Notes on Partial Differentiation

Matthew Smith

Abstract

People tend to get confused about how some variables are kept constant, and others varied, when calculating partial derivatives. The following is an attempt to explain what's going on, and also to collect a few results connected with partial differentiation.

1 Resumé of Total Differentiation

Given two variables y and x related by $y = f(x)$, we can define the derivative of y with respect to x as

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x} = \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x} = f'(x). \tag{1}$$

Essentially[1] this means that by making a small enough secant[2] line, we can make its slope as close as we please to the derivative of the curve at either point.

It also means that the first order approximation to δy for a small secant line is

$$\delta y \approx \frac{\mathrm{d}y}{\mathrm{d}x} \delta x$$

[1]Formally, the limit is defined by

$$\lim_{\delta \to 0} f(\delta) = L \iff \lim_{\delta \to 0+} f(\delta) = \lim_{\delta \to 0+} f(-\delta) = L$$

where

$$\lim_{\delta \to 0+} f(\delta) = L \iff \forall \epsilon > 0 \quad \exists \delta > 0 \quad \text{such that} \quad |f(\zeta) - L| < \epsilon \quad \text{whenever} \quad \zeta \leq \delta,$$

where $\forall$ means 'for all' and $\exists$ means 'there exists a'.

[2]A *secant* line cuts a curve at least twice, changing sides each time. A tangent is the limit of secants between narrowly separated points.

or, more precisely,

$$\delta y = \frac{\mathrm{d}y}{\mathrm{d}x}\delta x + O(\delta x^2). \tag{2}$$

From the definition follow the rules for differentiating polynomials and exponential and trigonometric functions, as well as the simple chain rule for functions of functions, and the product rule.

It can make sense to calculate the total derivative when several variables are involved. For example, suppose that z depends two variables x and y, such that $z = g(x, y)$, but that y itself depends on x, as $y = f(x)$. In this case we can certainly calculate $\frac{\mathrm{d}z}{\mathrm{d}x}$, as it is $\frac{\mathrm{d}}{\mathrm{d}x}g(x, f(x))$. We can also calculate $\frac{\mathrm{d}z}{\mathrm{d}y}$, which is $\frac{\mathrm{d}}{\mathrm{d}y}g(f^{-1}(y), y)$. (The former divided by the latter will, in accordance with the usual rules for manipulating derivatives, give $\frac{\mathrm{d}y}{\mathrm{d}x}$.) This works because, once we know x (or y) we can calculate z without having to know anything else. Thus it makes sense to refer to a change in z corresponding to a certain change in x (or y).

However, if we don't know that y and x are related, we can't calculate $\frac{\mathrm{d}z}{\mathrm{d}x}$, because there is no answer to the question "Given a variation δx in x, how much does y vary?". Furthermore, even when we know that y and x are related, it can be useful to think separately about the changes in z due to x directly, and those due to y (or, if you like, due to x through y). This is what partial differentiation is for.

2 The Chain Rule

Consider a function of several variables: $w = f(x, y, z)$. At this stage we don't know whether x, y and z are connected in any way. It might be that, once we know one, we can work out the others, as in section 1. It might be possible to specify all three of them independently. It might be that one needs to know two of them to work out the third. At this stage we don't know, and don't care, for we can define

$$\left.\frac{\partial w}{\partial x}\right|_{y,z} = \lim_{\delta x \to 0} \frac{f(x + \delta x, y, z) - f(x, y, z)}{\delta x} \tag{3}$$

as the partial derivative of w with respect to x, keeping y and z constant. As with total differentiation in equation (2), we have

$$f(x, y + \delta y, z) = f(x, y, z) + \left.\frac{\partial f(x, y, z)}{\partial y}\right|_{x,z} \delta y + O(\delta y^2). \tag{4}$$

This allows us to see what happens to w if we make small changes in one of x, y or z, without doing anything to the others.

Now what happens if we change all of x, y and z to new values? By applying equation (4) to $f(x+\delta x, y, z)$,

$$f(x+\delta x, y+\delta y, z) = f(x+\delta x, y, z) + \left.\frac{\partial f(x+\delta x, y, z)}{\partial y}\right|_{x,z} \delta y + O(\delta y^2),$$

and applying equation (4) again, but to x this time,

$$\begin{aligned} f(x+\delta x, y+\delta y, z) &= f(x,y,z) + \left.\frac{\partial f(x,y,z)}{\partial x}\right|_{y,z} \delta x + O(\delta x^2) + \\ &\quad + \left.\frac{\partial [f(x,y,z) + O(\delta x)]}{\partial y}\right|_{x,z} \delta y + O(\delta y^2). \end{aligned}$$

Applying the same procedure to z, we arrive at

$$\begin{aligned} \delta w &= f(x+\delta x, y+\delta y, z+\delta z) - f(x,y,z) \\ &= \left.\frac{\partial w}{\partial x}\right|_{y,z} \delta x + \left.\frac{\partial w}{\partial y}\right|_{x,z} \delta y + \left.\frac{\partial w}{\partial z}\right|_{x,y} \delta z + O(\delta^2). \end{aligned}$$

This is the change in w corresponding to small changes in all of x, y and z. Dividing by δ (anything), and taking the limit as it tends to zero, we get a relation between total derivatives. These can all be summarized by the *chain rule*, which is the central result relating partial and total derivatives:

$$\mathrm{d}w = \left.\frac{\partial w}{\partial x}\right|_{y,z} \mathrm{d}x + \left.\frac{\partial w}{\partial y}\right|_{x,z} \mathrm{d}y + \left.\frac{\partial w}{\partial z}\right|_{y,x} \mathrm{d}z. \tag{5}$$

The chain rule can easily be extended to more variables or reduced to fewer. It isn't possible to summarize everything worth knowing about partial derivatives in one equation, but, if it were, this would be that equation.

The chain rule can be used both ways round. If you know the partial derivatives, it gives you a relation between $\mathrm{d}w$ and $\mathrm{d}x$, $\mathrm{d}y$ etc. On the other hand, if you know that $\mathrm{d}\theta = P\mathrm{d}u + Q\mathrm{d}v$, you can infer that $P = \left.\frac{\partial \theta}{\partial u}\right|_v$ and $Q = \left.\frac{\partial \theta}{\partial v}\right|_u$, which is often useful.

3 Degrees of Freedom

Let us return to the question of how x, y and z might be related to each other. Any relation of this sort is a restriction on the values of x, y, and z to a subset of (x, y, z)-space. Thus if x and y can be specified independently, but z is fixed by $z = f(x, y)$, the possible values of x, y and z are restricted to a two-dimensional surface in (x, y, z)-space. This could just as well[3] be expressed as $y = g(x, z)$ or $x = h(x, y)$, or, more generally, as $F(x, y, z) = 0$, where in this case $F(x, y, z) = f(x, y) - z$. There are then two degrees of freedom. If only one variable can be specified independently, because it determines the others, then we have a curve in (x, y, z)-space, two constraint equations, and only one degree of freedom.

Given some $w(x, y, z)$ which could be defined everywhere in (x, y, z) space, but only realized on some surface $z = f(x, y)$, how can $\mathrm{d}w$ be expressed in terms of $\mathrm{d}x$ and $\mathrm{d}y$? Using equation (5) for $\mathrm{d}w$ in terms of all three differentials, but substituting in for $\mathrm{d}z$ in terms of $\mathrm{d}x$ and $\mathrm{d}y$ using equation (5) for f, we get

$$\mathrm{d}w = \left.\frac{\partial w}{\partial x}\right|_{y,z} \mathrm{d}x + \left.\frac{\partial w}{\partial y}\right|_{x,z} \mathrm{d}y + \left.\frac{\partial w}{\partial z}\right|_{x,y} \left(\left.\frac{\partial z}{\partial x}\right|_{y} \mathrm{d}x + \left.\frac{\partial z}{\partial y}\right|_{x} \mathrm{d}y \right).$$

This expression works out $\mathrm{d}w$ for a given $\mathrm{d}x$ and $\mathrm{d}y$, taking account of the fact that we have to stay on the constraint surface by adjusting $\mathrm{d}z$ to keep us there. As the condition 'staying on the constraint surface' ought to go without saying, it makes sense to define

$$\mathrm{d}w = \left.\frac{\partial w}{\partial x}\right|_{y} \mathrm{d}x + \left.\frac{\partial w}{\partial y}\right|_{x} \mathrm{d}y,$$

where it can be seen that

$$\left.\frac{\partial w}{\partial x}\right|_{y} = \left.\frac{\partial w}{\partial x}\right|_{y,z} + \left.\frac{\partial w}{\partial z}\right|_{x,y} \left.\frac{\partial z}{\partial x}\right|_{y},$$

Note that we could define $\left.\frac{\partial w}{\partial x}\right|_{z}$ in the same way, but that it would be different. This is the same issue which motivated the definition in equation (3) in the first place: the derivative of w with respect to x depends on which direction we take in (x, y, z) space. The solution there, to move parallel to the x-axis and thus form $\left.\frac{\partial w}{\partial x}\right|_{y,z}$, is unphysical if we're required to keep to the constraint surface. There are different directions for movement within the surface, though: we can decide whether to keep y or z constant.

4 Change of Basis

Suppose that we have a system of related variables with two degrees of freedom, so that we can work out all of them once we know the value of any two.

[3]. . . provided that the surface doesn't bend back on itself, which would make some of these functions multi-valued

The thermodynamics of some simple systems, like a fixed amount of gas, can be described in this way. It is possible to use the chain rule, equation (5), in order to change from one 'basis' of variables to another. For example, if $\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V$, where, of course, $T = \left.\frac{\partial U}{\partial T}\right|_S$ and $p = -\left.\frac{\partial U}{\partial V}\right|_S$, then we can express $\mathrm{d}U$ in terms of $\mathrm{d}V$ and $\mathrm{d}T$ (say) by substituting for $\mathrm{d}S$ in terms of $\mathrm{d}V$ and $\mathrm{d}T$ using the chain rule. This works out to be

$$\mathrm{d}U = T\left(\left.\frac{\partial S}{\partial V}\right|_T \mathrm{d}V + \left.\frac{\partial S}{\partial T}\right|_V \mathrm{d}T\right) - p\mathrm{d}V$$

so that

$$\mathrm{d}U = \left(T\left.\frac{\partial S}{\partial V}\right|_T - p\right)\mathrm{d}V + T\left.\frac{\partial S}{\partial T}\right|_V \mathrm{d}T.$$

This sort of manipulation can generate all kinds of useful identities. Applying it to a simple set of three variables, x, y, and z, again with two degrees of freedom, we get from

$$\mathrm{d}x = \left.\frac{\partial x}{\partial y}\right|_z \mathrm{d}y + \left.\frac{\partial x}{\partial z}\right|_y \mathrm{d}x$$

to

$$\left(1 - \left.\frac{\partial x}{\partial y}\right|_z \left.\frac{\partial y}{\partial x}\right|_z\right)\mathrm{d}x = \left(\left.\frac{\partial x}{\partial y}\right|_z \left.\frac{\partial y}{\partial z}\right|_x + \left.\frac{\partial x}{\partial z}\right|_y\right)\mathrm{d}z.$$

Since we can move about in (x, z)-space however we like, providing that we adjust y to compensate, we can move in such a way that $\frac{\mathrm{d}x}{\mathrm{d}z} = 0$, or, alternatively, that $\frac{\mathrm{d}z}{\mathrm{d}x}$=0. These are just changes which keep x and z constant, respectively. The equation above must remain true in both cases, and this can only happen if the coefficients of $\mathrm{d}x$ and $\mathrm{d}z$ are both zero. Using the $\mathrm{d}x$ coefficient,

$$\left.\frac{\partial y}{\partial x}\right|_z = \left(\left.\frac{\partial x}{\partial y}\right|_z\right)^{-1}, \tag{6}$$

as might have been expected (though it's always good to make sure that intuition can be confirmed). Using equation (6) on the $\mathrm{d}z$ coefficient,

$$\left.\frac{\partial x}{\partial y}\right|_z \left.\frac{\partial y}{\partial z}\right|_x \left.\frac{\partial z}{\partial x}\right|_y = -1,$$

which is perhaps less obvious. This last relation is known at the *reciprocity law.*

5 Total Differentials and the Legendre Transformation

Any differential of a variable is known as a total differential. For example, if $f = xy^2$, then $\mathrm{d}f = y^2\mathrm{d}x + 2xy\mathrm{d}y$ is a total differential. It's obvious that it's a total differential from the left-hand side, but, if you just had an expression like the right-hand side, how could you decide whether or not it was a total differential? In other words, faced with an expression like $P\mathrm{d}x + Q\mathrm{d}y$, how do you decide whether there is a function f such that $\mathrm{d}f = P\mathrm{d}x + Q\mathrm{d}y$? One way is to try to construct f by integrating P with respect to x, thus determining f up to an arbitrary function of y, and by integrating Q with respect to y, and seeing if these two methods can be made to agree. However, integration can be difficult or even impossible, so it would be nicer to have a test based on differentiation, so that there's no need to construct f itself. The chain rule tells us that, if the expression is a total differential, $P = \left.\frac{\partial f}{\partial x}\right|_y$ and $Q = \left.\frac{\partial f}{\partial y}\right|_x$. Thus $\left.\frac{\partial P}{\partial y}\right|_x$ would be $\frac{\partial^2 f}{\partial x \partial y}$, and $\left.\frac{\partial Q}{\partial x}\right|_y$ would be $\frac{\partial^2 f}{\partial y \partial x}$. Now by using the definition in equation (3) on itself, it can be shown that these are equal: the order of differentiation doesn't matter.[4] Thus

$$\left.\frac{\partial P}{\partial y}\right|_x = \left.\frac{\partial Q}{\partial x}\right|_y \tag{7}$$

whenever $P\mathrm{d}x + Q\mathrm{d}y$ is a total differential.

In fact this is a sufficient as well as a necessary condition: it turns out that $P\mathrm{d}x + Q\mathrm{d}y$ is a total differential whenever $\left.\frac{\partial P}{\partial y}\right|_x = \left.\frac{\partial Q}{\partial x}\right|_y$. This will be useful later on.

Some important identities in thermodynamics can be derived in this way. For example, given (as before) that $\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V$,[5] we can infer that $\left.\frac{\partial T}{\partial V}\right|_S = \left.\frac{\partial p}{\partial S}\right|_V$, which is one of the *Maxwell relations*. More such relations can be generated by means of the *Legendre transformation*: instead of considering

[4] In fact, they both equal

$$\lim_{\delta x \to 0} \lim_{\delta y \to 0} \frac{f(x+\delta x, y+\delta y) - f(x+\delta x, y) - f(x, y+\delta y) + f(x,y)}{\delta x \delta y},$$

provided that we are allowed to interchange the order of the limits, which is possible provided that f is a reasonably 'well-behaved' function.

[5] This is one of the most fundamental relations in thermodynamics, which is true for any system that can only do work mechanically, such as a gas in a cylinder. U is the internal energy, which is the total amount of energy in the gas, T is its temperature, p is the pressure which it exerts on the cylinder, and V is the volume which it takes up. S is the *entropy*, a quantity the meaning of which I despair of explaining in a footnote.

the differential of the function U, consider instead the differential of the newly-invented function $F = U - TS$. It's easy to show that $\mathrm{d}F = -S\mathrm{d}T - p\mathrm{d}V$, and since $\mathrm{d}F$ is a total differential, it follows that $\left.\frac{\partial S}{\partial V}\right|_T = \left.\frac{\partial p}{\partial T}\right|_V$. This is another of the Maxwell relations. In fact F has an important meaning in thermodynamics — it's called the Helmholtz Free Energy — but here it's just a function invented to have the right sort of differential.

6 Conservative Vector Fields

A vector field $\mathbf{F}$ is called *conservative* if there exists a scalar field ϕ such that $\mathbf{F} = \nabla\phi$. There's an important connexion with total differentials here: as

$$\mathbf{F} = \begin{pmatrix} F_x \\ F_y \end{pmatrix} = \begin{pmatrix} \partial\phi/\partial x \\ \partial\phi/\partial y \end{pmatrix},$$

it follows that

$$\mathbf{F} \cdot \mathrm{d}\mathbf{x} = \begin{pmatrix} \partial\phi/\partial x \\ \partial\phi/\partial y \end{pmatrix} \cdot \begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \end{pmatrix} = \frac{\partial\phi}{\partial x}\mathrm{d}x + \frac{\partial\phi}{\partial y}\mathrm{d}y = \mathrm{d}\phi.$$

In other words, if $\mathbf{F}$ is conservative, $\mathbf{F} \cdot \mathrm{d}\mathbf{x}$ is a total differential.

If we integrate $\mathbf{F}$ along some path from $\mathbf{x} = \mathbf{a}$ to $\mathbf{x} = \mathbf{b}$ in (x, y)-space, we get

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{F} \cdot \mathrm{d}\mathbf{x} = \int_{\mathbf{a}}^{\mathbf{b}} \mathrm{d}\phi = \phi(\mathbf{b}) - \phi(\mathbf{a}).$$

This shows that the integral of a conservative vector field along a path depends only on the start and end points on the path, not on the route taken between them.[6] Similarly it follows that, around a *closed* path Γ_C,

$$\oint_{\Gamma_C} \mathbf{F} \cdot \mathrm{d}\mathbf{x} = 0.$$

If $\mathbf{F}$ represents a force field, then this integral is the work done by the force on an object moving along the path. Since the integral around a closed loop is always zero, no work is done overall on an object moving around a closed loop. This means that the positive work done by the force on the object must have exactly cancelled with the work done by the object against the force. No energy can be gained or lost by going round a closed loop, which

[6] Remember that integration is the inverse of differentiation, and is a limit of adding up many little pieces. If we add up all of the little pieces of change in ϕ, we must get the total change in ϕ whichever way we do it.

is why the field is called conservative. The scalar $-\phi$ is the potential energy of the particle in the field.

Using equation (7), another condition on a conservative vector field is that $\frac{\partial F_x}{\partial y} = \frac{\partial F_y}{\partial x}$, i.e. that the z-component of the curl of $\mathbf{F}$ is zero. In fact, this all generalizes to three dimensions, and the following conditions for conservative vector fields are all equivalent:

$$\exists \phi \quad \text{such that} \quad \nabla\phi = \mathbf{F} \quad \Longleftrightarrow \quad \oint_{\Gamma_C} \mathbf{F}\cdot \mathrm{d}\mathbf{x} = 0 \quad \forall\, \Gamma_C$$

$$\Longleftrightarrow \quad \int_{\mathbf{a}}^{\mathbf{b}} \mathbf{F}\cdot\mathrm{d}\mathbf{x} \quad \text{is independent of path} \quad \Longleftrightarrow \quad \nabla\times\mathbf{F} = \mathbf{0}.$$

The second and fourth conditions are equivalent from Stokes' theorem, since the integral of $\mathbf{F}$ around a closed path must always be zero if $\nabla\times\mathbf{F}$ integrates to zero over any surface.