MAP 2220 – FUNDAMENTOS DE ANÁLISE NUMÉRICA

2º Semestre - 2017

Prof. Dr. Luis Carlos de Castro Santos

lsantos@ime.usp.br/lccs13@yahoo.com

10 Numerical Solutions of Nonlinear Systems of Equations 629

CHAPTER

10 Numerical Solutions of Nonlinear Systems of Equations

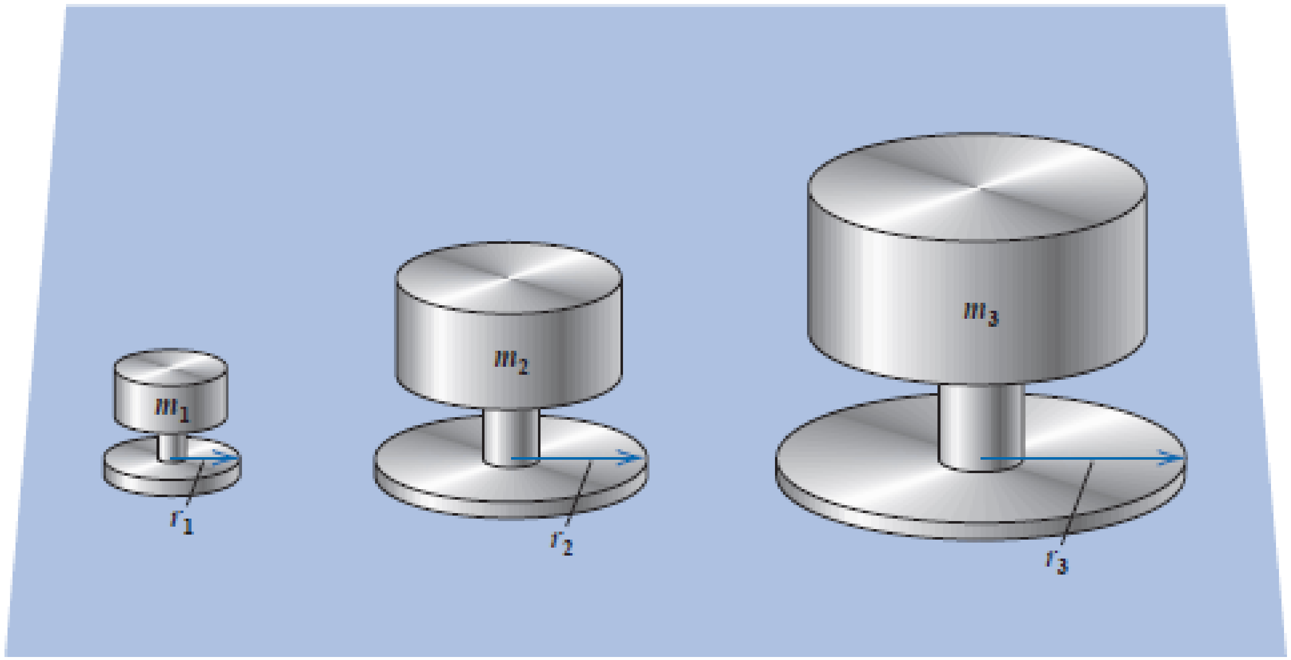
Introduction

The amount of pressure required to sink a large heavy object into soft, homogeneous soil lying above a hard base soil can be predicted by the amount of pressure required to sink smaller objects in the same soil. Specifically, the amount of pressure p to sink a circular plate of radius r a distance d in the soft soil, where the hard base soil lies a distance $D > d$ below the surface, can be approximated by an equation of the form

$$p = k_1 e^{k_2 r} + k_3 r,$$

where k_1, k_2, and k_3 are constants depending on d and the consistency of the soil, but not on the radius of the plate.

There are three unknown constants in this equation, so three small plates with differing radii are sunk to the same distance. This will determine the minimal size plate required to sustain a large load. The loads required for this sinkage are recorded, as shown in the accompanying figure.

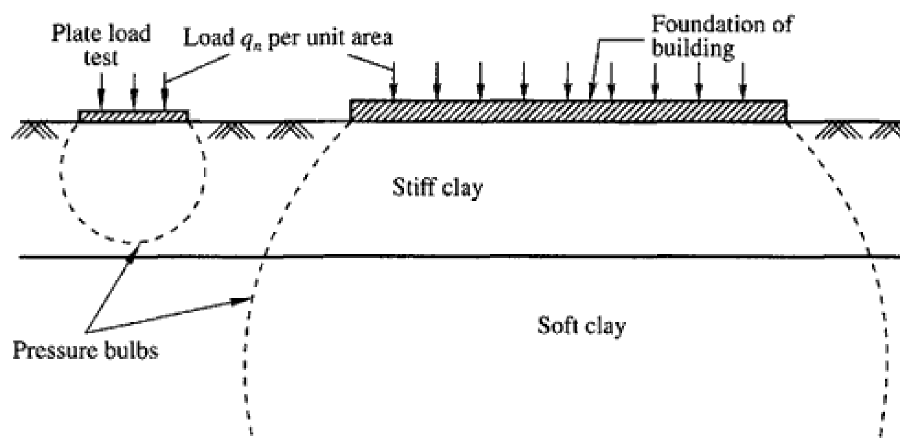


This produces the three nonlinear equations

$$m_1 = k_1 e^{k_2 r_1} + k_3 r_1,$$
$$m_2 = k_1 e^{k_2 r_2} + k_3 r_2,$$
$$m_3 = k_1 e^{k_2 r_3} + k_3 r_3,$$

in the three unknowns k_1, k_2, and k_3. Numerical approximation methods are usually needed for solving systems of equations when the equations are nonlinear. Exercise 12 of Section 10.2 concerns an application of the type described here.

Busca: Soil Bearing Capacity Test

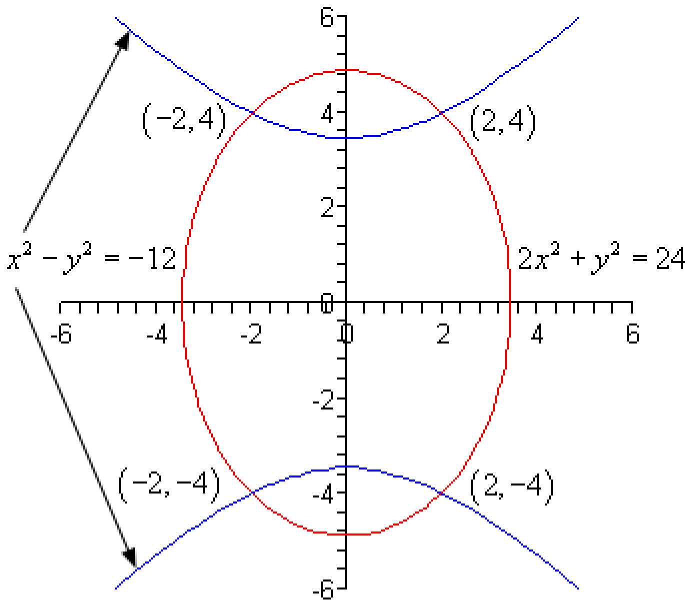


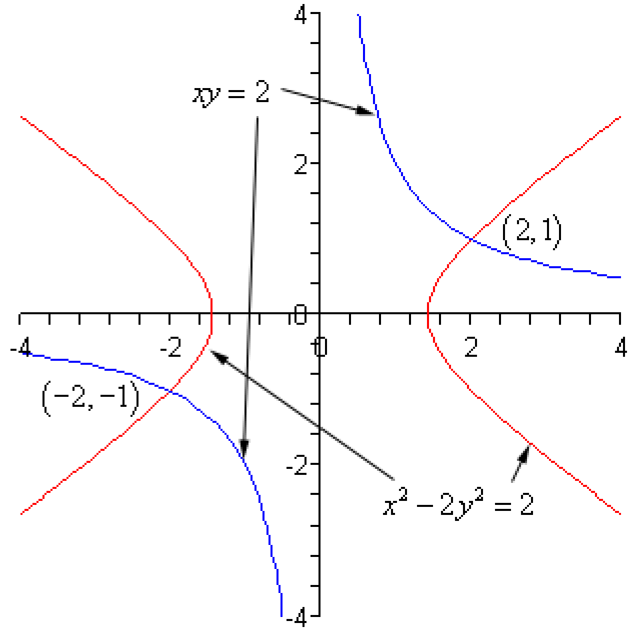
Solving a system of nonlinear equations is a problem that is avoided when possible, customarily by approximating the nonlinear system by a system of linear equations. When this is unsatisfactory, the problem must be tackled directly. The most straightforward approach is to adapt the methods from Chapter 2, which approximate the solutions of a single nonlinear equation in one variable, to apply when the single-variable problem is replaced by a vector problem that incorporates all the variables.

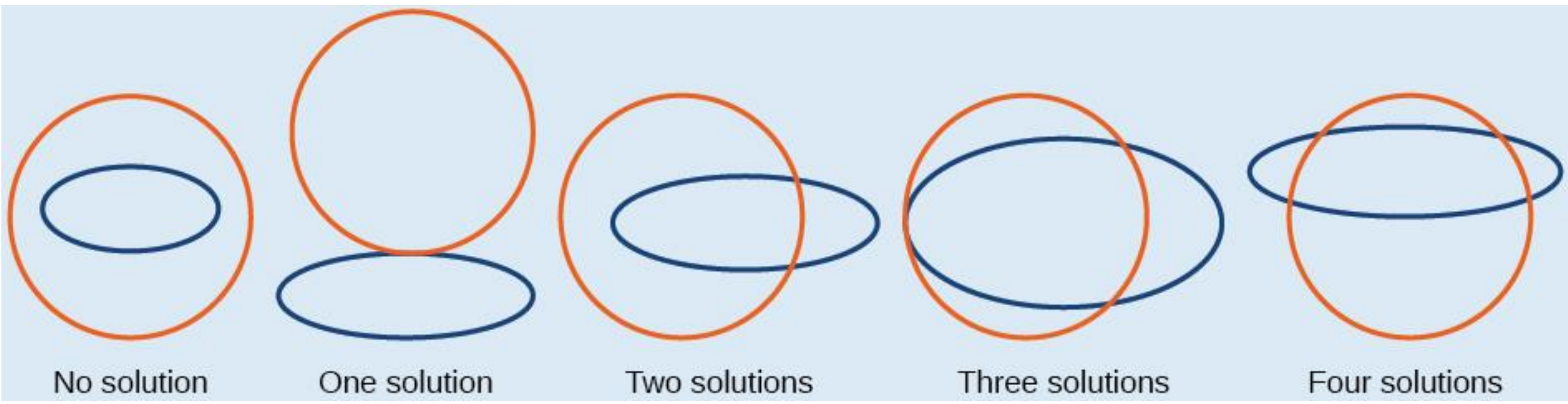
The principal tool in Chapter 2 was Newton's method, a technique that is generally quadratically convergent. This is the first technique we modify to solve systems of nonlinear equations. Newton's method, as modified for systems of equations, is quite costly to apply, and in Section 10.3 we describe how a modified Secant method can be used to obtain approximations more easily, although with a loss of the extremely rapid convergence that Newton's method can produce.

Section 10.4 describes the method of Steepest Descent. It is only linearly convergent, but it does not require the accurate starting approximations needed for more rapidly converging techniques. It is often used to find a good initial approximation for Newton's method or one of its modifications.

In Section 10.5, we give an introduction to continuation methods, which use a parameter to move from a problem with an easily determined solution to the solution of the original nonlinear problem.


$x^2 - y^2 = -12$
$2x^2 + y^2 = 24$
$(-2,4)$
$(2,4)$
$(-2,-4)$
$(2,-4)$


$xy = 2$
$(2,1)$
$(-2,-1)$
$x^2 - 2y^2 = 2$


No solution
One solution
Two solutions
Three solutions
Four solutions

10.1 Fixed Points for Functions of Several Variables

A system of nonlinear equations has the form

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0, \\ f_2(x_1, x_2, \ldots, x_n) &= 0, \\ \vdots \qquad & \quad \vdots \\ f_n(x_1, x_2, \ldots, x_n) &= 0, \end{aligned} \tag{10.1}$$

where each function f_i can be thought of as mapping a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ of the n-dimensional space $\mathbb{R}^n$ into the real line $\mathbb{R}$. A geometric representation of a nonlinear system when $n = 2$ is given in Figure 10.1.

This system of n nonlinear equations in n unknowns can also be represented by defining a function $\mathbf{F}$ mapping $\mathbb{R}^n$ into $\mathbb{R}^n$ as

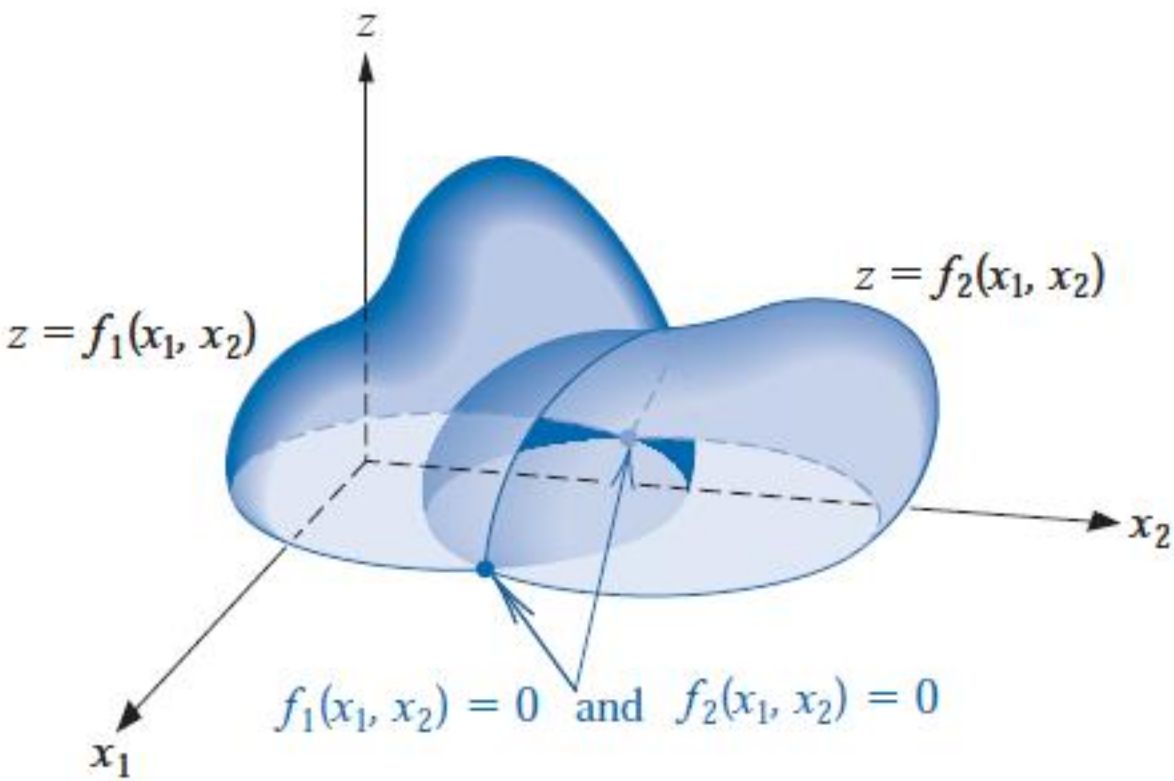
$$\mathbf{F}(x_1, x_2, \ldots, x_n) = (f_1(x_1, x_2, \ldots, x_n), f_2(x_1, x_2, \ldots, x_n), \ldots, f_n(x_1, x_2, \ldots, x_n))^t.$$

If vector notation is used to represent the variables $x_1, x_2, \ldots, x_n$, then system (10.1) assumes the form

$$\mathbf{F}(\mathbf{x}) = \mathbf{0}. \tag{10.2}$$

The functions $f_1, f_2, \ldots, f_n$ are called the **coordinate functions** of $\mathbf{F}$.

Figure 10.1



Example 1 Place the 3×3 nonlinear system

$$
\begin{aligned}
3x_1 - \cos(x_2x_3) - \frac{1}{2} &= 0,\\
x_1^2 - 81(x_2+0.1)^2 + \sin x_3 + 1.06 &= 0,\\
e^{-x_1x_2} + 20x_3 + \frac{10\pi-3}{3} &= 0
\end{aligned}
$$

in the form (10.2).

Solution Define the three coordinate functions f_1, f_2, and f_3 from $\mathbb{R}^3$ to $\mathbb{R}$ as

$$
\begin{aligned}
f_1(x_1,x_2,x_3) &= 3x_1 - \cos(x_2x_3) - \frac{1}{2},\\
f_2(x_1,x_2,x_3) &= x_1^2 - 81(x_2+0.1)^2 + \sin x_3 + 1.06,\\
f_3(x_1,x_2,x_3) &= e^{-x_1x_2} + 20x_3 + \frac{10\pi-3}{3},
\end{aligned}
$$

Then define $\mathbf{F}$ from $\mathbb{R}^3 \to \mathbb{R}^3$ by

$$
\begin{aligned}
\mathbf{F}(\mathbf{x}) &= \mathbf{F}(x_1,x_2,x_3)\\
&= (f_1(x_1,x_2,x_3), f_2(x_1,x_2,x_3), f_3(x_1,x_2,x_3))^t\\
&= \left(3x_1 - \cos(x_2x_3) - \frac{1}{2}, x_1^2 - 81(x_2+0.1)^2\right.\\
&\qquad \left. + \sin x_3 + 1.06, e^{-x_1x_2} + 20x_3 + \frac{10\pi-3}{3}\right)^t.
\end{aligned}
$$

■

Definition 10.1 Let f be a function defined on a set $D \subset \mathbb{R}^n$ and mapping into $\mathbb{R}$. The function f is said to have the **limit** L at $\mathbf{x}_0$, written

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = L,$$

if, given any number $\varepsilon > 0$, a number $\delta > 0$ exists with

$$|f(\mathbf{x}) - L| < \varepsilon,$$

whenever $\mathbf{x} \in D$ and

$$0 < ||\mathbf{x} - \mathbf{x}_0|| < \delta.$$

The existence of a limit is also independent of the particular vector norm being used, as discussed in Section 7.1. Any convenient norm can be used to satisfy the condition in this definition. The specific value of δ will depend on the norm chosen, but the existence of a δ is independent of the norm.

The notion of a limit permits us to define continuity for functions from $\mathbb{R}^n$ into $\mathbb{R}$. Although various norms can be used, continuity is independent of the particular choice.

Definition 10.2

Let f be a function from a set $D \subset \mathbb{R}^n$ into $\mathbb{R}$. The function f is **continuous** at $\mathbf{x}_0 \in D$ provided $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ exists and

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

Moreover, f is **continuous** on a set D if f is continuous at every point of D. This concept is expressed by writing $f \in C(D)$. ■

We can now define the limit and continuity concepts for functions from $\mathbb{R}^n$ into $\mathbb{R}^n$ by considering the coordinate functions from $\mathbb{R}^n$ into $\mathbb{R}$.

Definition 10.3 Let $\mathbf{F}$ be a function from $D \subset \mathbb{R}^n$ into $\mathbb{R}^n$ of the form

$$\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_n(\mathbf{x}))^t,$$

where f_i is a mapping from $\mathbb{R}^n$ into $\mathbb{R}$ for each i. We define

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{F}(\mathbf{x}) = \mathbf{L} = (L_1, L_2, \ldots, L_n)^t,$$

if and only if $\lim_{\mathbf{x} \to \mathbf{x}_0} f_i(\mathbf{x}) = L_i$, for each $i = 1, 2, \ldots, n$. ■

The function $\mathbf{F}$ is **continuous** at $\mathbf{x}_0 \in D$ provided $\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{F}(\mathbf{x})$ exists and $\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_0)$. In addition, $\mathbf{F}$ is continuous on the set D if $\mathbf{F}$ is continuous at each $\mathbf{x}$ in D. This concept is expressed by writing $\mathbf{F} \in C(D)$.

For functions from $\mathbb{R}$ into $\mathbb{R}$, continuity can often be shown by demonstrating that the function is differentiable (see Theorem 1.6). Although this theorem generalizes to functions of several variables, the derivative (or total derivative) of a function of several variables is quite involved and will not be presented here. Instead we state the following theorem, which relates the continuity of a function of n variables at a point to the partial derivatives of the function at the point.

Theorem 10.4 Let f be a function from $D \subset \mathbb{R}^n$ into $\mathbb{R}$ and $\mathbf{x}_0 \in D$. Suppose that all the partial derivatives of f exist and constants $\delta > 0$ and $K > 0$ exist so that whenever $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ and $\mathbf{x} \in D$, we have

$$\left| \frac{\partial f(\mathbf{x})}{\partial x_j} \right| \le K, \quad \text{for each } j = 1, 2, \ldots, n.$$

Then f is continuous at $\mathbf{x}_0$. ■

Fixed Points in $\mathbb{R}^n$

In Chapter 2, an iterative process for solving an equation $f(x) = 0$ was developed by first transforming the equation into the fixed-point form $x = g(x)$. A similar procedure will be investigated for functions from $\mathbb{R}^n$ into $\mathbb{R}^n$.

Definition 10.5 A function $\mathbf{G}$ from $D \subset \mathbb{R}^n$ into $\mathbb{R}^n$ has a **fixed point** at $\mathbf{p} \in D$ if $\mathbf{G}(\mathbf{p}) = \mathbf{p}$. ■

The following theorem extends the Fixed-Point Theorem 2.4 on page 62 to the n-dimensional case. This theorem is a special case of the Contraction Mapping Theorem, and its proof can be found in [Or2], p. 153.

Theorem 10.6 Let $D = \{ (x_1, x_2, \ldots, x_n)^t \mid a_i \leq x_i \leq b_i$, for each $i = 1, 2, \ldots, n \}$ for some collection of constants $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$. Suppose $\mathbf{G}$ is a continuous function from $D \subset \mathbb{R}^n$ into $\mathbb{R}^n$ with the property that $\mathbf{G}(\mathbf{x}) \in D$ whenever $\mathbf{x} \in D$. Then $\mathbf{G}$ has a fixed point in D.

Moreover, suppose that all the component functions of $\mathbf{G}$ have continuous partial derivatives and a constant $K < 1$ exists with

$$\left| \frac{\partial g_i(\mathbf{x})}{\partial x_j} \right| \leq \frac{K}{n}, \quad \text{whenever } \mathbf{x} \in D,$$

for each $j = 1, 2, \ldots, n$ and each component function g_i. Then the sequence $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ defined by an arbitrarily selected $\mathbf{x}^{(0)}$ in D and generated by

$$\mathbf{x}^{(k)} = G(\mathbf{x}^{(k-1)}), \quad \text{for each } k \geq 1$$

converges to the unique fixed point $\mathbf{p} \in D$ and

$$\left\| \mathbf{x}^{(k)} - \mathbf{p} \right\|_\infty \leq \frac{K^k}{1-K} \left\| \mathbf{x}^{(1)} - \mathbf{x}^{(0)} \right\|_\infty . \tag{10.3}$$

■

Example 2 Place the nonlinear system

$$3x_1 - \cos(x_2 x_3) - \frac{1}{2} = 0,$$
$$x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06 = 0,$$
$$e^{-x_1 x_2} + 20x_3 + \frac{10\pi - 3}{3} = 0.$$

in a fixed-point form $\mathbf{x} = \mathbf{G}(\mathbf{x})$ by solving the ith equation for x_i, show that there is a unique solution on

$$D = \{ (x_1, x_2, x_3)^t \mid -1 \le x_i \le 1, \quad \text{for each } i = 1, 2, 3\}.$$

and iterate starting with $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$ until accuracy within 10^{-5} in the l_∞ norm is obtained.

Solution Solving the ith equation for x_i gives the fixed-point problem

$$
\begin{aligned}
x_1 &= \frac{1}{3}\cos(x_2x_3) + \frac{1}{6},\\
x_2 &= \frac{1}{9}\sqrt{x_1^2 + \sin x_3 + 1.06} - 0.1, \qquad (10.4)\\
x_3 &= -\frac{1}{20}e^{-x_1x_2} - \frac{10\pi - 3}{60}.
\end{aligned}
$$

Let $\mathbf{G} : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $\mathbf{G}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), g_3(\mathbf{x}))^t$, where

$$
\begin{aligned}
g_1(x_1, x_2, x_3) &= \frac{1}{3}\cos(x_2x_3) + \frac{1}{6},\\
g_2(x_1, x_2, x_3) &= \frac{1}{9}\sqrt{x_1^2 + \sin x_3 + 1.06} - 0.1,\\
g_3(x_1, x_2, x_3) &= -\frac{1}{20}e^{-x_1x_2} - \frac{10\pi - 3}{60}.
\end{aligned}
$$

Theorems 10.4 and 10.6 will be used to show that $\mathbf{G}$ has a unique fixed point in

$$D = \{\, (x_1, x_2, x_3)^t \mid -1 \le x_i \le 1, \quad \text{for each } i = 1, 2, 3\}.$$

For $\mathbf{x} = (x_1, x_2, x_3)^t$ in D,

$$|g_1(x_1, x_2, x_3)| \leq \frac{1}{3}|\cos(x_2 x_3)| + \frac{1}{6} \leq 0.50,$$

$$|g_2(x_1, x_2, x_3)| = \left|\frac{1}{9}\sqrt{x_1^2 + \sin x_3 + 1.06} - 0.1\right| \leq \frac{1}{9}\sqrt{1 + \sin 1 + 1.06} - 0.1 < 0.09,$$

and

$$|g_3(x_1, x_2, x_3)| = \frac{1}{20}e^{-x_1 x_2} + \frac{10\pi - 3}{60} \leq \frac{1}{20}e + \frac{10\pi - 3}{60} < 0.61.$$

So we have, for each $i = 1, 2, 3,$

$$-1 \leq g_i(x_1, x_2, x_3) \leq 1.$$

Thus $\mathbf{G}(\mathbf{x}) \in D$ whenever $\mathbf{x} \in D$.

Finding bounds for the partial derivatives on D gives

$$\left|\frac{\partial g_1}{\partial x_1}\right| = 0, \quad \left|\frac{\partial g_2}{\partial x_2}\right| = 0, \quad \text{and} \quad \left|\frac{\partial g_3}{\partial x_3}\right| = 0,$$

as well as

$$\left|\frac{\partial g_1}{\partial x_2}\right| \le \frac{1}{3}|x_3| \cdot |\sin x_2 x_3| \le \frac{1}{3}\sin 1 < 0.281, \quad \left|\frac{\partial g_1}{\partial x_3}\right| \le \frac{1}{3}|x_2| \cdot |\sin x_2 x_3| \le \frac{1}{3}\sin 1 < 0.281,$$

$$\left|\frac{\partial g_2}{\partial x_1}\right| = \frac{|x_1|}{9\sqrt{x_1^2 + \sin x_3 + 1.06}} < \frac{1}{9\sqrt{0.218}} < 0.238,$$

$$\left|\frac{\partial g_2}{\partial x_3}\right| = \frac{|\cos x_3|}{18\sqrt{x_1^2 + \sin x_3 + 1.06}} < \frac{1}{18\sqrt{0.218}} < 0.119,$$

$$\left|\frac{\partial g_3}{\partial x_1}\right| = \frac{|x_2|}{20}e^{-x_1 x_2} \le \frac{1}{20}e < 0.14, \quad \text{and} \quad \left|\frac{\partial g_3}{\partial x_2}\right| = \frac{|x_1|}{20}e^{-x_1 x_2} \le \frac{1}{20}e < 0.14.$$

The partial derivatives of g_1, g_2, and g_3 are all bounded on D, so Theorem 10.4 implies that these functions are continuous on D. Consequently, $\mathbf{G}$ is continuous on D. Moreover, for every $\mathbf{x} \in D$,

$$\left|\frac{\partial g_i(\mathbf{x})}{\partial x_j}\right| \leq 0.281, \quad \text{for each } i = 1, 2, 3 \quad \text{and} \quad j = 1, 2, 3,$$

and the condition in the second part of Theorem 10.6 holds with $K = 3(0.281) = 0.843$.

In the same manner it can also be shown that $\partial g_i/\partial x_j$ is continuous on D for each $i = 1, 2, 3$ and $j = 1, 2, 3$. (This is considered in Exercise 3.) Consequently, $\mathbf{G}$ has a unique fixed point in D, and the nonlinear system has a solution in D.

Note that $\mathbf{G}$ having a unique fixed point in D does not imply that the solution to the original system is unique on this domain, because the solution for x_2 in (10.4) involved the choice of the principal square root. Exercise 7(d) examines the situation that occurs if the negative square root is instead chosen in this step.

To approximate the fixed point $\mathbf{p}$, we choose $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$. The sequence of vectors generated by

$$x_1^{(k)} = \frac{1}{3}\cos x_2^{(k-1)} x_3^{(k-1)} + \frac{1}{6},$$

$$x_2^{(k)} = \frac{1}{9}\sqrt{\left(x_1^{(k-1)}\right)^2 + \sin x_3^{(k-1)} + 1.06} - 0.1,$$

$$x_3^{(k)} = -\frac{1}{20}e^{-x_1^{(k-1)}x_2^{(k-1)}} - \frac{10\pi - 3}{60}$$

converges to the unique solution of the system in (10.4). The results in Table 10.1 were generated until

$$\left\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\right\|_\infty < 10^{-5}.$$

Table 10.1

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\left\lVert \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)} \right\rVert_\infty$
0	0.10000000	0.10000000	−0.10000000	
1	0.49998333	0.00944115	−0.52310127	0.423
2	0.49999593	0.00002557	−0.52336331	9.4×10^{-3}
3	0.50000000	0.00001234	−0.52359814	2.3×10^{-4}
4	0.50000000	0.00000003	−0.52359847	1.2×10^{-5}
5	0.50000000	0.00000002	−0.52359877	3.1×10^{-7}

We could use the error bound (10.3) with $K = 0.843$ in the previous example. This gives

$$\|\mathbf{x}^{(5)} - \mathbf{p}\|_\infty \le \frac{(0.843)^5}{1 - 0.843}(0.423) < 1.15,$$

which does not indicate the true accuracy of $\mathbf{x}^{(5)}$. The actual solution is

$$\mathbf{p} = \left(0.5, 0, -\frac{\pi}{6}\right)^t \approx (0.5, 0, -0.5235987757)^t, \quad \text{so} \quad \|\mathbf{x}^{(5)} - \mathbf{p}\|_\infty \le 2 \times 10^{-8}.$$

Accelerating Convergence

One way to accelerate convergence of the fixed-point iteration is to use the latest estimates $x_1^{(k)}, \ldots, x_{i-1}^{(k)}$ instead of $x_1^{(k-1)}, \ldots, x_{i-1}^{(k-1)}$ to compute $x_i^{(k)}$, as in the Gauss-Seidel method for linear systems. The component equations for the problem in the example then become

$$x_1^{(k)} = \frac{1}{3}\cos\left(x_2^{(k-1)}x_3^{(k-1)}\right) + \frac{1}{6},$$

$$x_2^{(k)} = \frac{1}{9}\sqrt{\left(x_1^{(k)}\right)^2 + \sin x_3^{(k-1)} + 1.06} - 0.1,$$

$$x_3^{(k)} = -\frac{1}{20}e^{-x_1^{(k)}x_2^{(k)}} - \frac{10\pi - 3}{60}.$$

With $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$, the results of these calculations are listed in Table 10.2.

Table 10.2

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_\infty$
0	0.10000000	0.10000000	−0.10000000	
1	0.49998333	0.02222979	−0.52304613	0.423
2	0.49997747	0.00002815	−0.52359807	2.2×10^{-2}
3	0.50000000	0.00000004	−0.52359877	2.8×10^{-5}
4	0.50000000	0.00000000	−0.52359877	3.8×10^{-8}

The iterate $\mathbf{x}^{(4)}$ is accurate to within 10^{-7} in the l_∞ norm; so the convergence was indeed accelerated for this problem by using the Gauss-Seidel method. However, this method does not *always* accelerate the convergence.

EXERCISE SET 10.1

8. Use functional iteration to find solutions to the following nonlinear systems, accurate to within 10^{-5}, using the l_∞ norm.

b.
$$3x_1^2 - x_2^2 = 0,$$
$$3x_1x_2^2 - x_1^3 - 1 = 0.$$

MAP2220

$$x_1 = \frac{x_2}{\sqrt{3}}$$

$$x_2 = \sqrt{\frac{x_1^3 + 1}{3x_1}}$$

iteração	x1	x2	norm_2	f1	f2
0	1	2		-1	10
1	1,154701	0,816497	1,424613	3,333333	-0,2302
2	0,471405	0,856224	0,468472	-0,06645	-0,06797
3	0,494341	0,883844	0,001289	-0,04806	0,037706
4	0,510288	0,869342	0,000465	0,025425	0,024083
5	0,501915	0,860247	0,000153	0,015732	-0,01215
6	0,496664	0,864925	4,95E-05	-0,00807	-0,00786
7	0,499365	0,867969	1,66E-05	-0,00527	0,004095
8	0,501122	0,866393	5,57E-06	0,002734	0,002638
9	0,500212	0,86538	1,85E-06	0,001755	-0,00136
10	0,499627	0,865903	6,16E-07	-0,00091	-0,00088

10.2 Newton's Method

The problem in Example 2 of Section 10.1 is transformed into a convergent fixed-point problem by algebraically solving the three equations for the three variables x_1, x_2, and x_3. It is, however, unusual to be able to find an explicit representation for all the variables. In this section, we consider an algorithmic procedure to perform the transformation in a more general situation.

To construct the algorithm that led to an appropriate fixed-point method in the one-dimensional case, we found a function ϕ with the property that

$$g(x) = x - \phi(x) f(x)$$

gives quadratic convergence to the fixed point p of the function g (see Section 2.4). From this condition Newton's method evolved by choosing $\phi(x) = 1/f'(x)$, assuming that $f'(x) \neq 0$.

A similar approach in the n-dimensional case involves a matrix

$$A(\mathbf{x}) = \begin{bmatrix} a_{11}(\mathbf{x}) & a_{12}(\mathbf{x}) & \cdots & a_{1n}(\mathbf{x}) \\ a_{21}(\mathbf{x}) & a_{22}(\mathbf{x}) & \cdots & a_{2n}(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ a_{n1}(\mathbf{x}) & a_{n2}(\mathbf{x}) & \cdots & a_{nn}(\mathbf{x}) \end{bmatrix}, \tag{10.5}$$

where each of the entries $a_{ij}(\mathbf{x})$ is a function from $\mathbb{R}^n$ into $\mathbb{R}$. This requires that $A(\mathbf{x})$ be found so that

$$\mathbf{G}(\mathbf{x}) = \mathbf{x} - A(\mathbf{x})^{-1}\mathbf{F}(\mathbf{x})$$

gives quadratic convergence to the solution of $\mathbf{F}(\mathbf{x}) = \mathbf{0}$, assuming that $A(\mathbf{x})$ is nonsingular at the fixed point $\mathbf{p}$ of $\mathbf{G}$.

The following theorem parallels Theorem 2.8 on page 80. Its proof requires being able to express $\mathbf{G}$ in terms of its Taylor series in n variables about $\mathbf{p}$.

Theorem 10.7 Let $\mathbf{p}$ be a solution of $\mathbf{G}(\mathbf{x}) = \mathbf{x}$. Suppose a number $\delta > 0$ exists with

(i) $\partial g_i/\partial x_j$ is continuous on $N_\delta = \{\, \mathbf{x} \mid \|\mathbf{x} - \mathbf{p}\| < \delta \,\}$, for each $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$;

(ii) $\partial^2 g_i(\mathbf{x})/(\partial x_j \partial x_k)$ is continuous, and $|\partial^2 g_i(\mathbf{x})/(\partial x_j \partial x_k)| \leq M$ for some constant M, whenever $\mathbf{x} \in N_\delta$, for each $i = 1, 2, ..., n$, $j = 1, 2, \ldots, n$, and $k = 1, 2, \ldots, n$;

(iii) $\partial g_i(\mathbf{p})/\partial x_k = 0$, for each $i = 1, 2, \ldots, n$ and $k = 1, 2, \ldots, n$.

Then a number $\hat{\delta} \leq \delta$ exists such that the sequence generated by $\mathbf{x}^{(k)} = \mathbf{G}(\mathbf{x}^{(k-1)})$ converges quadratically to $\mathbf{p}$ for any choice of $\mathbf{x}^{(0)}$, provided that $\|\mathbf{x}^{(0)} - \mathbf{p}\| < \hat{\delta}$. Moreover,

$$\|\mathbf{x}^{(k)} - \mathbf{p}\|_\infty \leq \frac{n^2 M}{2} \|\mathbf{x}^{(k-1)} - \mathbf{p}\|_\infty^2, \quad \text{for each } k \geq 1. \qquad \blacksquare$$

To apply Theorem 10.7, suppose that $A(\mathbf{x})$ is an $n \times n$ matrix of functions from $\mathbb{R}^n$ into $\mathbb{R}$ in the form of Eq. (10.5), where the specific entries will be chosen later. Assume, moreover, that $A(\mathbf{x})$ is nonsingular near a solution $\mathbf{p}$ of $\mathbf{F}(\mathbf{x}) = \mathbf{0}$, and let $b_{ij}(\mathbf{x})$ denote the entry of $A(\mathbf{x})^{-1}$ in the ith row and jth column.

For $\mathbf{G}(\mathbf{x}) = \mathbf{x} - A(\mathbf{x})^{-1}\mathbf{F}(\mathbf{x})$, we have $g_i(\mathbf{x}) = x_i - \sum_{j=1}^{n} b_{ij}(\mathbf{x}) f_j(\mathbf{x})$. So

$$\frac{\partial g_i}{\partial x_k}(\mathbf{x}) = \begin{cases} 1 - \displaystyle\sum_{j=1}^{n}\left(b_{ij}(\mathbf{x})\frac{\partial f_j}{\partial x_k}(\mathbf{x}) + \frac{\partial b_{ij}}{\partial x_k}(\mathbf{x}) f_j(\mathbf{x})\right), & \text{if } i = k, \\ -\displaystyle\sum_{j=1}^{n}\left(b_{ij}(\mathbf{x})\frac{\partial f_j}{\partial x_k}(\mathbf{x}) + \frac{\partial b_{ij}}{\partial x_k}(\mathbf{x}) f_j(\mathbf{x})\right), & \text{if } i \neq k. \end{cases}$$

Theorem 10.7 implies that we need $\partial g_i(\mathbf{p})/\partial x_k = 0$, for each $i = 1, 2, \ldots, n$ and $k = 1, 2, \ldots, n$. This means that for $i = k$,

$$0 = 1 - \sum_{j=1}^{n} b_{ij}(\mathbf{p}) \frac{\partial f_j}{\partial x_i}(\mathbf{p}),$$

that is,

$$\sum_{j=1}^{n} b_{ij}(\mathbf{p}) \frac{\partial f_j}{\partial x_i}(\mathbf{p}) = 1. \tag{10.6}$$

When $k \neq i$,

$$0 = -\sum_{j=1}^{n} b_{ij}(\mathbf{p}) \frac{\partial f_j}{\partial x_k}(\mathbf{p}),$$

so

$$\sum_{j=1}^{n} b_{ij}(\mathbf{p}) \frac{\partial f_j}{\partial x_k}(\mathbf{p}) = 0. \tag{10.7}$$

The Jacobian Matrix

Define the matrix $J(\mathbf{x})$ by

$$J(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}) & \frac{\partial f_n}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}) \end{bmatrix}. \tag{10.8}$$

Then conditions (10.6) and (10.7) require that

$$A(\mathbf{p})^{-1}J(\mathbf{p}) = I, \text{ the identity matrix,} \quad \text{so } A(\mathbf{p}) = J(\mathbf{p}).$$

An appropriate choice for $A(\mathbf{x})$ is, consequently, $A(\mathbf{x}) = J(\mathbf{x})$ since this satisfies condition (iii) in Theorem 10.7. The function $\mathbf{G}$ is defined by

$$\mathbf{G}(\mathbf{x}) = \mathbf{x} - J(\mathbf{x})^{-1}\mathbf{F}(\mathbf{x}),$$

and the functional iteration procedure evolves from selecting $\mathbf{x}^{(0)}$ and generating, for $k \geq 1$,

$$\mathbf{x}^{(k)} = \mathbf{G}\left(\mathbf{x}^{(k-1)}\right) = \mathbf{x}^{(k-1)} - J\left(\mathbf{x}^{(k-1)}\right)^{-1}\mathbf{F}\left(\mathbf{x}^{(k-1)}\right). \tag{10.9}$$

This is called **Newton's method for nonlinear systems**, and it is generally expected to give quadratic convergence, provided that a sufficiently accurate starting value is known and that $J(\mathbf{p})^{-1}$ exists. The matrix $J(\mathbf{x})$ is called the **Jacobian** matrix and has a number of applications in analysis. It might, in particular, be familiar to the reader due to its application in the multiple integration of a function of several variables over a region that requires a change of variables to be performed.

A weakness in Newton's method arises from the need to compute and invert the matrix $J(\mathbf{x})$ at each step. In practice, explicit computation of $J(\mathbf{x})^{-1}$ is avoided by performing the operation in a two-step manner. First, a vector $\mathbf{y}$ is found that satisfies $J(\mathbf{x}^{(k-1)})\mathbf{y} = -\mathbf{F}(\mathbf{x}^{(k-1)})$. Then the new approximation, $\mathbf{x}^{(k)}$, is obtained by adding $\mathbf{y}$ to $\mathbf{x}^{(k-1)}$. Algorithm 10.1 uses this two-step procedure.

ALGORITHM 10.1

Newton's Method for Systems

To approximate the solution of the nonlinear system $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ given an initial approximation $\mathbf{x}$:

INPUT number n of equations and unknowns; initial approximation $\mathbf{x} = (x_1, \ldots, x_n)^t$, tolerance TOL; maximum number of iterations N.

OUTPUT approximate solution $\mathbf{x} = (x_1, \ldots, x_n)^t$ or a message that the number of iterations was exceeded.

Step 1 Set $k = 1$.

Step 2 While ($k \leq N$) do Steps 3–7.

- *Step 3* Calculate $\mathbf{F}(\mathbf{x})$ and $J(\mathbf{x})$, where $J(\mathbf{x})_{i,j} = (\partial f_i(\mathbf{x})/\partial x_j)$ for $1 \leq i, j \leq n$.
- *Step 4* Solve the $n \times n$ linear system $J(\mathbf{x})\mathbf{y} = -\mathbf{F}(\mathbf{x})$.
- *Step 5* Set $\mathbf{x} = \mathbf{x} + \mathbf{y}$.
- *Step 6* If $||\mathbf{y}|| < TOL$ then OUTPUT ($\mathbf{x}$);
  (*The procedure was successful.*)
  STOP.
- *Step 7* Set $k = k + 1$.

Step 8 OUTPUT ('Maximum number of iterations exceeded');
(*The procedure was unsuccessful.*)
STOP.

Example 1 The nonlinear system

$$3x_1 - \cos(x_2x_3) - \frac{1}{2} = 0,$$

$$x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06 = 0,$$

$$e^{-x_1x_2} + 20x_3 + \frac{10\pi - 3}{3} = 0$$

was shown in Example 2 of Section 10.1 to have the approximate solution $(0.5, 0, -0.52359877)^t$. Apply Newton's method to this problem with $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$.

Solution Define

$$\mathbf{F}(x_1,x_2,x_3) = (f_1(x_1,x_2,x_3),\ f_2(x_1,x_2,x_3),\ f_3(x_1,x_2,x_3))^t,$$

where

$$f_1(x_1,x_2,x_3) = 3x_1 - \cos(x_2x_3) - \frac{1}{2},$$
$$f_2(x_1,x_2,x_3) = x_1^2 - 81(x_2+0.1)^2 + \sin x_3 + 1.06,$$

and

$$f_3(x_1,x_2,x_3) = e^{-x_1x_2} + 20x_3 + \frac{10\pi - 3}{3}.$$

The Jacobian matrix $J(\mathbf{x})$ for this system is

$$J(x_1,x_2,x_3) = \begin{bmatrix} 3 & x_3 \sin x_2x_3 & x_2 \sin x_2x_3 \\ 2x_1 & -162(x_2+0.1) & \cos x_3 \\ -x_2e^{-x_1x_2} & -x_1e^{-x_1x_2} & 20 \end{bmatrix}.$$

Let $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$. Then $\mathbf{F}(\mathbf{x}^{(0)}) = (-0.199995, -2.269833417, 8.462025346)^t$ and

$$J(\mathbf{x}^{(0)}) = \begin{bmatrix} 3 & 9.999833334 \times 10^{-4} & 9.999833334 \times 10^{-4} \\ 0.2 & -32.4 & 0.9950041653 \\ -0.09900498337 & -0.09900498337 & 20 \end{bmatrix}.$$

Solving the linear system, $J(\mathbf{x}^{(0)})\mathbf{y}^{(0)} = -\mathbf{F}(\mathbf{x}^{(0)})$ gives

$$\mathbf{y}^{(0)} = \begin{bmatrix} 0.3998696728 \\ -0.08053315147 \\ -0.4215204718 \end{bmatrix} \quad \text{and} \quad \mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{y}^{(0)} = \begin{bmatrix} 0.4998696782 \\ 0.01946684853 \\ -0.5215204718 \end{bmatrix}.$$

Continuing for $k = 2, 3, \ldots,$ we have

$$\begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{bmatrix} = \begin{bmatrix} x_1^{(k-1)} \\ x_2^{(k-1)} \\ x_3^{(k-1)} \end{bmatrix} + \begin{bmatrix} y_1^{(k-1)} \\ y_2^{(k-1)} \\ y_3^{(k-1)} \end{bmatrix},$$

where

$$\begin{bmatrix} y_1^{(k-1)} \\ y_2^{(k-1)} \\ y_3^{(k-1)} \end{bmatrix} = -\left(J\left(x_1^{(k-1)}, x_2^{(k-1)}, x_3^{(k-1)}\right)\right)^{-1} \mathbf{F}\left(x_1^{(k-1)}, x_2^{(k-1)}, x_3^{(k-1)}\right).$$

Thus, at the kth step, the linear system $J\left(\mathbf{x}^{(k-1)}\right)\mathbf{y}^{(k-1)} = -\mathbf{F}\left(\mathbf{x}^{(k-1)}\right)$ must be solved, where

$$J\left(\mathbf{x}^{(k-1)}\right) = \begin{bmatrix} 3 & x_3^{(k-1)} \sin x_2^{(k-1)} x_3^{(k-1)} & x_2^{(k-1)} \sin x_2^{(k-1)} x_3^{(k-1)} \\ 2x_1^{(k-1)} & -162\left(x_2^{(k-1)} + 0.1\right) & \cos x_3^{(k-1)} \\ -x_2^{(k-1)} e^{-x_1^{(k-1)} x_2^{(k-1)}} & -x_1^{(k-1)} e^{-x_1^{(k-1)} x_2^{(k-1)}} & 20 \end{bmatrix},$$

$$\mathbf{y}^{(k-1)} = \begin{bmatrix} y_1^{(k-1)} \\ y_2^{(k-1)} \\ y_3^{(k-1)} \end{bmatrix},$$

and

$$\mathbf{F}\left(\mathbf{x}^{(k-1)}\right) = \begin{bmatrix} 3x_1^{(k-1)} - \cos x_2^{(k-1)} x_3^{(k-1)} - \frac{1}{2} \\ \left(x_1^{(k-1)}\right)^2 - 81\left(x_2^{(k-1)} + 0.1\right)^2 + \sin x_3^{(k-1)} + 1.06 \\ e^{-x_1^{(k-1)} x_2^{(k-1)}} + 20x_3^{(k-1)} + \frac{10\pi - 3}{3} \end{bmatrix}.$$

The results using this iterative procedure are shown in Table 10.3. ■

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_\infty$
0	0.1000000000	0.1000000000	−0.1000000000	
1	0.4998696728	0.0194668485	−0.5215204718	0.4215204718
2	0.5000142403	0.0015885914	−0.5235569638	1.788×10^{-2}
3	0.5000000113	0.0000124448	−0.5235984500	1.576×10^{-3}
4	0.5000000000	8.516×10^{-10}	−0.5235987755	1.244×10^{-5}
5	0.5000000000	-1.375×10^{-11}	−0.5235987756	8.654×10^{-10}

The previous example illustrates that Newton's method can converge very rapidly once a good approximation is obtained that is near the true solution. However, it is not always easy to determine good starting values, and the method is comparatively expensive to employ. In the next section, we consider a method for overcoming the latter weakness. Good starting values can usually be found using the Steepest Descent method, which will be discussed in Section 10.4.

EXERCISE SET 10.2

1. Use Newton's method with $\mathbf{x}^{(0)} = \mathbf{0}$ to compute $\mathbf{x}^{(2)}$ for each of the following nonlinear systems.

a.

$$4x_1^2 - 20x_1 + \frac{1}{4}x_2^2 + 8 = 0,$$

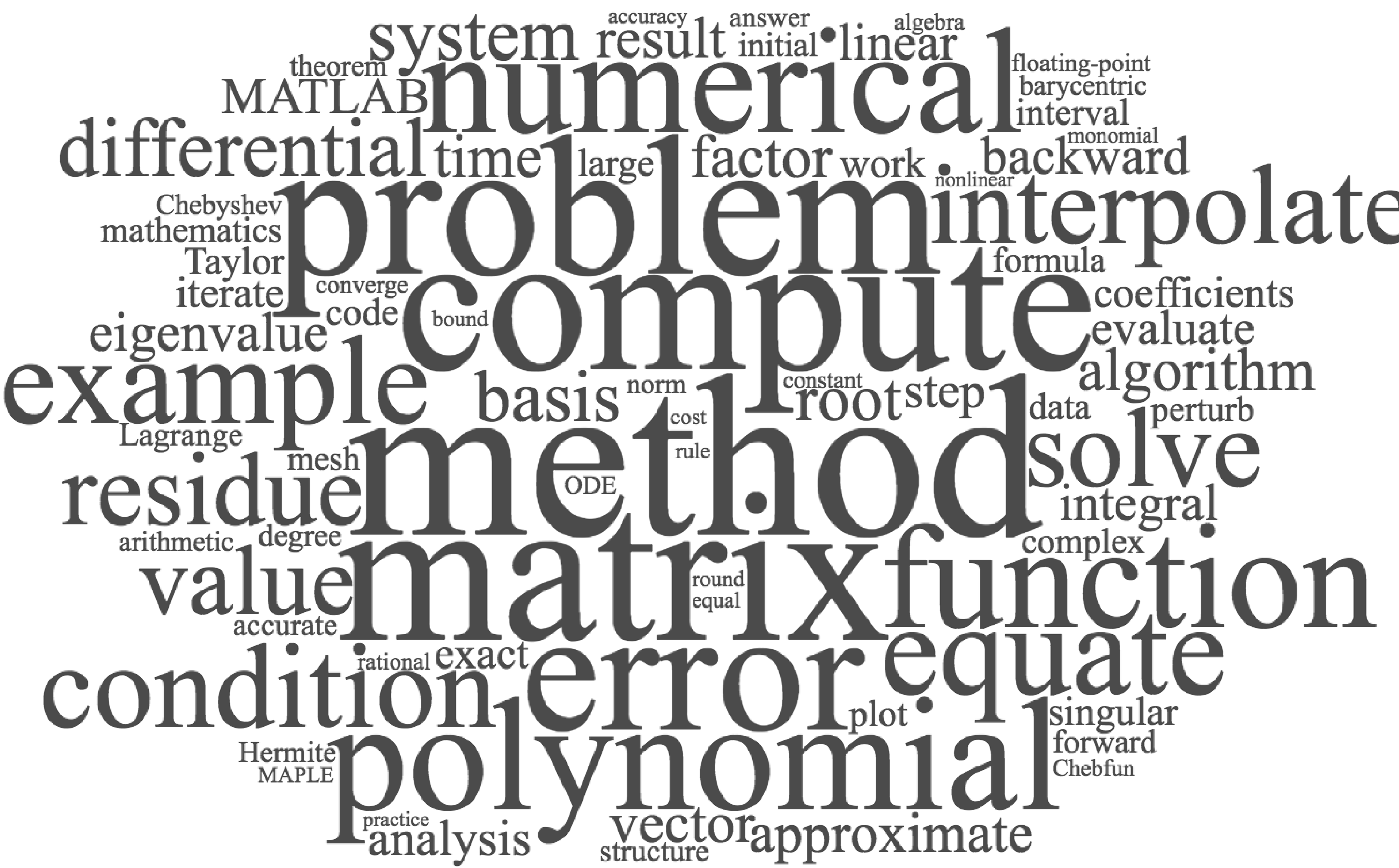
$$\frac{1}{2}x_1x_2^2 + 2x_1 - 5x_2 + 8 = 0.$$

MAP2220

iteração	x1	x2	f1	f2	df1dx1	df1dx2	df2dx1	df2dx2	detJ	y1	y2	norm_2
0	0	0	8	8	-20	0	2	-5	100	0,4	1,76	3,2576
1	0,4	1,76	1,4144	0,61952	-16,8	0,88	3,5488	-4,296	69,04986	0,095894	0,223423	0,059114
2	0,495894	1,983423	0,04926185	0,050085	-16,0329	0,991712	3,966984	-4,01643	60,46077	0,004094	0,016514	0,000289
3	0,499988	1,999937	0,000135218	0,000202	-16,0001	0,999969	3,999874	-4,00006	60,00155	1,24E-05	6,3E-05	4,12E-09
4	0,5	2	1,60427E-09	2,55E-09	-16	1	4	-4	60	1,49E-10	7,87E-10	6,42E-19
5	0,5	2	0	0	-16	1	4	-4	60	0	0	0

10 Numerical Solutions of Nonlinear Systems of Equations 629

MAP2220


accuracy
answer
algebra
system
result
initial
linear
theorem
floating-point
MATLAB
numerical
barycentric
interval
monomial
differential
time
large
factor
work
backward
nonlinear
Chebyshev
problem
interpolate
mathematics
Taylor
formula
iterate
converge
compute
coefficients
code
bound
evaluate
eigenvalue
algorithm
example
basis
norm
constant
root
step
data
perturb
Lagrange
cost
method
solve
mesh
rule
ODE
residue
integral
arithmetic
degree
complex
matrix
function
value
round
equal
accurate
rational
exact
error
equate
condition
plot
singular
Hermite
forward
MAPLE
polynomial
Chebfun
practice
analysis
vector
approximate
structure