(i) The sequence

$$\tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}, \tfrac{1}{32}, \ldots$$

with general term $x_k = 2^{-k}$, for $k = 1, 2, \ldots$.

(ii) The sequence

$$1.05, 1.0005, 1.000005, \ldots$$

with general term $x_k = 1 + 5 \times 10^{-2k}$, for $k = 1, 2, \ldots$.

(iii) The sequence with general term $x_k = 2^{-2^k}$.

(iv) The sequence with general term $x_k = 3^{-k^2}$.

(v) The sequence with general term $x_k = 1 - 2^{-2^k}$ for k odd, and $x_k = 1 + 2^{-k}$ for k even.

5.2. Consider the sequence defined by $x_0 = a > 0$ and

$$x_{k+1} = \tfrac{1}{2}\left(x_k + \frac{a}{x_k}\right).$$

Prove that this sequence converges to $x_* = \sqrt{a}$ and that the convergence rate is quadratic, and determine the rate constant.

5.3. Consider a convergent sequence $\{ x_k \}$ and define a second sequence $\{ y_k \}$ with $y_k = c x_k$ where c is some nonzero constant. What is the relationship between the convergence rates and rate constants of the two sequences?

5.4. Let $\{ x_k \}$ and $\{ c_k \}$ be convergent sequences, and assume that

$$\lim_{k\to\infty} c_k = c \neq 0.$$

Consider the sequence $\{ y_k \}$ with $y_k = c_k x_k$. Is this sequence guaranteed to converge? If so, can its convergence rate and rate constant be determined from the rates and rate constants for the sequences $\{ x_k \}$ and $\{ c_k \}$?

2.6 Taylor Series

The Taylor series is a tool for approximating a function f near a specified point x_0. The approximation obtained is a polynomial, i.e., a function that is easy to manipulate. The Taylor series is a general tool—it can be applied whenever the function has derivatives—and it has many uses:

- It allows you to estimate the value of the function near the given point (when the function is difficult to evaluate directly).
- The derivatives and integral of the approximation can be used to estimate the derivatives and integral of the original function.
- It is used to derive many algorithms for finding zeroes of functions (see below), for minimizing functions, etc.

Since many problems are difficult to solve exactly, and an approximate solution is often adequate (the data for the problem may be inaccurate), the Taylor series is widely used, both theoretically and practically. Even if the data are exact, an approximate solution may be adequate, and in any case it is all we can hope for under most circumstances.

How does it work? We first consider the case of a one-dimensional function f with n continuous derivatives. Let x_0 be a specified point (say $x_0 = 17.5$ or $x_0 = 0$). Then the nth order Taylor series approximation is

$$f(x_0 + p) \approx f(x_0) + pf'(x_0) + \frac{1}{2}p^2 f''(x_0) + \cdots + \frac{p^n}{n!}f^{(n)}(x_0).$$

Here $f^{(n)}(x_0)$ is the nth derivative of f at the point x_0, and $n! = n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$. Notice that $\frac{1}{2}p^2 f''(x_0) = (p^2/2!)f^{(2)}(x_0)$. In this formula, p is a variable; we will decide later what values p will take. The approximation will normally only be accurate for small values of p.

Example 2.3 (Taylor Series)**.** Let $f(x) = \sqrt{x}$ and let $x_0 = 1$. Then

$$\begin{aligned}
f(x_0) &= \sqrt{x_0} = \sqrt{1} = 1 \\
f'(x_0) &= \tfrac{1}{2}x_0^{-\frac{1}{2}} = \tfrac{1}{2}1^{-\frac{1}{2}} = \tfrac{1}{2} \\
f''(x_0) &= -\tfrac{1}{4}x_0^{-\frac{3}{2}} = -\tfrac{1}{4}1^{-\frac{3}{2}} = -\tfrac{1}{4} \\
f'''(x_0) &= \tfrac{3}{8}x_0^{-\frac{5}{2}} = \tfrac{3}{8}1^{-\frac{5}{2}} = \tfrac{3}{8} \\
&\ \vdots
\end{aligned}$$

Hence, substituting into the formula for the Taylor series,

$$\begin{aligned}
\sqrt{1+p} &= f(x_0 + p) \\
&\approx f(x_0) + pf'(x_0) + \tfrac{1}{2}p^2 f''(x_0) + \tfrac{1}{6}p^3 f'''(x_0) \\
&= 1 + p(\tfrac{1}{2}) + \tfrac{1}{2}p^2(-\tfrac{1}{4}) + \tfrac{1}{6}p^3(\tfrac{3}{8}) \\
&= 1 + \tfrac{1}{2}p - \tfrac{1}{8}p^2 + \tfrac{1}{16}p^3.
\end{aligned}$$

How do we use this? Suppose we want to approximate $f(1.6)$. Then $x_0 + p = 1 + p = 1.6$, and so $p = 0.6$:

$$\begin{aligned}
\sqrt{1.6} &= \sqrt{1 + 0.6} \\
&\approx 1 + \tfrac{1}{2}(0.6) - \tfrac{1}{8}(0.6)^2 + \tfrac{1}{16}(0.6)^3 \approx 1.2685.
\end{aligned}$$

The true value is $1.264911\ldots$; the approximation is accurate to three digits. ∎

The first two terms of the Taylor series give us the formula for the tangent line for the function f at the point x_0. We commonly define the tangent line in terms of a general point x, and not in terms of p. Since $x_0 + p = x$, we can rearrange to get $p = x - x_0$. Substitute this into the first two terms of the series to get the tangent line:

$$y = f(x_0) + (x - x_0)f'(x_0).$$

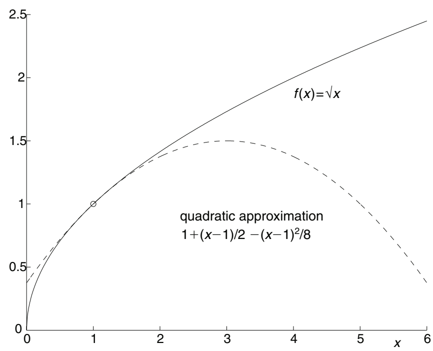

Figure 2.9. *Taylor series approximation.*

For the example above we get

$$y = 1 + (x - 1)\tfrac{1}{2} \quad \text{or} \quad y = \tfrac{1}{2}(x + 1).$$

The first three terms of the Taylor series give a quadratic approximation to the function f at the point x_0. This is illustrated in Figure 2.9.

So far we have only considered a Taylor series for a function of one variable. The Taylor series can also be derived for real-valued functions of many variables. If we use matrix and vector notation, then there is an obvious analogy between the two cases:

$$\begin{aligned} \text{1-variable:} \quad & f(x_0 + p) = f(x_0) + pf'(x_0) + \tfrac{1}{2}p^2 f''(x_0) + \cdots \\ n\text{-variables:} \quad & f(x_0 + p) = f(x_0) + p^T\nabla f(x_0) + \tfrac{1}{2}p^T\nabla^2 f(x_0)p + \cdots. \end{aligned}$$

In the second line above x_0 and p are both vectors. The notation $\nabla f(x_0)$ refers to the gradient of the function f at the point $x = x_0$. The notation $\nabla^2 f(x_0)$ represents the Hessian of f at the point $x = x_0$. (See Appendix B.4.) The higher-order terms of the Taylor series can also be written down, but the notation is more complex and they will not be required in this book.

Example 2.4 (Multidimensional Taylor Series)**.** Consider the function

$$f(x_1, x_2) = x_1^3 + 5x_1^2x_2 + 7x_1x_2^2 + 2x_2^3$$

at the point

$$x_0 = (-2, 3)^T.$$

The gradient of this function is

$$\nabla f(x) = \begin{pmatrix} 3x_1^2 + 10x_1x_2 + 7x_2^2 \\ 5x_1^2 + 14x_1x_2 + 6x_2^2 \end{pmatrix}$$

and the Hessian matrix is

$$\nabla^2 f(x) = \begin{pmatrix} 6x_1 + 10x_2 & 10x_1 + 14x_2 \\ 10x_1 + 14x_2 & 14x_1 + 12x_2 \end{pmatrix}.$$

At the point $x_0 = (-2, 3)^T$ these become

$$\nabla f(x_0) = \begin{pmatrix} 15 \\ -10 \end{pmatrix} \quad \text{and} \quad \nabla^2 f(x_0) = \begin{pmatrix} 18 & 22 \\ 22 & 8 \end{pmatrix}.$$

If $p = (p_1, p_2)^T = (0.1, 0.2)^T$, then

$$\begin{aligned} f(-1.9, 3.2) &= f(-2 + 0.1, 3 + 0.2) \\ &= f(x_0 + p) \\ &\approx f(x_0) + p^T \nabla f(x_0) + \frac{1}{2} p^T \nabla^2 f(x_0) p \\ &= -20 + (\, 0.1 \quad 0.2\,) \begin{pmatrix} 15 \\ -10 \end{pmatrix} + \frac{1}{2} (\, 0.1 \quad 0.2\,) \begin{pmatrix} 18 & 22 \\ 22 & 8 \end{pmatrix} \begin{pmatrix} 0.1 \\ 0.2 \end{pmatrix} \\ &= -20 - 0.5 + 0.69 = -19.81. \end{aligned}$$

The true value is $f(-1.9, 3.2) = -19.755$, so the approximation is accurate to three digits. ∎

The Taylor series for multidimensional problems can also be derived using summations rather than matrix-vector notation:

$$f(x_0 + p) = f(x_0) + \sum_{i=1}^{n} p_i \frac{\partial f(x)}{\partial x_i}\Big|_{x=x_0} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_j \frac{\partial^2 f(x)}{\partial x_i \partial x_j}\Big|_{x=x_0} + \cdots.$$

The formula is the same as before; only the notation has changed.

There is an alternate form of the Taylor series that is often used, called the *remainder* form. If three terms are used it looks like

$$\begin{aligned} \text{1-variable:} \quad & f(x_0 + p) = f(x_0) + p f'(x_0) + \tfrac{1}{2} p^2 f''(\xi) \\ n\text{-variables:} \quad & f(x_0 + p) = f(x_0) + p^T \nabla f(x_0) + \tfrac{1}{2} p^T \nabla^2 f(\xi) p. \end{aligned}$$

The point ξ is an *unknown* point lying between x_0 and $x_0 + p$. In this form the series is exact, but it involves an unknown point, so it cannot be evaluated. This form of the series is often used for theoretical purposes, or to derive bounds on the accuracy of the series. The accuracy of the series can be analyzed by establishing bounds on the final "remainder" term.

If the remainder form of the series is used, but with only two terms, then we obtain

$$\begin{aligned} \text{1-variable:} \quad & f(x_0 + p) = f(x_0) + p f'(\xi) \\ n\text{-variables:} \quad & f(x_0 + p) = f(x_0) + p^T \nabla f(\xi). \end{aligned}$$

This result is known as the *mean-value theorem.*

Exercises

6.1. Find the first four terms of the Taylor series for

$$f(x) = \log(1 + x)$$