CHAPTER 13

Numerical differentiation of functions of two variables

So far, most of the functions we have encountered have only depended on one variable, but both within mathematics and in applications there is often a need for functions of several variables. In this chapter we will deduce methods for numerical differentiation of functions of two variables. The methods are simple extensions of the numerical differentiation methods for functions of one variable.

13.1 Functions of two variables

In this section we will review some basic results on functions of two variables, in particular the definition of partial and directional derivatives. For proofs, the reader is referred to a suitable calculus book.

13.1.1 Basic definitions

Functions of two variables are natural generalisations of functions of one variable that act on pairs of numbers rather than a single number. We assume that you are familiar with their basic properties already, but we repeat the definition and some basic notation.

Definition 13.1 (Function of two variables). *A (scalar) function f of two variables is a rule that to a pair of numbers (x, y) assigns a number $f(x, y)$.*

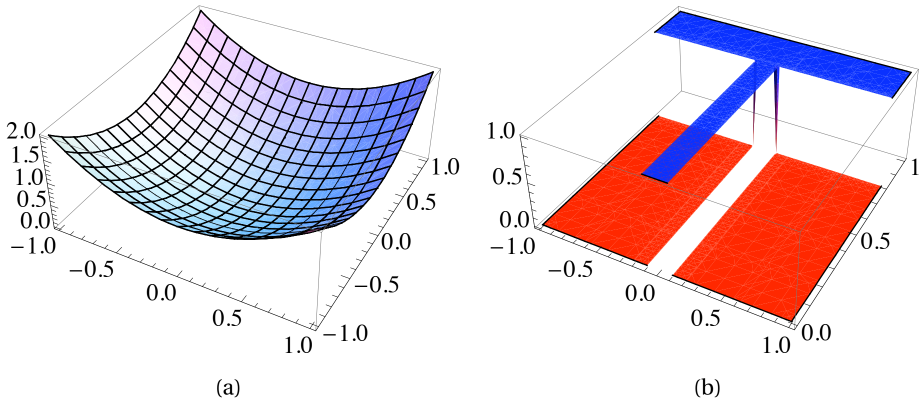

Figure 13.1. The plot in (a) shows the function $f(x,y) = x^2 + y^2$ with x and y varying in the interval $[-1,1]$. The function in (b) is defined by the rule that $f(x,y) = 0$ except in a small area around the y-axis and the line $y = 1$, where the value is $f(x,y) = 1$.

The obvious interpretation is that $f(x,y)$ gives the height above the point in the plane given by (x,y). This interpretation lets us plot functions of two variables, see figure 13.1.

The rule f can be given by a formula like $f(x,y) = x^2 + y^2$, but this is not necessary, we just need to be able to determine $f(x,y)$ from x and y. In figure 13.1 the function in (a) is given by a formula, while the function in (b) is given by the rule

$$f(x,y) = \begin{cases} 1, & \text{if } |x| \le 0.1. \text{ and } 0 \le y \le 1; \\ 1, & \text{if } |y-1| \le 0.1 \text{ and } -1 \le x \le 1; \\ 0, & \text{otherwise.} \end{cases}$$

We will sometimes use vector notation and refer to (x,y) as the point $\boldsymbol{x}$; then $f(x,y)$ can be written simply as $f(\boldsymbol{x})$. There is also convenient notation for a set of pairs of numbers that are assembled from two intervals.

Notation 13.2. *Let the two sets of numbers $\mathbb{A}$ and $\mathbb{B}$ be given. The set of all pairs of numbers from $\mathbb{A}$ and $\mathbb{B}$ is denoted $\mathbb{A} \times \mathbb{B}$,*

$$\mathbb{A} \times \mathbb{B} = \{(a,b) \mid a \in \mathbb{A} \text{ and } b \in \mathbb{B}\}.$$

The set $\mathbb{A} \times \mathbb{A}$ is denoted $\mathbb{A}^2$.

The most common set of pairs of numbers is $\mathbb{R}^2$, the set of all pairs of real numbers.

To define differentiation we need the concept of an interior point of a set. This is defined in terms of small discs.

Notation 13.3. *The disc with radius* r *and centre* $\boldsymbol{x} \in \mathbb{R}^2$ *is denoted* $B(\boldsymbol{x}; r)$*. A point* $\boldsymbol{x}$ *in a subset* $\mathbb{A}$ *of* $\mathbb{R}^2$ *is called an* interior point *of* $\mathbb{A}$ *if there is a real number* $\epsilon > 0$ *such that the disc* $B(\boldsymbol{x}; \epsilon)$ *lies completely in* $\mathbb{A}$*. The disc* $B(\boldsymbol{x}; \epsilon)$ *is called a* neighbourhood *of* $\boldsymbol{x}$*.*

More informally, an interior point of $\mathbb{A}$ is a point which is completely surrounded by points from $\mathbb{A}$.

13.1.2 Differentiation

Differentiation generalises to functions of two variables in a simple way: We keep one variable fixed and differentiate the resulting function as a function of one variable.

Definition 13.4 (Partial derivatives)**.** *Let* f *be a function defined on a set* $\mathbb{A} \subseteq \mathbb{R}^2$*. The* partial derivatives *of* f *at an interior point* $(a, b) \in \mathbb{A}$ *are given by*

$$\frac{\partial f}{\partial x}(a,b) = \lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h},$$
$$\frac{\partial f}{\partial y}(a,b) = \lim_{h \to 0} \frac{f(a,b+h) - f(a,b)}{h}.$$

From the definition we see that the partial derivative $\partial f / \partial x$ is obtained by fixing $y = b$ and differentiating the function $g_1(x) = f(x, b)$ at $x = a$. Similarly, the partial derivative with respect to y is obtained by fixing $x = a$ and differentiating the function $g_2(y) = f(a, y)$ at $y = b$.

Geometrically, the partial derivatives give the slope of f at (a, b) in the directions parallel to the two coordinate axes. The *directional derivative* gives the slope in a general direction.

Definition 13.5. *Suppose the function* f *is defined on the set* $\mathbb{A} \subseteq \mathbb{R}^2$ *and that* $\boldsymbol{a}$ *is an interior point of* $\mathbb{A}$*. The* directional derivative *at* $\boldsymbol{a}$ *in the direction* $\boldsymbol{r}$ *is given by the limit*

$$f'(\boldsymbol{a}; \boldsymbol{r}) = \lim_{h \to 0} \frac{f(\boldsymbol{a} + h\boldsymbol{r}) - f(\boldsymbol{a})}{h},$$

provided the limit exists.

It turns out that for reasonable functions, the directional derivative can be computed in terms of partial derivatives.

Theorem 13.6. *Suppose the function is defined on the set $\mathbb{A} \subseteq \mathbb{R}^2$ and that $\boldsymbol{a}$ is an interior point of $\mathbb{A}$. If the two partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ exist in a neighbourhood of $\boldsymbol{a}$ and are continuous at $\boldsymbol{a}$, then the directional derivative $f'(\boldsymbol{a};\boldsymbol{r})$ exists for all directions $\boldsymbol{r} = (r_1, r_2)$ and*

$$f'(\boldsymbol{a};\boldsymbol{r}) = r_1 \frac{\partial f}{\partial x}(\boldsymbol{a}) + r_2 \frac{\partial f}{\partial y}(\boldsymbol{a}).$$

The conditions in theorem 13.6 are not very strict, but should be kept in mind. In particular you should be on guard when you need to compute directional derivatives near points where the partial derivatives do not exist.

If we consider a function like $f(x, y) = x^3 y + x^2 y^2$, the partial derivatives are $\partial f/\partial x = 3x^2 y + 2xy^2$ and $\partial f/\partial y = x^3 + 2x^2 y$. Each of these can of course be differentiated again,

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= 6xy + 2y^2, & \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = 3x^2 + 4xy, \\
\frac{\partial^2 f}{\partial y^2} &= 2x^2, & \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = 3x^2 y + 4xy.
\end{aligned}$$

We notice that the two mixed derivatives are equal. In general the derivatives

$$\frac{\partial^2 f}{\partial x \partial y}(\boldsymbol{a}), \quad \frac{\partial^2 f}{\partial y \partial x}(\boldsymbol{a})$$

are equal if they both exist in a neighbourhood of $\boldsymbol{a}$ and are continuous at $\boldsymbol{a}$. All the functions we consider here have mixed derivatives that are equal. We can of course consider partial derivatives of any order.

Notation 13.7 (Higher order derivatives)**.** *The expression*

$$\frac{\partial^{n+m} f}{\partial x^n \partial y^m}$$

denotes the result of differentiating f, first m times with respect to y, and then differentiating the result n times with respect to x.

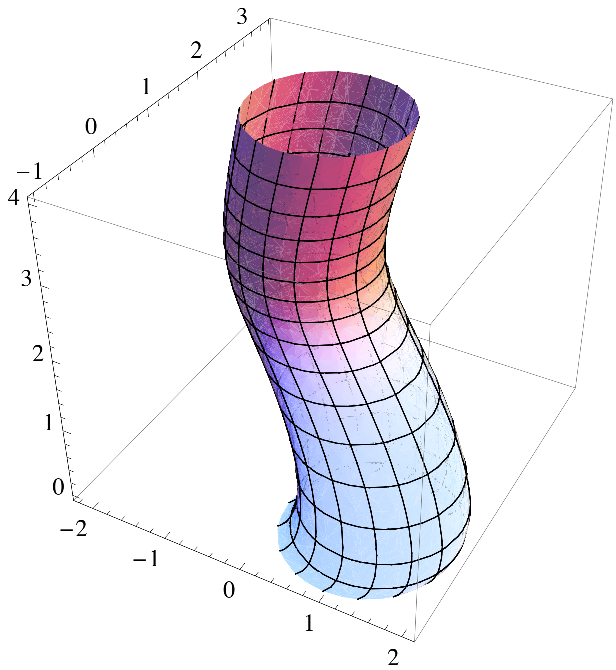

Figure 13.2. An example of a parametric surface.

13.1.3 Vector functions of several variables

The theory of functions of two variables extends nicely to functions of an arbitrary number of variables and functions where the scalar function value is replaced by a vector. We are only going to define these functions, but the whole theory of differentiation works in this more general setting.

Definition 13.8 (General functions)**.** *A function* $\boldsymbol{f}:\mathbb{R}^n \mapsto \mathbb{R}^m$ *is a rule that to* n *numbers* $\boldsymbol{x} = (x_1, \ldots, x_n)$ *assigns* m *numbers* $\boldsymbol{f}(\boldsymbol{x}) = \big(f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})\big)$.

Apart from the case $n = 2$, $m = 1$ which we considered above, we are interested in the case $n = 2$, $m = 3$.

Definition 13.9. *A function from* $\boldsymbol{f}:\mathbb{R}^2 \mapsto \mathbb{R}^3$ *is called a* parametric surface.

An example of a parametric surface is shown in figure 13.2. Parametric surfaces can take on almost any shape and are therefore used for representing geometric form in computer programs for geometric design. These kinds of programs are used for designing cars, aircrafts and other industrial objects as well as the 3D objects and characters in animated movies.

13.2 Numerical differentiation

The reason that we may want to compute derivatives numerically are the same for functions of two variables as for functions of one variable: The function may only be known via some procedure or computer program that can compute function values.

Theorem 13.6 shows that we can compute directional derivatives very easily as long as we can compute partial derivatives. The basic problem in numerical differentiation is therefore to find numerical approximations to the partial derivatives. Since only one variable varies in the definition of a first-order partial derivative, we can actually use the approximations that we obtained for functions of one variable. The simplest approximation is the following.

Proposition 13.10. *Let f be a function defined on a set $\mathbb{A} \subseteq \mathbb{R}^2$ and suppose that the points (a,b), $(a+rh_1,b)$ and $(a,b+rh_2)$ all lie in $\mathbb{A}$ for any $r \in [0,1]$. Then the two partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ can be approximated by*

$$\frac{\partial f}{\partial x}(a,b) \approx \frac{f(a+h_1,b)-f(a,b)}{h_1},$$
$$\frac{\partial f}{\partial y}(a,b) \approx \frac{f(a,b+h_2)-f(a,b)}{h_2}.$$

The errors in the two estimates are

$$\frac{\partial f}{\partial x}(a,b) - \frac{f(a+h_1,b)-f(a,b)}{h_1} = \frac{h_1}{2}\frac{\partial^2 f}{\partial x^2}(c_1,b), \tag{13.1}$$
$$\frac{\partial f}{\partial y}(a,b) - \frac{f(a,b+h_2)-f(a,b)}{h_2} = \frac{h_2}{2}\frac{\partial^2 f}{\partial y^2}(a,c_2), \tag{13.2}$$

where c_1 is a number in $(a,a+h_1)$ and c_2 is a number in $(a,a+h_2)$.

Proof. We will just consider the first approximation. For this we define the function $g(x) = f(x,b)$. From 'Setning 9.15' in the Norwegian notes we know that

$$g'(x) = \frac{g(a+h_1)-g(a)}{h_1} + \frac{h_1}{2}g''(c_1)$$

where c_1 is a number in the interval $(a,a+h_1)$. From this the relation (13.1) follows. ∎

The other approximations to the derivatives in chapter 9 of the Norwegian notes lead directly to approximations of partial derivatives that are not mixed. For example we have

$$\frac{\partial f}{\partial x} = \frac{f(a+h,b)-f(a-h,b)}{2h} + \frac{h^2}{6}\frac{\partial^3 f}{\partial x^3}(c,b) \tag{13.3}$$

where $c \in (a-h, a+h)$. A common approximation of a second derivative is

$$\frac{\partial^2 f}{\partial x^2} \approx \frac{-f(a-h,b)+2f(a,b)-f(a+h,b)}{h^2},$$

with error bounded by

$$\frac{h^2}{12}\max_{z\in(a-h,a+h)}\left|\frac{\partial^4 f}{\partial x^4}(z,b)\right|,$$

see exercise 9.11 in the Norwegian notes. These approximations of course work equally well for non-mixed derivatives with respect to y.

Approximation of mixed derivatives requires that we use estimates for the derivatives both in the x- and y-directions. This makes it more difficult to keep track of the error. In fact, the easiest way to estimate the error is with the help of Taylor polynomials with remainders for functions of two variables. However, this is beyond the scope of these notes.

Let us consider an example of how an approximation to a mixed derivative can be deduced.

Example 13.11. Let us consider the simplest mixed derivative,

$$\frac{\partial^2 f}{\partial x \partial y}(a,b).$$

If we set

$$g(a) = \frac{\partial f}{\partial y}(a,b) \tag{13.4}$$

we can use the approximation

$$g'(a) \approx \frac{g(a+h_1)-g(a-h_1)}{2h_1}.$$

If we insert (13.4) in this approximation we obtain

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) \approx \frac{\dfrac{\partial f}{\partial y}(a+h_1,b) - \dfrac{\partial f}{\partial y}(a-h_1,b)}{2h_1}. \tag{13.5}$$

Now we can use the same kind of approximation for the two first-order partial derivatives in (13.5),

$$\frac{\partial f}{\partial y}(a+h_1,b) \approx \frac{f(a+h_1,b+h_2)-f(a+h_1,b-h_2)}{2h_2},$$
$$\frac{\partial f}{\partial y}(a-h_1,b) \approx \frac{f(a-h_1,b+h_2)-f(a-h_1,b-h_2)}{2h_2}.$$

If we insert these expressions in (13.5) we obtain the final approximation

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) \approx \frac{f(a+h_1,b+h_2)-f(a+h_1,b-h_2)-f(a-h_1,b+h_2)+f(a-h_1,b-h_2)}{4h_1h_2}.$$

If we introduce the notation

$$\begin{aligned} f(a-h_1,b-h_2) &= f_{-1,-1}, & f(a+h_1,b-h_2) &= f_{1,-1},\\ f(a-h_1,b+h_2) &= f_{-1,1}, & f(a+h_1,b+h_2) &= f_{1,1}, \end{aligned} \tag{13.6}$$

we can write the approximation more compactly as

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) \approx \frac{f_{1,1}-f_{1,-1}-f_{-1,1}+f_{-1,-1}}{4h_1h_2}.$$

These approximations require f to be a 'nice' function. A sufficient condition is that all partial derivatives up to order four are continuous in a disc that contains the rectangle with corners $(a-h_1,b-h_2)$ and $(a+h_1,b+h_2)$. ∎

We record the approximation in example 13.11 in a proposition. We do not have the right tools to estimate the error, but just indicate how it behaves.

Proposition 13.12 (Approximation of a mixed derivative). *Suppose that f has continuous derivatives up to order four in a disc that contains the rectangle with corners $(a-h_1,b-h_2)$ and $(a+h_1,b+h_2)$. Then the mixed second derivative of f can be approximated by*

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) \approx \frac{f_{1,1}-f_{1,-1}-f_{-1,1}+f_{-1,-1}}{4h_1h_2}, \tag{13.7}$$

where the notation is defined in (13.6). The error is proportional to $h_1^2h_2^2$.

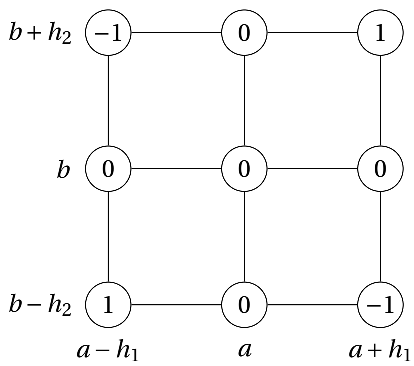

Figure 13.3. The weights involved in computing the mixed second derivative with the approximation in example 13.11. This kind of figure is referred to as the *computational molecule* of the approximation.

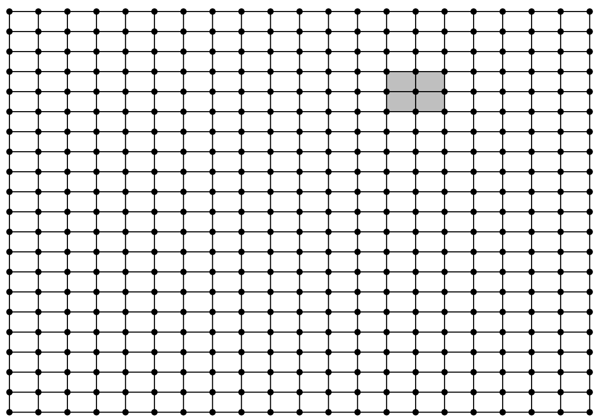

Figure 13.4. Numerical approximations to partial derivatives are often computed at all points of a grid like the one shown here by sliding around the grid a computational molecule like the one in figure 13.3.

Numerical approximations of other mixed partial derivatives can be derived with the same technique as in example 13.11, see exercise 1.

A formula like (13.7) is often visualised with a drawing like the one in figure 13.3 which is called a *computational molecule*. The arguments of the function values involved in the approximation are placed in a rectangular grid together with the corresponding coefficients of the function values. More complicated approximations will usually be based on additional values and involve more complicated coefficients.

Approximations to derivatives are usually computed at many points, and often the points form a rectangular grid as in figure 13.4. The computations can be performed by moving the computational molecule of the approximation across the grid and computing the approximation at each point, as indicated by the grey area in figure 13.4.

Exercises

13.1 In this exercise we are going to derive approximations to mixed derivatives.

a) Use the approximation $g'(a) = \big(g(a+h) - g(a-h)\big)/(2h)$ repeatedly as in example 13.11 and deduce the approximation

$$\frac{\partial^3 f}{\partial x^2 \partial y} \approx \frac{f_{2,1} - 2f_{0,1} + f_{-2,1} - f_{2,-1} + 2f_{0,-1} - f_{-2,-1}}{8h_1^2 h_2}.$$

Hint: Use the approximation in (13.7).

b) Use the same technique as in (a) and deduce the approximation

$$\frac{\partial^4 f}{\partial x^2 \partial y^2} \approx \frac{f_{2,2} - 2f_{0,2} + f_{-2,2} - 2f_{2,0} + 4f_{0,0} - 2f_{-2,0} + f_{2,-2} - 2f_{0,-2} + f_{-2,-2}}{16h_1^2 h_2^2}.$$

Hint: Use the approximation in (a) as a starting point.

13.2 Determine approximations to the two mixed derivatives

$$\frac{\partial^3 f}{\partial x^2 \partial y}, \quad \frac{\partial^4 f}{\partial x^2 \partial y^2},$$

in exercise 1, but use the approximation $g'(a) = \big(g(a+h) - g(a)\big)/h$ at every stage.