Multivariable Calculus

As Wikipedia states: 'Multivariable calculus is used in many fields of natural and social science and engineering to model and study high-dimensional systems that exhibit deterministic behaviour'. Multivariable calculus is a popular topic (chapter 2) in FP3. As is made clear in the introduction to that chapter, an important reason for studying functions of more than one variable is that in many applications the value of the quantity you want to know depends on several other quantities (variables). Although a realistic model may use many variables, we shall concentrate on functions of two or three variables since the new ideas can be seen clearly in these cases.

Functions of two variables

An example is $z = f(x, y) = x^2 + 2y^2 + 1$

This can be represented as a 2-dimensional *surface* in 3 dimensional space.

It is sometimes helpful to draw a *contour diagram* consisting of curves of the form $z = k$ e.g. the contour $z = 5$ gives the ellipse $x^2 + 2y^2 = 4$.

Alternatively, *sections* of the form $x = a$ or $y = b$ can be useful.
In the above example, the section $x = 3$ reduces to the parabola $z = 2y^2 + 10$.

Derivatives

The key concept in this chapter is *differentiation*. Unlike for functions of one variable, the two interpretations of *differentiation* as

1) **Find the rate of change of a function**
2) **Find the tangent line at a point**

are essentially different for two or more variables.

Partial Derivatives

In higher dimensions, the rate of change of a function is complicated by having to specify 'rate of change with respect to what?'

In the example, z depends on x and y. Just knowing how one of these variables changes is insufficient to determine the behaviour of z. However, it does give important information. The process of finding the rate of change of one variable with respect to another, while keeping all other variables fixed is known as *partial differentiation*.

In the above example, differentiating z with respect to x while keeping y fixed would be written: $\frac{\partial z}{\partial x} = 2x$ or $\frac{\partial f}{\partial x} = 2x$ or $f_x(x, y) = 2x$

The process of obtaining partial derivatives is the single most important technique in this chapter and students must know how to do it, but, if they are competent at 'normal' differentiation, they will pick it up quickly.

The split nature of differentiation in two or more dimensions becomes clear when we consider the analogy to a tangent line. Functions of two variables like the above can be represented as surfaces in three dimensions and the natural geometric equivalent of a tangent line is a *tangent plane*.

The equation of the tangent plane at a point on the surface is obtained by considering the partial derivatives at this point.

So, if $z = \mathrm{f}(x,y)$, the tangent plane at the point P = (a,b,c) is given by:

$$z - c = f_x(x-a) + f_y(y-b) \qquad (1)$$

$$\text{or } \mathrm{d}z = \frac{\partial z}{\partial x}\mathrm{d}x + \frac{\partial z}{\partial y}\mathrm{d}y \text{ where the partial derivatives are evaluated at P} \qquad (2)$$

$$\text{or } \mathrm{d}z = \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y \qquad (3)$$

(2) is analogous to: $\mathrm{d}y = \frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x$ where dx and dy are *differentials* (see appendix 1).

This situation and these equations are best explained by a diagram! (see appendix 2).

In the example $z = x^2 + 2y^2 + 1$, at the point (2,3) the partial derivatives are:

$$f_x(2,3) = 4 \text{ and } f_y(2,3) = 12$$

The equation of the tangent plane at this point is

$$z - 23 = 4(x-2) + 12(y-3)$$

$$\text{or } 4x + 12y - z = 21$$

A one-dimensional function is *differentiable* at a point if we can construct the tangent line.

There are plenty of examples of functions for which this is not possible

$$\text{eg. } y = |x| \text{ at } (0,0).$$

This happens because the function changes direction ‘sharply’ at (0,0).

Similarly, a two-dimensional function is *differentiable* at a point if we can construct the tangent plane.

A function may not be differentiable (i.e. not have a tangent plane) at a point because it is ‘sharp’ and not ‘smooth’ there.

The key difference from functions of one variable is that a function may have partial derivatives (i.e. it may be possible to differentiate with respect to x and y separately) but not be *differentiable,* (i.e. it may not have a tangent plane).

For example, consider the surface $z = \sqrt{x^2 y^2}$ near the origin (0,0,0).

From the 3-d plot it can be seen that, as we move away from the origin along the x-axis or the y-axis:

$$\frac{\partial z}{\partial x} = 0 \text{ and } \frac{\partial z}{\partial y} = 0$$

But no suitable tangent plane can be defined which approximates the function ‘successfully’ near the origin. For this and other functions Autograph is a useful tool for drawing diagrams to clarify understanding.

This is rather an artificial example and most functions that students will meet in this unit are differentiable. Nevertheless it is important to maintain a distinction between the concept of *differentiability* (the existence of an approximating tangent plane) and ‘derivability’ (the existence of the rate of change of a function with respect to a single variable (the partial derivative)).

For completion, if the partial derivatives of a function exist near a point *and are continuous there,* then the function is differentiable and has a tangent plane at this point.

Directional derivatives and the vector grad f

The ‘gradient’ on a surface depends on the direction taken (as with partial derivatives).

The *directional derivative* $\frac{\mathrm{d}z}{\mathrm{d}s}$ is the rate of change of z in the direction $\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix}$ where s is the length of $\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix}$. It can be thought of as the scalar change in z needed to stay on the tangent plane after a **unit** horizontal step **û** parallel to $\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix}$.

$\mathrm{d}z = \dfrac{\partial f}{\partial x}\mathrm{d}x + \dfrac{\partial f}{\partial y}\mathrm{d}y$ may be written as a scalar product: $\mathrm{d}z = \begin{bmatrix}\mathrm{d}x \\ \mathrm{d}y\end{bmatrix} . \begin{bmatrix}\dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y}\end{bmatrix}$ (4)

The vector $\begin{bmatrix}\dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y}\end{bmatrix}$ is called **grad** f and can be written as ∇f or ∇z

So the directional derivative, $\dfrac{\mathrm{d}z}{\mathrm{d}s}$ = **û.grad f**

The significance of **grad** f (≠ **0**) at a point is that it is a vector normal to the contour through that point It gives the magnitude and direction of the greatest rate of change of f. (If **grad** f = **0**, then both partial derivatives = 0 and the point is *stationary* (see below).)

If d*z* = 0 (i.e. we stay on the contour line) then **grad f** is *perpendicular* to $\begin{bmatrix}\mathrm{d}x \\ \mathrm{d}y\end{bmatrix}$

Also, the directional derivative is a maximum when **grad f** is *parallel* to $\begin{bmatrix}\mathrm{d}x \\ \mathrm{d}y\end{bmatrix}$

i.e. **grad f** is a vector in the direction of greatest slope.

So, in our example if we set out from the point (2,3,23) and move 1 unit in a direction of 40˚ with the positive *x*-axis, then d*x* = cos40˚ and d*y* = sin40˚.

The scalar change d*z* needed to stay on the tangent plane is given by:

$$\begin{aligned}\mathrm{d}z &= \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y \\ &= 4\cos 40^\circ + 12\sin 40^\circ \\ &\approx 6.27\end{aligned}$$

Stationary points

If z = f(*x,y*), a point at which the tangent plane is horizontal is called a *stationary point*.

At a stationary point, $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$ are **both** zero.

A stationary point could be a maximum or minimum or a *saddle* point but there are other possibilities (e.g. a ridge). The best way for students to investigate the nature of these is to consider values of the function near the point.

Functions of more than two variables

As said at the beginning, examples of functions of many variables are common. We will concentrate here on functions of **three** variables.

If $w = \mathrm{g}(x,y,z)$

then
$$\mathrm{d}w = \frac{\partial w}{\partial x}\mathrm{d}x + \frac{\partial w}{\partial y}\mathrm{d}y + \frac{\partial w}{\partial z}\mathrm{d}z \quad \text{or} \quad \mathrm{d}w = \frac{\partial g}{\partial x}\mathrm{d}x + \frac{\partial g}{\partial y}\mathrm{d}y + \frac{\partial g}{\partial z}\mathrm{d}z \qquad (6)$$

gives the equation of the 'tangent hyperplane'. This is difficult to visualize (!) but can be useful in three ways:

(1) If the function g is differentiable and $\delta x = \mathrm{d}x$, $\delta y = \mathrm{d}y$ and $\delta z = \mathrm{d}z$, then the change in w, δw is given by:

$$\delta w \approx \frac{\partial w}{\partial x}\delta x + \frac{\partial w}{\partial y}\delta y + \frac{\partial w}{\partial z}\delta z$$

This result may be used to find approximations to the value of the **function** (dw gives the change on the 'tangent hyperplane') near a known point.

(2) If we define **grad** g as the vector $\begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial y} \\ \frac{\partial g}{\partial z} \end{bmatrix}$ and $\hat{\mathbf{u}}$ as the *unit* vector $\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{bmatrix}$ then from (6)

$$\mathrm{d}w = \hat{\mathbf{u}}.\mathbf{grad}\ \mathrm{g}$$

gives the change in w on the 'tangent hyperplane' for a unit change in x,y and z.

As before, this is the *directional derivative* (the rate of change of w in a particular direction) and may be written $\frac{\mathrm{d}w}{\mathrm{d}s}$ with s defined as before**.**

(3) The (implicit) equation g*(x,y,z)* = k defines a surface in **three** dimensions. Such a surface is a 3-d equivalent of a contour of the form f*(x,y)* = k defined above.

Now, if $\mathbf{u} = \begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{bmatrix}$ is *tangential* to such a surface, then (6) may be written:

$$\mathbf{u}.\mathbf{grad}\ g = 0 \quad (\text{or } \mathbf{u}.\mathbf{grad}\ w = 0) \qquad *$$

(since dw = 0).

This means that the (2-d!) tangent *plane* can easily be found since **grad** g is the normal vector to all vectors in the tangent plane.

As in the A2 unit C4, once students are able to find the normal vector to a plane in 3-d space, most algebraic objects can be found easily (for example the normal line at a point has direction vector **grad** g).

The theory in (3) above may be used to find the tangent plane for functions of the form $z = f(x,y)$.

In the original example $z = x^2 + 2y^2 + 1$, we can define $w = g(x, y, z) = x^2 + 2y^2 - z$

and the equation may be written: $g(x, y, z) = -1$ or $w = 1$

Then the tangent plane at the 'point' (2,3,23,-1) can be found from **u**.**grad** g = 0

$$\begin{bmatrix} x-2 \\ y-3 \\ z-23 \end{bmatrix} . \begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial y} \\ \frac{\partial g}{\partial z} \end{bmatrix} = 0$$

$$\begin{bmatrix} x-2 \\ y-3 \\ z-23 \end{bmatrix} . \begin{bmatrix} 4 \\ 12 \\ -1 \end{bmatrix} = 0$$

$$4(x-2) + 12(y-3) - 1(z-23) = 0$$

$$4x + 12y - z = 21 \text{ as before.}$$

This result (*) (with its offshoots) is really the only one that is required in this chapter. Hence you can teach the contents of this chapter 'at pace' provided that your students learn and understand the significance of *. The technical requirements of partial differentiation and vector algebra should already be familiar to them from C3/C4.

Appendix 1: meaning of differentials

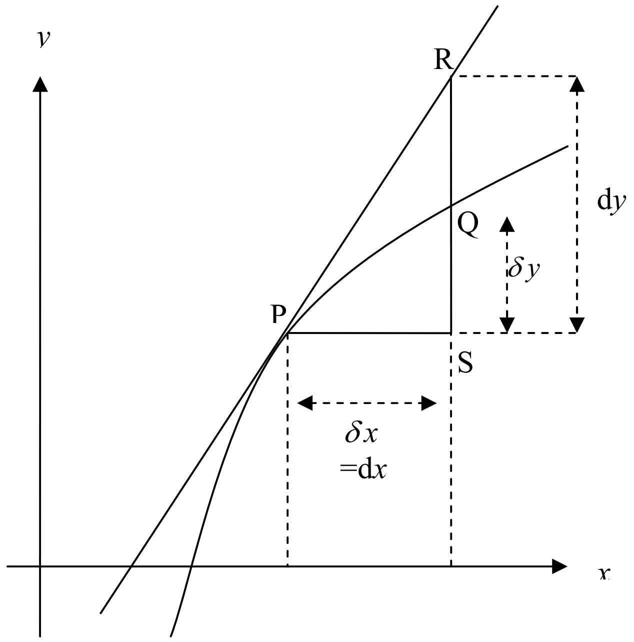
In the Leibnitz notation, $\frac{\mathrm{d}y}{\mathrm{d}x}$ is regarded as a single entity, not as a ratio of two separate quantities dy and dx. In fact, dy and dx can be given separate meanings in such a way that their ratio is equal to the derivative.

Let $y = \mathrm{f}(x)$ where f is a differentiable function of x.

The **differential d*x*** is defined as an independent variable, ie dx can take any real value.
The **differential d*y*** is then defined by the equation: $\mathrm{d}y = \mathrm{f}'(x)\mathrm{d}x$ or $\mathrm{d}y = \frac{dy}{dx}\mathrm{d}x$
(So dy is also a variable, but is dependent on x and dx.)
and the 'derivative' $\frac{\mathrm{d}y}{\mathrm{d}x}$ can be interpreted as a ratio of differentials.

In 2 dimensions it is easy to interpret the differentials geometrically.

Consider points $\mathrm{P}(x, y)$ and $\mathrm{Q}(x+\delta x, y+\delta y)$



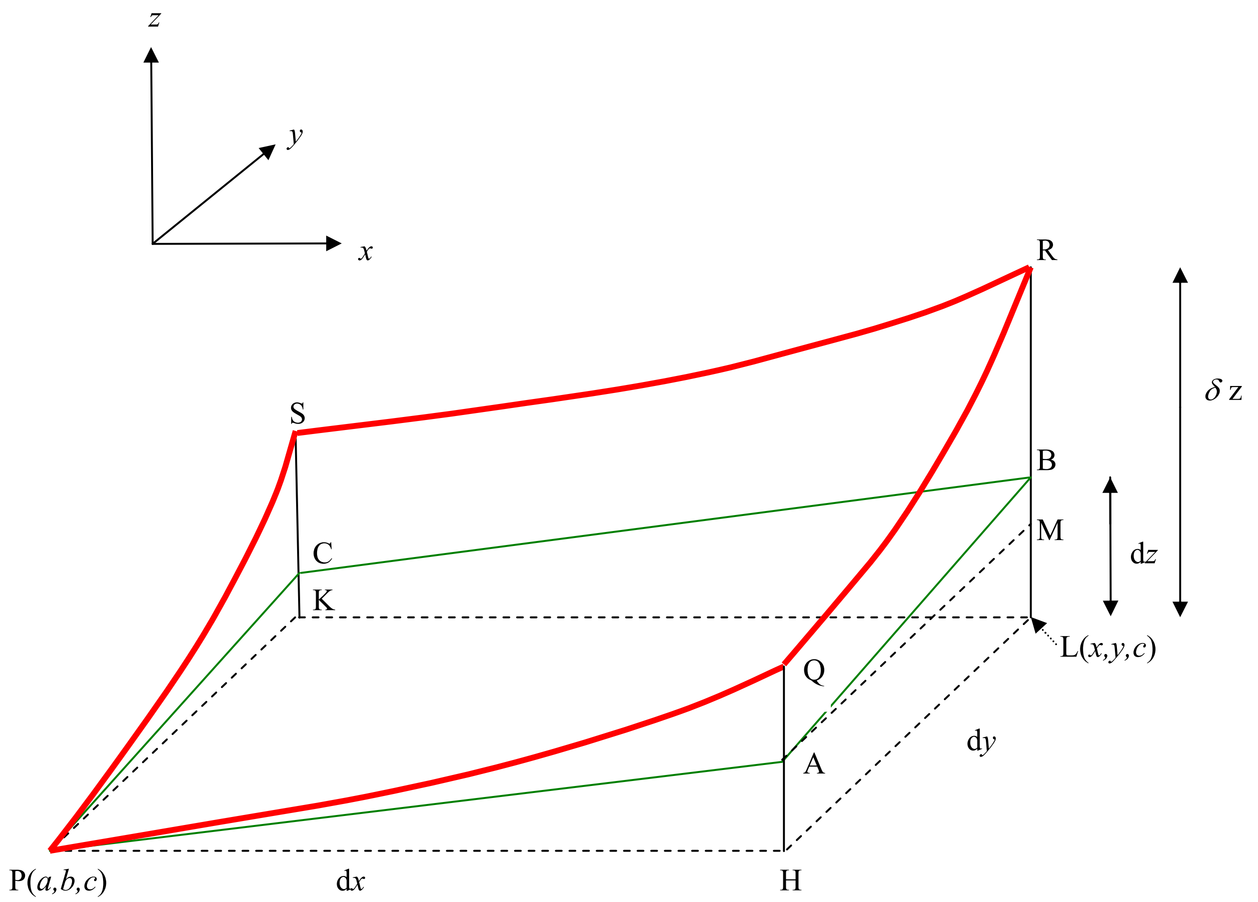
Gradient at P = slope of tangent PR ie. $\frac{dy}{dx} = \frac{\mathrm{RS}}{\mathrm{PS}}$

$$\mathrm{RS} = \frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{PS} = \frac{\mathrm{d}y}{\mathrm{d}x}\delta x = \frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x = \mathrm{d}y$$

The differential dy is the change in y to stay on the **tangent** line when x changes by dx.

δy is the change in y to stay on the **curve** when x changes by δx (= dx).

Appendix 2: the tangent plane



PQRS represents the surface $z = f(x,y)$.

PABC represents the tangent plane to the surface at P.

So $\mathrm{d}z = \mathrm{LM} + \mathrm{BM}$

$= \mathrm{AH} + \mathrm{CK}$

$$\mathrm{d}z = \frac{\partial z}{\partial x}\mathrm{d}x + \frac{\partial z}{\partial y}\mathrm{d}y$$ where the partial derivatives are evaluated at P

(which is equation (2) on page 2).

John Cooper: 28.6.10