1

The approximation problem and existence of best approximations

1.1 Examples of approximation problems

A simple example of an approximation problem is to draw a straight line that fits the curve shown in Figure 1.1. Alternatively we may require a straight line fit to the data shown in Figure 1.2. Three possible fits to the discrete data are shown in Figure 1.3, and it seems that lines B and C are better than line A. Whether B or C is preferable depends on our confidence in the highest data point, and to choose between the two straight lines we require a measure of the quality of the trial approximations. These examples show the three main ingredients of an approximation calculation, which are as follows: (1) A function, or some data, or

Figure 1.1. A function to be approximated.

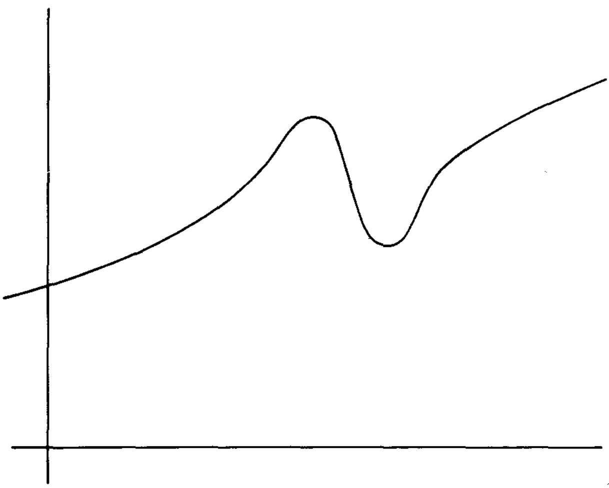


Figure 1.2. Some data to be approximated.

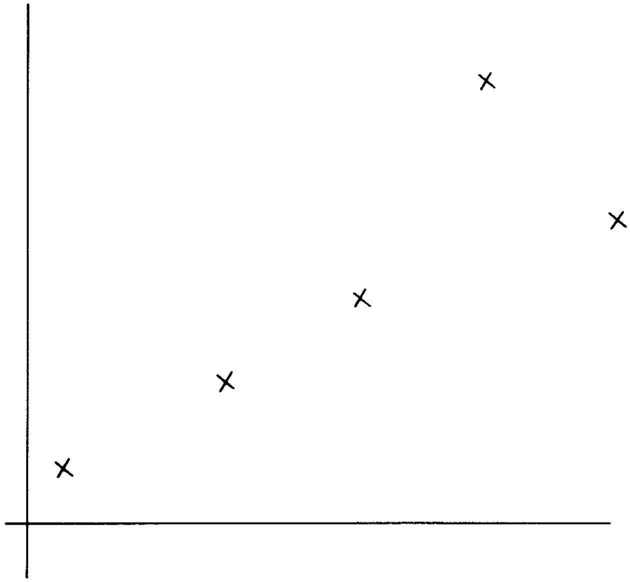
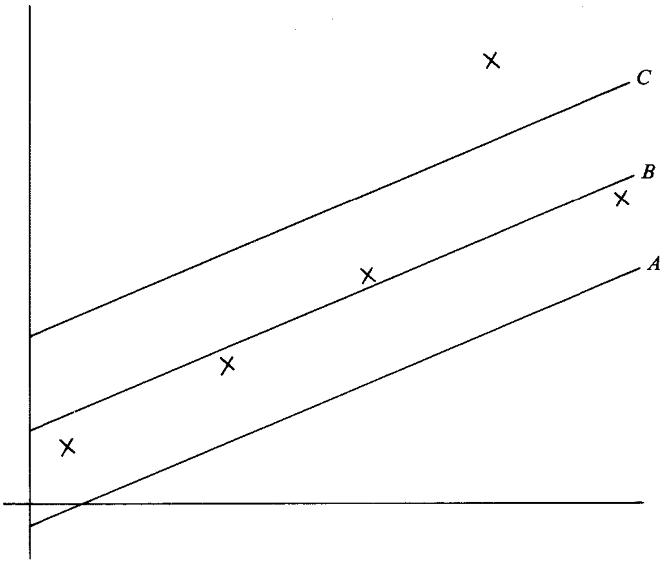


Figure 1.3. Three straight-line fits to the data of Figure 1.2.



more generally a member of a set, that is to be approximated. We call it f. (2) A set, $\mathscr{A}$ say, of approximations, which in the case of the given examples is the set of all straight lines. (3) A means of selecting an approximation from $\mathscr{A}$.

Approximation problems of this type arise frequently. For instance we may estimate the solution of a differential equation by a function of a certain simple form that depends on adjustable parameters, where the measure of goodness of the approximation is a scalar quantity that is derived from the residual that occurs when the approximating function is substituted into the differential equation. Another example comes from the choice of components in electrical circuits. The function f may be the required response from the circuit, and the range of available components gives a set $\mathscr{A}$ of attainable responses. We have to approximate f by a member of $\mathscr{A}$, and we require a criterion that selects suitable components. Moreover, in computer calculations of mathematical functions, the mathematical function is usually approximated by one that is easy to compute.

Many closely related questions are of interest also. Given f and $\mathscr{A}$, we may wish to know whether any member of $\mathscr{A}$ satisfies a fixed tolerance condition, and, if suitable approximations exist, we may be willing to accept any one. It is often useful to develop methods for selecting a member of $\mathscr{A}$ such that the error of the chosen approximation is always within a certain factor of the least error that can be achieved. It may be possible to increase the size of $\mathscr{A}$ if necessary, for example $\mathscr{A}$ may be a linear space of polynomials of any fixed degree, and we may wish to predict the improvement in the best approximation that comes from enlarging $\mathscr{A}$ by increasing the degree. At the planning stage of a numerical method we may know only that f will be a member of a set $\mathscr{B}$, in which case it is relevant to discover how well any member of $\mathscr{B}$ can be approximated from $\mathscr{A}$. Further, given $\mathscr{B}$, it may be valuable to compare the suitability of two different sets of approximating functions, $\mathscr{A}_0$ and $\mathscr{A}_1$. Numerical methods for the calculation of approximating functions are required. This book presents much of the basic theory and algorithms that are relevant to these questions, and the material is selected and described in a way that is intended to help the reader to develop suitable techniques for himself.

1.2 Approximation in a metric space

The framework of metric spaces provides a general way of measuring the goodness of an approximation, because one of the basic properties of a metric space is that it has a distance function. Specifically, the distance function $d(x, y)$ of a metric space $\mathscr{B}$ is a real-valued function, that is defined for all pairs of points (x, y) in $\mathscr{B}$, and that has the following properties. If $x \neq y$, then $d(x, y)$ is positive and is equal to $d(y, x)$. If $x = y$, then the value of $d(x, y)$ is zero. The triangle inequality

$$d(x, y) \leqslant d(x, z) + d(z, y) \tag{1.1}$$

must hold, where x, y and z are any three points in $\mathscr{B}$.

In most approximation problems there exists a suitable metric space that contains both f and the set of approximations $\mathscr{A}$. Then it is natural to decide that $a_0 \in \mathscr{A}$ is a better approximation than $a_1 \in \mathscr{A}$ if the inequality

$$d(a_0, f) < d(a_1, f) \tag{1.2}$$

is satisfied. We define $a^* \in \mathscr{A}$ to be a best approximation if the condition

$$d(a^*, f) \leqslant d(a, f) \tag{1.3}$$

holds for all $a \in \mathscr{A}$.

The metric space should be chosen so that it provides a measure of the error of each trial approximation. For example, in the problem of fitting the data of Figure 1.2 by a straight line, we approximate a set of points $\{(x_i, y_i); i = 1, 2, 3, 4, 5\}$ by a function of the form

$$p(x) = c_0 + c_1 x, \tag{1.4}$$

where c_0 and c_1 are scalar coefficients. Because we are interested in only five values of x, the most convenient space is $\mathscr{R}^5$. The fact that $p(x)$ depends on two parameters is not relevant to the choice of metric space. We measure the goodness of the approximation (1.4) as the distance, according to the metric we have chosen, from the vector of function values $\{p(x_i); i = 1, 2, 3, 4, 5\}$ to the data values $\{y_i; i = 1, 2, 3, 4, 5\}$.

It may be important to know whether or not a best approximation exists. One reason is that many methods of calculation are derived from properties that are obtained by a best approximation. The following theorem shows existence in the case when $\mathscr{A}$ is compact.

Theorem 1.1

If $\mathscr{A}$ is a compact set in a metric space $\mathscr{B}$, then, for every f in $\mathscr{B}$, there exists an element $a^* \in \mathscr{A}$, such that condition (1.3) holds for all $a \in \mathscr{A}$.

Proof. Let d^* be the quantity

$$d^* = \inf_{a \in \mathscr{A}} d(a, f). \tag{1.5}$$

If there exists a^* in $\mathscr{A}$ such that this bound on the distance is achieved, then there is nothing to prove. Otherwise there is a sequence $\{a_i; i = 1, 2, \ldots\}$ of points in $\mathscr{A}$ which gives the limit

$$\lim_{i\to\infty} d(a_i, f) = d^*. \tag{1.6}$$

By compactness the sequence has at least one limit point in $\mathscr{A}$, a^+ say. Expression (1.6) and the definition of a^+ imply that, for any $\varepsilon > 0$, there exists an integer k such that the inequalities

$$d(a_k, f) < d^* + \tfrac{1}{2}\varepsilon \tag{1.7}$$

and

$$d(a_k, a^+) < \tfrac{1}{2}\varepsilon \tag{1.8}$$

are obtained. Hence the triangle inequality (1.1) provides the bound

$$\begin{aligned} d(a^+, f) &\leq d(a^+, a_k) + d(a_k, f) \\ &< d^* + \varepsilon. \end{aligned} \tag{1.9}$$

Because ε can be arbitrarily small, the distance $d(a^+, f)$ is not greater than d^*. Therefore a^+ is a best approximation. □

When $\mathscr{A}$ is not compact it is easy to find examples to show that best approximations may not exist. For instance, let $\mathscr{B}$ be the Euclidean space $\mathscr{R}^2$ and let $\mathscr{A}$ be the set of points that are strictly inside the unit circle. There is no best approximation to any point of $\mathscr{B}$ that is outside or on the unit circle.

1.3 Approximation in a normed linear space

The properties of metric spaces are not sufficiently strong for most of our work, so it is assumed that $\mathscr{A}$ and f are contained in a normed linear space, which we call $\mathscr{B}$ also when we want to refer to it. The norm is a real-valued function $\|x\|$ that is defined for all $x \in \mathscr{B}$. Its properties are such that the function

$$d(x, y) = \|x - y\| \tag{1.10}$$

is suitable as a distance function. Therefore, by letting z be zero in expression (1.1) and by reversing the sign of y, we may deduce the triangle inequality

$$\|x + y\| \leq \|x\| + \|y\|. \tag{1.11}$$

Moreover, the norm must satisfy the homogeneity condition

$$\|\lambda x\| = |\lambda|\, \|x\| \tag{1.12}$$

for all $x \in \mathscr{B}$ and for all scalars λ.

The specialization from metric spaces to normed linear spaces does not exclude any of the approximation problems that we will consider. Therefore mostly we use the distance function (1.10). It occurs naturally in the approximation calculations that are of practical interest, and it allows the existence of a best approximation to be proved when $\mathscr{A}$ is a linear space.

Theorem 1.2

If $\mathscr{A}$ is a finite-dimensional linear space in a normed linear space $\mathscr{B}$, then, for every $f \in \mathscr{B}$, there exists an element of $\mathscr{A}$ that is a best approximation from $\mathscr{A}$ to f.

Proof. Let the subset $\mathscr{A}_0$ contain the elements of $\mathscr{A}$ that satisfy the condition

$$\|a\| \leqslant 2\|f\|. \tag{1.13}$$

It is compact because it is a closed and bounded subset of a finite-dimensional space. It is not empty: for example it contains the zero element. Therefore, by Theorem 1.1, there is a best approximation from $\mathscr{A}_0$ to f which we call a_0^*. By definition the inequality

$$\|a - f\| \geqslant \|a_0^* - f\|, \qquad a \in \mathscr{A}_0, \tag{1.14}$$

holds. Alternatively, if the element a is in $\mathscr{A}$ but is not in $\mathscr{A}_0$ then, because condition (1.13) is not obtained we have the bound

$$\begin{aligned} \|a - f\| &\geqslant \|a\| - \|f\| \\ &> \|f\| \\ &\geqslant \|a_0^* - f\|, \end{aligned} \tag{1.15}$$

where the last line makes further use of the fact that the zero element is in $\mathscr{A}_0$. Hence expression (1.14) is satisfied for all a in $\mathscr{A}$, which proves that a_0^* is a best approximation. □

1.4 The L_p-norms

In most of the approximation problems that we consider, f and $\mathscr{A}$ are in the space $\mathscr{C}[a\ b]$, which is the set of continuous real-valued functions that are defined on the interval $[a, b]$ of the real line. Occasionally we turn to discrete problems, where f and $\mathscr{A}$ are in $\mathscr{R}^m$, which is the set of real m-component vectors. Both of these spaces are linear and we have a choice of norms.

We study the three norms that are used most frequently, namely the L_p-norms in the cases when $p = 1$, 2 and ∞. For finite p the L_p-norm in

$\mathscr{C}[a, b]$ is defined to have the value

$$\|f\|_p = \left[\int_a^b |f(x)|^p \, dx\right]^{1/p}, \qquad 1 \leq p < \infty, \tag{1.16}$$

and in $\mathscr{R}^m$ it has the value

$$\|f\|_p = \left[\sum_{i=1}^{m} |y_i|^p\right]^{1/p}, \qquad 1 \leq p < \infty, \tag{1.17}$$

where $\{y_i; i = 1, 2, \ldots, m\}$ are the components of f. The ∞-norms are the expressions

$$\|f\|_\infty = \max_{a \leq x \leq b} |f(x)| \tag{1.18}$$

and

$$\|f\|_\infty = \max_{1 \leq i \leq m} |y_i| \tag{1.19}$$

respectively.

There are excellent reasons for giving our attention to the 1-, 2- and ∞-norms. The 1-norm is the least used of the three, but it has one remarkable property that makes it highly suitable for fitting to discrete data in the case when it is possible that there may be some gross errors in the data due to blunders. It is that the magnitude of a blunder makes no difference to the final approximation. This statement will be made clear in Chapter 14. Further, we find later that an understanding of the conditions that are obtained by best approximations in the 1-norm is necessary to analyse some error expressions that occur in the approximation of functionals.

The 2-norm, or perhaps a weighted 2-norm of the form

$$\|f\|_2 = \left[\int_a^b w(x)|f(x)|^2 \, dx\right]^{\frac{1}{2}}, \tag{1.20}$$

where w is a fixed positive function, occurs naturally in theoretical studies of Hilbert spaces. The practical reasons for considering the 2-norm are even stronger. Statistical considerations show that it is the most appropriate choice for data fitting when the errors in the data have a normal distribution. Moreover, when $\mathscr{A}$ is a linear space, the calculation of the best approximation in the 2-norm reduces to a system of linear equations, which allows highly efficient algorithms to be developed. Often the 2-norm is preferred because it is known that the best approximation calculation is straightforward to solve.

The ∞-norm provides the foundation of much of approximation theory, for our next theorem shows that, if we succeed in finding an approximation $a \in \mathscr{A}$ such that the ∞-norm distance function $d(f, a)$ is small, then the 2-norm and 1-norm distance functions are small also. However, an example that follows the theorem shows that the converse statement may not be true. A practical reason for using the ∞-norm is that, when in computer calculations a complicated mathematical function, f say, is estimated by one that is easy to calculate, p say, then it is usually necessary to ensure that the greatest value of the error function $\{|f(x)-p(x)|;\ a \leq x \leq b\}$ is less than a fixed amount, which is just the required accuracy of the approximation. In other words we have a condition on the norm $\|f-p\|_\infty$.

Theorem 1.3
For all e in $\mathscr{C}[a, b]$ the inequalities

$$\|e\|_1 \leq (b-a)^{\frac{1}{2}}\|e\|_2 \leq (b-a)\|e\|_\infty \tag{1.21}$$

hold.

Proof. The Cauchy–Schwarz inequality provides the bound

$$\begin{aligned} \|e\|_1 &= \int_a^b |e(x)|\,|1|\,\mathrm{d}x \\ &\leq \left[\int_a^b |e(x)|^2\,\mathrm{d}x\right]^{\frac{1}{2}} \left[\int_a^b \mathrm{d}x\right]^{\frac{1}{2}} \\ &= (b-a)^{\frac{1}{2}}\|e\|_2, \end{aligned} \tag{1.22}$$

which is the first part of the required result. Moreover, by replacing an integrand by its maximum value, we obtain the inequality

$$\begin{aligned} \|e\|_2 &= \left[\int_a^b |e(x)|^2\,\mathrm{d}x\right]^{\frac{1}{2}} \\ &\leq \left[\int_a^b \|e\|_\infty^2\,\mathrm{d}x\right]^{\frac{1}{2}} \\ &= (b-a)^{\frac{1}{2}}\|e\|_\infty, \end{aligned} \tag{1.23}$$

which completes the proof of the theorem. □

It is interesting to consider the statement of Theorem 1.3, when e is the error in approximating the constant function $\{f(x)=1;\ 0 \leq x \leq 1\}$ by $\{x^\lambda;\ 0 \leq x \leq 1\}$, where λ is a positive parameter. Straightforward calculation shows that the norms have the values

$$\|e\|_1 = \lambda/(\lambda+1), \tag{1.24}$$

$$\|e\|_2 = [2\lambda^2/(\lambda+1)(2\lambda+1)]^{\frac{1}{2}}, \tag{1.25}$$

and

$$\|e\|_\infty = 1. \tag{1.26}$$

We see that, if λ becomes arbitrarily small, then $\|e\|_1$ and $\|e\|_2$ tend to zero, but $\|e\|_\infty$ remains at one. Hence it is not always possible to reduce the ∞-norm of an error function by making small its 2-norm or its 1-norm. In order to develop algorithms that give approximations with small errors in the 1-, 2- and ∞-norms, we just have to ensure that the algorithm is suitable for the ∞-norm.

The ∞-norm is sometimes called the uniform or minimax norm, and the 2-norm is sometimes called the least squares or Euclidean norm.

1.5 A geometric view of best approximations

In the case when f and $\mathscr{A}$ are contained in a normed linear space $\mathscr{B}$, and when we require a best approximation from $\mathscr{A}$ to f, it is sometimes helpful to think of the balls of different radii whose centres are at f. The ball of radius r is defined to be the set

$$\mathscr{N}(f, r) \equiv \{g: \|g-f\| \leq r, g \in \mathscr{B}\}. \tag{1.27}$$

It follows that, if $r_1 > r_0$, then $\mathscr{N}(f, r_0) \subset \mathscr{N}(f, r_1)$. Hence, if $f \notin \mathscr{A}$, and if r is allowed to increase from zero there exists a scalar, r^* say, such that, for $r > r^*$, there are points of $\mathscr{A}$ that are in $\mathscr{N}(f, r)$, but, for $r < r^*$, the intersection of $\mathscr{N}(f, r)$ and $\mathscr{A}$ is empty. The value of r^* is equal to expression (1.5), and we know from Theorem 1.2 that, if $\mathscr{A}$ is a finite-dimensional linear space, then the equation

$$r^* = \inf_{a \in \mathscr{A}} \|f-a\| = \|f-a^*\| \tag{1.28}$$

is obtained for a point a^* in $\mathscr{A}$.

For example, suppose that $\mathscr{B}$ is the two-dimensional Euclidean space $\mathscr{R}^2$, and that we are using the 2-norm. Let f be the point whose components are (2, 1), and let $\mathscr{A}$ be the linear space of vectors

$$\mathscr{A} = \{(\lambda, \lambda); -\infty < \lambda < \infty\}, \tag{1.29}$$

where λ is a real parameter. Figure 1.4 shows the set $\mathscr{A}$ and the three balls, centre f, whose radii are $\frac{1}{2}$, $\sqrt{\frac{1}{2}}$ and 1. If we imagine that the value of r is allowed to increase from zero, we see that the best approximation is the point where the ball of radius $\sqrt{\frac{1}{2}}$ touches $\mathscr{A}$.

The shapes of balls in two-dimensional space for the 1-, 2- and ∞-norms are interesting, because they indicate some of the implications of the choice of norm. The boundaries of the three unit balls centred on the origin are shown in Figure 1.5. We note that, if the 2-norm is replaced

Figure 1.4. An approximation problem in $\mathscr{R}^2$.

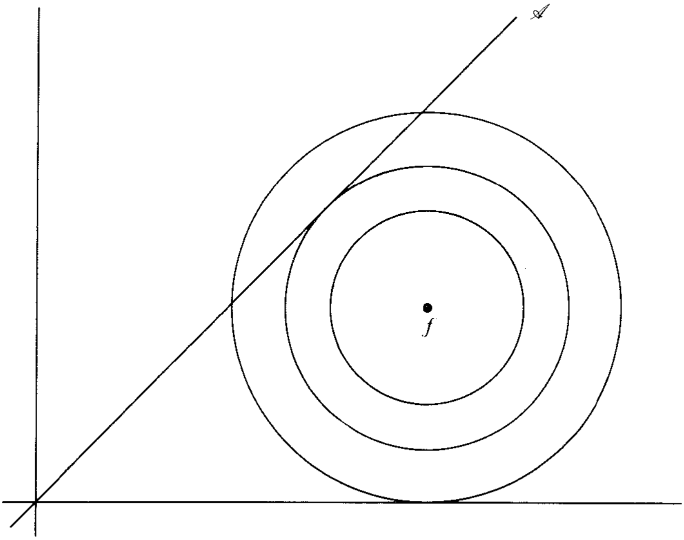
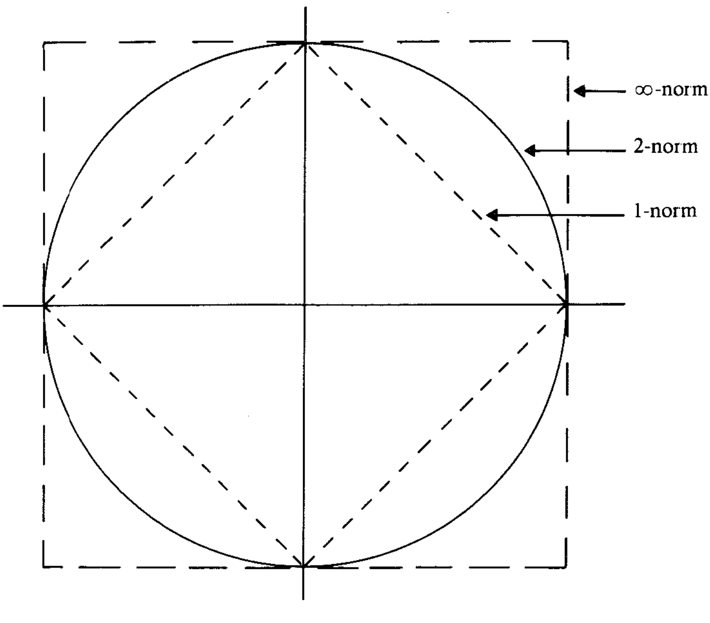


Figure 1.5. The unit balls of the 1-, 2- and ∞-norms.



by the 1-norm in Figure 1.4, and if the radius of the ball centred at f is again allowed to increase from zero, then we find that many points of $\mathscr{A}$ are best approximations to f. The question of the uniqueness of best approximations is considered in the next chapter.