1 Partial differentiation and the chain rule

In this section we review and discuss certain notations and relations involving partial derivatives.

The more general case can be illustrated by considering a function $f(x, y, z)$ of three variables x, y and z. If y and z are held constant and only x is allowed to vary, the partial derivative of f with respect to x is denoted by $\frac{\partial f}{\partial x}$ and defined by

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y, z) - f(x, y, z)}{\Delta x} \tag{1}$$

Similarly we define

$$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y, z) - f(x, y, z)}{\Delta y} \tag{2}$$

$$\frac{\partial f}{\partial z} = \lim_{\Delta z \to 0} \frac{f(x, y, z + \Delta z) - f(x, y, z)}{\Delta z} \tag{3}$$

These formulae are direct generalisations of the well known definition of the derivative of a function $f(x)$ of one variable x

$$\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{4}$$

Example

Let $f(x, y, z) = x^2yz + ye^z$, then

$$\frac{\partial f}{\partial x} = 2xyz$$
$$\frac{\partial f}{\partial y} = x^2z + e^z$$
$$\frac{\partial f}{\partial z} = x^2y + ye^z$$

1.1 The chain rule

We present the chain for the function $f(x, y, z)$. We consider 3 cases.

(1) If x, y and z are all functions of a single variable t, then f can be considered as a function of t and

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} \tag{5}$$

Example

Let $f(x, y, z) = x^2yz$, $x = e^t$, $y = t$ and $z = 1 + t$.

Method 1: Substitute the expressions for x, y and z into f. This yields

$$f = e^{2t}t(1 + t) \tag{6}$$

and differentiate (6). This gives

$$\frac{df}{dt} = e^{2t}(2t + 1) + 2e^{2t}t(1 + t) \tag{7}$$

Method 2: Use the chain rule (5). This gives

$$\frac{df}{dt} = (2xyz)e^t + x^2z + x^2y \tag{8}$$

Subsituting the expressions for x, y and z into (8) gives (7).

(2) More generally if x, y and z are all functions of two (or more variables), say s and t, then we can consider f as a function of s and t and

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t} \tag{9}$$

In writing (9) we consider the variables t and s as 'associated' in the sense that in $\frac{\partial f}{\partial t}$ we assume that s is held constant. Similarly the variables x, y and z are associated in the sense that in $\frac{\partial f}{\partial x}$ we assume that y and z are held constant. Alternatively (and this is more elegant and safer) we can define

$$F(s,t) = f[x(s,t), y(s,t), z(s,t)] \tag{10}$$

and write

$$\frac{\partial F}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t} \tag{11}$$

Example

Let $f(x,y,z) = xyz$, $x = st$, $y = s + t$, $z = t$. We consider f as a function of s and t and we want to calculate $\frac{\partial f}{\partial t}$.

Method 1: Substitute the expressions for x, y and z into f. This gives

$$F(s,t) = s^2t^2 + st^3 \tag{12}$$

Diffirentiating 12 with respect to t gives

$$\frac{dF}{dt} = 2s^2t + 3st^2 \tag{13}$$

Method 2: Use the chain rule (9) or (10). Here we use (9). This gives

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t} = yzs + xz + xy \tag{14}$$

Subsituting the expressions for x, y and z into (14) gives (13).

(3) We know suppose that y and z are functions of x. Then $f(x,y,z)$ can be considered as a function of x and we have

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial z}\frac{dz}{dx} \tag{15}$$

Example

Let $f(x,y,z) = xyz$, $y = x$ and $z = x^2$

Method 1: Substitute the expressions for y and z into f. This gives

$$f = xyz = x^4 \tag{16}$$

and

$$\frac{df}{dx} = 4x^3 \tag{17}$$

Method 2: Use the chain rule 15. This gives

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial z}\frac{dz}{dx} = yz + xz + 2x^2y \tag{18}$$

Substituting the expressions for y and z into (18) gives (17).

1.2 Higher order derivatives

Consider the function $f(x, y, z)$. We define the second order partial derivatives by the formulae

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\frac{\partial f}{\partial x} \tag{19}$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y} \tag{20}$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\frac{\partial f}{\partial x} \tag{21}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\frac{\partial f}{\partial y} \tag{22}$$

It is not always true that

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \tag{23}$$

However (23) holds if all the partial derivatives of f up to second order are continuous. This condition is usually satisfied in applications and in particular in all the examples considered in this course.

The following alternative notation for partial derivatives is often convenient and more economical.

$$f_x = \frac{\partial f}{\partial x}$$

$$f_y = \frac{\partial f}{\partial y}$$

$$f_{xx} = \frac{\partial^2 f}{\partial x^2}$$

$$f_{yy} = \frac{\partial^2 f}{\partial y^2}$$

$$f_{xy} = \frac{\partial^2 f}{\partial y \partial x}$$

$$f_{yx} = \frac{\partial^2 f}{\partial x \partial y}$$

Example

Let $f(x, y) = x^3 \cos y$. Then

$$\frac{\partial f}{\partial x} = 3x^2 \cos y$$

$$\frac{\partial f}{\partial y} = -x^3 \sin y$$

$$\frac{\partial^2 f}{\partial x^2} = 6x \cos y$$

$$\frac{\partial^2 f}{\partial y^2} = -x^3 \cos y$$

$$\frac{\partial^2 f}{\partial y \partial x} = -3x^2 \sin y$$

$$\frac{\partial^2 f}{\partial x \partial y} = -3x^2 \sin y$$

We note that $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$ in accordance with our previous remark.

2 Taylor series

In previous courses you encountered Taylor series for a function of one variable . For example the Taylor series of a function $F(t)$ about $t = 0$ is

$$F(t) = F(0) + F'(0)t + F''(0)\frac{t^2}{2!} + R_3 \tag{24}$$

where

$$R_3 = F'''(\tau)\frac{t^3}{3!} \tag{25}$$

for some τ between 0 and t. Here we only wrote the first 3 terms. We refer to (24) as a three term Taylor series. Similarly the two term Taylor series of $F(t)$ is

$$F(t) = F(0) + F'(0)t + R_2 \tag{26}$$

where

$$R_2 = F''(\tau)\frac{t^2}{2!} \tag{27}$$

for some τ between 0 and t.

In this section we generalise (24) and (26) to a function of several variables. We present the details for a function $f(x, y)$ of two variables. The idea is to apply (24) and (26) with the particular choice

$$F(t) = f(x + ht, y + kt) \tag{28}$$

and then to set $t = 1$.

We first note that (28) implies

$$F(0) = f(x, y) \tag{29}$$

Next applying the chain rule to (28) we obtain sucessively

$$F'(t) = hf_x(x + ht, y + kt) + kf_y(x + ht, y + kt) \tag{30}$$

$$F''(t) = h\frac{\partial}{\partial x}[hf_x(x+ht, y+kt)+kf_y(x+ht, y+kt)]+k\frac{\partial}{\partial y}[hf_x(x+ht, y+kt)+kf_y(x+ht, y+kt)] =$$

$$h^2 f_{xx}(x + ht, y + kt) + 2hkf_{xy}(x + ht, y + kt) + k^2 f_{yy}(x + ht, y + kt) \tag{31}$$

$$F'''(t) = h\frac{\partial}{\partial x}[h^2 f_{xx}(x + ht, y + kt) + 2hkf_{xy}(x + ht, y + kt) + k^2 f_{yy}(x + ht, y + kt)]+$$

$$k\frac{\partial}{\partial y}[h^2 f_{xx}(x + ht, y + kt) + 2hkf_{xy}(x + ht, y + kt) + k^2 f_{yy}(x + ht, y + kt)] =$$

$$h^3 f_{xxx} + 3h^2 kf_{xxy} + 3hk^2 f_{xyy} + k^3 f_{yyy} \tag{32}$$

Substituting (28)-(32) into (24) gives

$$f(x+ht,y+kt) = f(x,y)+[hf_x(x,y)+kf_y(x,y)]t+[h^2f_{xx}(x,y)+2hkf_{xy}(x,y)+k^2f_{yy}]\frac{t^2}{2}+R_3 \quad (33)$$

where

$$R_3 = \frac{t^3}{3!}[h^3f_{xxx}+3h^2kf_{xxy}+3kh^2f_{xyy}+k^3f_{yyy}]_{(x+\tau h,y+\tau k)} \quad (34)$$

for some τ between 0 and t.

Similarly substituting (28)-(32) into (26) we obtain

$$f(x+ht,y+kt) = f(x,y)+[hf_x(x,y)+kf_y(x,y)]t+R_2 \quad (35)$$

where

$$R_2 = \frac{t^2}{2}[h^2f_{xx}+2hkf_{xy}+k^2f_{yy}]_{(x+\tau h,y+\tau k)} \quad (36)$$

for some τ between 0 and t.

Our final formulae are obtained by setting $t = 1$ in (33)-(36). This yields the three term Taylor series

$$f(x+h,y+k) = f(x,y)+[hf_x(x,y)+kf_y(x,y)]+\frac{1}{2}[h^2f_{xx}(x,y)+2hkf_{xy}(x,y)+k^2f_{yy}]+R_3 \quad (37)$$

$$R_3 = \frac{1}{3!}[h^3f_{xxx}+3h^2kf_{xxy}+3kh^2f_{xyy}+k^3f_{yyy}]_{(x+\tau h,y+\tau k)} \quad (38)$$

and the two term Taylor series

$$f(x+h,y+k) = f(x,y)+[hf_x(x,y)+kf_y(x,y)]+R_2 \quad (39)$$

$$R_2 = \frac{1}{2}[h^2f_{xx}+2hkf_{xy}+k^2f_{yy}]_{(x+\tau h,y+\tau k)} \quad (40)$$

In (38) and (40), τ is some value between 0 and 1. Formulae (37) and (39) can also be rewritten in the more familiar way

$$f(x,y) = f(x_0,y_0)+(x-x_0)f_x(x_0,y_0)+(y-y_0)f_y(x_0,y_0)+$$

$$\frac{1}{2}[(x-x_0)^2f_{xx}(x_0,y_0)+2(x-x_0)(y-y_0)f_{xy}(x_0,y_0)+(y-y_0)^2f_{yy}(x_0,y_0)]+R_3 \quad (41)$$

$$f(x,y) = f(x_0,y_0)+(x-x_0)f_x(x_0,y_0)+(y-y_0)f_y(x_0,y_0)++R_2 \quad (42)$$

The Taylor formulae (37), (39), (41) and (42) have many applications. One of them is the investigation of the maxima and the minima of functions of several variables (see next section). Another is illustrated in the next example

Example

Consider the function $f(x,y) = x^3y^4$. Clearly $f(1,1) = 1$. We wish to use this information and the Taylor formulae (41) and (42) to find an approximate value for $f(1.03, 1.05)$. Setting $x_0 = 1$, $y_0 = 1$, $x = 1.03$ and $y = 1.05$ in (42) and (41) give $f(1.03, 1.05) \approx 1.29$ and $f(1.03, 1.05) \approx 1.3255$ respectively. On the other hand the exact value of $f(1.03, 1.05)$ is 1.3282 (to 4 decimal places). As expected (41) give a better approximation than (42).

In the previous calculation the accuracy of the approximation given by (42) is poor. Is it possible to improve it? The answer is yes. We just need to choose values of x and y closer to $x_0 = 1$ and $y_0 = 1$. For example the exact value of $f(1.003, 1.005)$ is 1.02936 (to 5 decimal places). Formula (42) predicts the approximate value 1.029.