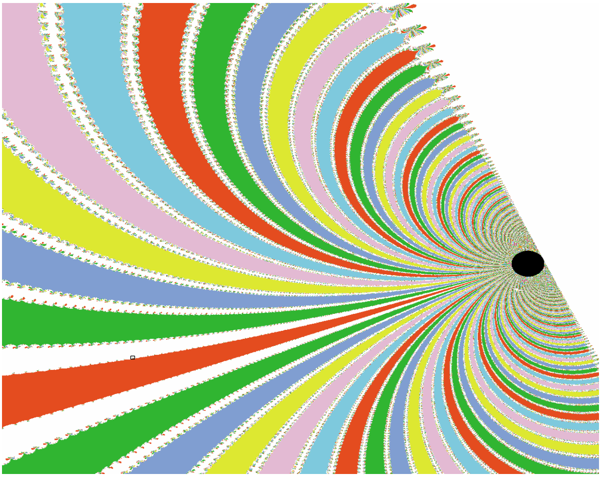

Fig. 4.8. The sets S_k in the region $2.4930 \le x \le 2.4960$, $2.2100 \le y \le 2.2130$ of the complex plane.

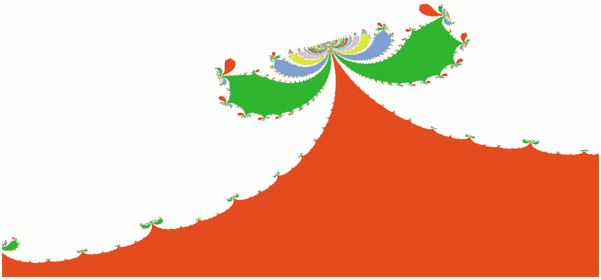

Fig. 4.9. The sets S_k in the region $2.493645 \le x \le 2.493665$, $2.21073 \le y \le 2.21075$ of the complex plane.