N. Non-independent Variables

1. Partial differentiation with non-independent variables.

Up to now in calculating partial derivatives of functions like $w = f(x, y)$ or $w = f(x, y, z)$, we have assumed the variables x, y (or x, y, z) were independent. However in real-world applications this is frequently not so. Computing partial derivatives then becomes confusing, but it is better to face these complications now while you are still in a calculus course, than wait to be hit with them at the same time that you are struggling to cope with the thermodynamics or economics or whatever else is involved.

For example, in thermodynamics, three variables that are associated with a contained gas are its

$$p = \text{pressure}, \qquad v = \text{volume}, \qquad T = \text{temperature},$$

and you can express other thermodynamic variables like the internal energy U and entropy S in terms of p, v, and T.

However, p, v, and T are not independent variables. If the gas is a so-called "ideal gas", they are related by the equation

$$(1) \qquad\qquad pv = nRT \qquad\qquad (n, R \ \text{constants}).$$

To see what complications this produces, let's consider first a purely mathematical example.

Example 1. Let $w = x^2 + y^2 + z^2$, where $z = x^2 + y^2$. Calculate $\dfrac{\partial w}{\partial x}$.

Discussion.

(a) If we think of x and y as the independent variables, then we can calculate $\dfrac{\partial w}{\partial x}$ by two different methods:

(i) using $z = x^2 + y^2$ to get rid of z, we get

$$
\begin{aligned}
w &= x^2 + y^2 + (x^2 + y^2)^2 \\
&= x^2 + y^2 + x^4 + 2x^2 y^2 + y^4; \\
\frac{\partial w}{\partial x} &= 2x + 4x^3 + 4xy^2
\end{aligned}
$$

(ii) or by using the chain rule, remembering z is a function of x and y,

$$
\begin{aligned}
w &= x^2 + y^2 + z^2 \\
\frac{\partial w}{\partial x} &= 2x + 2z \frac{\partial z}{\partial x} = 2x + 2z \cdot 2x \\
&= 2x + 2(x^2 + y^2) \cdot 2x,
\end{aligned}
$$

so the two methods agree.

(b) On the other hand, if we think of x and z as the independent variables, using say method (i) above, we get rid of y by using the relation $y^2 = z - x^2$, and get

$$
\begin{aligned}
w &= x^2 + y^2 + z^2 = x^2 + (z - x^2) + z^2 \\
&= z + z^2; \\
\frac{\partial w}{\partial x} &= 0.
\end{aligned}
$$

0

These answers are genuinely different — we cannot convert one into the other by using the relation $z = x^2 + y^2$. Will the right $\partial w / \partial x$ please stand up?

The answer is, that there *is* no one right answer, because the problem was not well-stated. *When the variables are not independent, an expression like $\partial w / \partial x$ has no definite meaning.*

To see why this is so, we interpret the above example geometrically. Saying that x, y, z satisfy the relation $z = x^2 + y^2$ means that the point (x, y, z) lies on the paraboloid surface formed by rotating $z = y^2$ about the z-axis. The function

$$w \;=\; x^2 + y^2 + z^2$$

measures the square of the distance from the origin. To be definite, let's suppose we are at the starting point $P = P_0 : (1, 0, 1)$ indicated, and we want to calculate $\partial w / \partial x$ at this point.
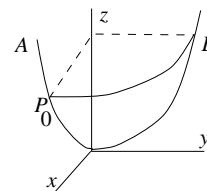
Case (a) If we take x and y to be the independent variables, then to find $\partial w / \partial x$, we *hold y fixed and let x vary.* So P moves in the xz-plane towards A, along the path shown.

As P moves along this path, evidently w, the square of its distance from the origin, is steadily increasing: $\dfrac{\partial w}{\partial x} > 0$ and in fact the calculations for **(a)** on the previous page show that $\dfrac{\partial w}{\partial x} = 6$.

Case (b) If we take x and z to be the independent variables, then to find $\partial w / \partial x$, we *hold z fixed and let x vary.* Now P moves in the plane $z = 1$, along the circular path towards B.

As P moves on this path, the square of its distance from the origin is *not changing*, and therefore $\dfrac{\partial w}{\partial x} = 0$,   as we calculated in **(b)** before.

To sum up, the value of $\partial w / \partial x$ depends on which variables we take to be independent, because we are measuring different rates of change, as P moves along different paths.

There is only one way out of our difficulty. When we ask for $\partial w / \partial x$, we must at the same time specify which variables are to be taken as the independent ones. This is done by using the following notation:

Case **(a)**: x, y are the independent variables: $\left(\dfrac{\partial w}{\partial x}\right)_y$

Case **(b)**: x, z are the independent variables: $\left(\dfrac{\partial w}{\partial x}\right)_z$

These are read, "the partial of w with respect to x, with y (resp. z) held constant".

Note how in each case the two lower letters give you the two independent variables. If we had more variables, we would use a similar notation. For instance if

(2) $w \;=\; f(x, y, z, t),$ where $xy = zt,$

then only three of the variables x, y, z, t can be independent; the fourth is then determined

by the equation on the right of (2). Thus we would write expressions like

$$\left(\frac{\partial w}{\partial x}\right)_{y,t} \qquad \text{"partial of } w \text{ with respect to } x; \ y \text{ and } t \text{ held constant";}$$

$$\left(\frac{\partial w}{\partial y}\right)_{x,z} \qquad \text{"partial of } w \text{ with respect to } y; \ x \text{ and } z \text{ held constant";}$$

in the first, x, y, t are the independent variables; in the second, x, y, z are independent.

2. Differentials vs. Chain Rule

An alternative way of calculating partial derivatives uses total differentials. We illustrate with an example, doing it first with the chain rule, then repeating it using differentials. By definition, the differential of a function of several variables, such as $w = f(x, y, z)$ is

$$(3) \qquad\qquad dw = f_x dx + f_y dy + f_z dz,$$

where the three partial derivatives f_x, f_y, f_z are the *formal* partial derivatives, i.e., the derivatives calculated as if x, y, z were independent.

Example 2. Find $\left(\dfrac{\partial w}{\partial y}\right)_{x,t}$, where $w = x^3 y - z^2 t$ and $xy = zt$.

Solution 1. Using the chain rule and the two equations in the problem, we have

$$\left(\frac{\partial w}{\partial y}\right)_{x,t} = x^3 - 2zt \left(\frac{\partial z}{\partial y}\right)_{x,t} = x^3 - 2zt\frac{x}{t} = x^3 - 2zx.$$

Solution 2. We take the differentials of both sides of the two equations in the problem:

$$(4) \qquad dw = 3x^2 y\, dx + x^3 dy - 2zt\, dz - z^2 dt, \qquad\qquad y\, dx + x\, dy = z\, dt + t\, dz.$$

Since the problem indicates that x, y, t are the independent variables, we eliminate dz from the equations in (4) by multiplying the second equation by $2z$, adding it to the first, then grouping the terms, which gives

$$dw = (3x^2 y - 2zy)\, dx + (x^3 - 2zx)\, dy + z^2 dt$$

Comparing this with (3) — after replacing z by t in (3) — we see that

$$\left(\frac{\partial w}{\partial x}\right)_{y,t} = 3x^2 y - 2zy, \qquad \left(\frac{\partial w}{\partial y}\right)_{x,t} = x^3 - 2zx, \qquad \left(\frac{\partial w}{\partial t}\right)_{x,y} = z^2 \ .$$

(The actual partial derivatives are the same as the formal partial derivatives w_x, w_y, w_t because x, y, t are independent variables.)

Notice that the differential method here takes a bit more calculation, but gives us three derivatives, not just one; this is fine if you want all three, but a little wasteful if you don't. The main thing to keep in mind for the method is that differentials are treated like vectors, with the dx, dy, $dz, \ldots$ playing the role of $\mathbf{i}, \mathbf{j}, \mathbf{k}, \ldots$. That is:

D1. *Differentials can be added, subtracted, and multiplied by scalar functions;*

D2. *If the variables $x, y, \ldots$ are independent, two differentials are equal if and only if their corresponding coefficients are equal:*

$$(5) \qquad A\,dx + B\,dy + \ldots = A_1\,dx + B_1\,dy + \ldots \qquad \Leftrightarrow \qquad A = A_1, \ B = B_1, \ldots \ ;$$

D3. *One differential can be substituted into another.*

Remarks.

1. In Example 2, Solution 2, we used the operations in **D1** to do the calculations. We used **D2** in the last step, taking advantage of the fact that the x, y, t were independent.

We could have done the calculations using **D3** instead, by solving the second equation in (4) for dz and substituting it into the first equation. **D3** is a consequence of the chain rule. Illustrations of its use will be given in the next section.

2. The main advantage of calculating with differentials is that one need not take into account whether the variables are dependent or not, or which variables depend on which others; the method does this automatically for you. Examples will illustrate.

3. If the variables are not independent, **D2** is emphatically *not* true; the second equation in (4) gives a counterexample.

Note also that in **D1**, there is no attempt to include a "multiplication" or "division" of differentials to the list of operations. If u and v are functions of several variables, then their "product" $du\,dv$ makes no sense as a differential, nor does their "quotient" du/dv, which despite appearances is not in general related to any derivative, or function, or even defined. (There is no elementary analogue of the dot and cross product of vectors, though in advanced differential geometry courses a certain type of product for differentials is defined and used for multiple integration.)

Example 3. Let $w = x^2 - yz + t^2$, where x, y, z, t satisfy the two equations
$$z^2 = x + y^2 \qquad \text{and} \qquad xy = zt.$$
Using these equations, we can express first z and then t in terms of x and y; this means that w can also be expressed in terms of x and y. Without actually calculating $w(x, y)$ explicitly, find its gradient vector $\nabla w(x, y)$.

Solution. Since we need both partial derivatives $(\partial w/\partial x)_y$ and $(\partial w/\partial y)_x$, it makes sense to use the differential method. Taking the differential of w and of the two equations connecting the variables gives us

$$(6) \quad dw = 2x\,dx - z\,dy - y\,dz + 2t\,dt, \qquad x\,dy + y\,dx = z\,dt + t\,dz, \qquad 2z\,dz = dx + 2y\,dy.$$

We want x and y to be the independent variables; using the operations in **D1**, first eliminate dt by solving for it in the second equation, and substituting for it into the first equation; then eliminate dz by solving for it in the last equation and substituting into the first equation; the result is

$$(7) \qquad dw = \left(2x - \frac{y}{2z} + \frac{2ty}{z} - \frac{t^2}{z^2}\right) dx + \left(-z - \frac{y^2}{z} + \frac{2xt}{z} - \frac{2t^2 y}{z^2}\right) dy.$$

Since x and y are independent, comparing the two expressions for dw in (7) and (3) (using x and y), and then using **D2**, shows that the two coefficients in (7) are respectively the two partial derivatives w_x and w_y, i.e., the two components of the gradient ∇w.

Example 4. Suppose the variables x, y, z satisfy an equation $g(x, y, z) = 0$. Assume the point $P : (1, 1, 1)$ lies on the surface $g = 0$ and that $(\nabla g)_P = \langle -1, 1, 2 \rangle$.

Let $f(x, y, z)$ be another function, and assume that $(\nabla f)_P = \langle 1, 2, 1 \rangle$.

Find the gradient of the function $w = f(x, y, z(x, y))$ of the two independent variables x and y, at the point $x = 1, \ y = 1$.

Solution. Using differentials, we have, by (3) and our hypotheses,

$$(dw)_P = dx + 2dy + dz; \qquad (dg)_P = -dx + dy + 2dz = 0, \quad \text{since } dg = 0 \text{ for all } x, y, z;$$

eliminating dz by solving the second equation for it and substituting into the first, or by dividing the second equation by 2 and substracting it from the first, we get

$$(dw)_P = \tfrac{3}{2}dx + \tfrac{3}{2}dy; \qquad\qquad (\nabla w)_P = \tfrac{3}{2}\,\mathbf{i} + \tfrac{3}{2}\,\mathbf{j}.$$

3. Abstract partial differentiation; rules relating partial derivatives

Often in applications, the function w is not given explicitly, nor are the equations connecting the variables. Thus you need to be able to work with functions and equations just given abstractly. The previous ideas work perfectly well, as we will illustrate. However, we will need (as in section 2) to distinguish between

formal partial derivatives, written here $f_x, \ f_y, \ldots$ (calculated as if all the variables were independent), and

actual partial derivatives, written $\partial f / \partial x, \ldots$, which take account of any relations between the variables.

Example 5. If $f(x, y, z) = xy^2z^4$, where $z = 2x + 3y$, the three formal derivatives are

$$f_x = y^2 z^4, \qquad\qquad f_y = 2xyz^4, \qquad\qquad f_z = 4xy^2 z^3,$$

while three of the many possible actual partial derivatives are (we use the chain rule)

$$\left(\frac{\partial f}{\partial x}\right)_y = f_x + f_z \left(\frac{\partial z}{\partial x}\right)_y = y^2 z^4 + 8xy^2 z^3;$$

$$\left(\frac{\partial f}{\partial y}\right)_x = f_y + f_z \left(\frac{\partial z}{\partial y}\right)_x = 2xyz^4 + 12xy^2 z^3;$$

$$\left(\frac{\partial f}{\partial z}\right)_x = f_y \left(\frac{\partial y}{\partial z}\right)_x + f_z = \tfrac{2}{3}xyz^4 + 4xy^2 z^3.$$

Rules connecting partial derivatives. These rules are widely used in the applications, especially in thermodynamics. Here we will use them as an excuse for further practice with the chain rule and differentials.

With an eye to thermodynamics, we assume a set of variables $t, u, v, w, x, y, z, \ldots$ connected by several equations in such a way that
- any *two* are independent;
- any *three* are connected by an equation.

Thus, one can choose any two of them to be the independent variables, and then each of the other variables can be expressed in terms of these two.

We give each rule in two forms—the second form is the one ordinarily used, while the first is easier to remember. (The first two rules are fairly simple in either form.)

(8a,b)
$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z = 1 \qquad\qquad \left(\frac{\partial x}{\partial y}\right)_z = \frac{1}{(\partial y/\partial x)_z} \quad \textbf{reciprocal rule}$$

(9a,b)
$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial t}\right)_z = \left(\frac{\partial x}{\partial t}\right)_z \qquad \left(\frac{\partial x}{\partial y}\right)_z = \frac{(\partial x/\partial t)_z}{(\partial y/\partial t)_z}, \quad \textbf{chain rule}$$

(10a,b)
$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1 \qquad \left(\frac{\partial x}{\partial y}\right)_z = -\frac{(\partial x/\partial z)_y}{(\partial y/\partial z)_x}, \quad \textbf{cyclic rule}$$

Note how the successive factors in the cyclic rule are formed: the variables are used in the successive orders x, y, z; y, z, x; z, x, y; one says they are permuted cyclically, and this explains the name.

Proof of the rules. The first two rules are simple: since z is being held fixed throughout, each variable becomes a function of just one other variable, and (9) is just the one-variable chain rule. Then (8) is just the special case of (9) where $x = t$.

The cyclic rule is less obvious — on the right side it looks almost like the chain rule, but different variables are being held constant in each of the differentiations, and this changes it entirely. To prove it, we suppose $f(x, y, z) = 0$ is the equation satisfied by x, y, z; taking y and z as the independent variables and differentiating $f(x, y, z) = 0$ with respect to y gives:

(11)
$$f_x \left(\frac{\partial x}{\partial y}\right)_z + f_y = 0; \qquad \text{therefore} \quad \left(\frac{\partial x}{\partial y}\right)_z = -\frac{f_y}{f_x} .$$

Permuting the variables in (11) and multiplying the resulting three equations gives (10a):

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -\frac{f_y}{f_x} \cdot -\frac{f_z}{f_y} \cdot -\frac{f_x}{f_z} = -1.$$

Example 6. Suppose $w = w(x, r)$, with $r = r(x, \theta)$. Give an expression for $\left(\dfrac{\partial w}{\partial r}\right)_\theta$ in terms of formal partial derivatives of w and r.

Solution. Evidently the independent variables are to be r and θ, since these are the ones that occur in the lower part of the partial derivative, with x dependent on them. Since θ is viewed as a constant, the chain rule gives

$$\left(\frac{\partial w}{\partial r}\right)_\theta = w_x \left(\frac{\partial x}{\partial r}\right)_\theta + w_r \;;$$
$$\left(\frac{\partial x}{\partial r}\right)_\theta = \frac{1}{(\partial r/\partial x)_\theta},$$

by the reciprocal rule (8). and therefore finally,

$$\left(\frac{\partial w}{\partial r}\right)_\theta = \frac{w_x}{r_x} + w_r .$$

4.   Changing the independent variables.* For those of you who will study thermodynamics, a major use of the rules of the preceding section is to change physical laws expressed in terms of one pair of independent variables to another pair which is better adapted to the particular problem at hand.

In thermodynamics, some of the variables associated with a confined gas are p (pressure), V (volume), T (temperature), U (internal energy), S (entropy), and H (enthalpy). Any two are independent, and their values then determine all the others.

To avoid confusion, it is better to state our general problem in terms of a neutral list of variables — we will use u, v, w, x, y. Then we can state the problem this way: a partial derivative $\left(\dfrac{\partial A}{\partial B}\right)_C$ is given, where the A, B, C are three of these variables, and we want to use x and y as the new independent variables; i.e., we want to express $\left(\dfrac{\partial A}{\partial B}\right)_C$ in terms of partial derivatives that look like $\left(\dfrac{\partial *}{\partial x}\right)_y$ and $\left(\dfrac{\partial *}{\partial y}\right)_x$, where $*$ stands for any of the variables.

It looks like there will be many cases, but outside of the trivial ones, the most commonly occurring ones are all handled by the rules of the previous section.

The trivial cases are when two of A, B, C are equal:

$$(12) \qquad \left(\frac{\partial A}{\partial B}\right)_C = \begin{cases} 1, & A = B; \\ 0, & A = C; \\ \text{undefined}, & B = C. \end{cases}$$

Two more "trivial" cases are when B and C are x and y, in either order, since then the partial derivative is already in the desired form.

The rest of the cases are non-trivial, but are covered by the rules. Remembering that x and y are to be the new variables, the commonly occurring cases are these two:

$$(13) \qquad \left(\frac{\partial A}{\partial B}\right)_x = \frac{(\partial A/\partial y)_x}{(\partial B/\partial y)_x}, \qquad \text{(chain rule (9))}$$

$$(14) \qquad \left(\frac{\partial x}{\partial y}\right)_u = -\frac{(\partial x/\partial u)_y}{(\partial y/\partial u)_x} = -\frac{(\partial u/\partial y)_x}{(\partial u/\partial x)_y}, \qquad \text{by (10b) and (8b)}$$

In the above, x and y can be interchanged; A, B, C stand for any variables; u, v, w are any variables other than x or y. The reciprocal rule can be used as a preliminary step to put a given partial derivative into one of the above forms.

Example 7. One of the laws of thermodynamics is expressed by the equation

$$\left(\frac{\partial U}{\partial p}\right)_T + T\left(\frac{\partial V}{\partial T}\right)_p + p\left(\frac{\partial V}{\partial p}\right)_T = 0.$$

What is the equation for this law when V and T are the independent variables?

Solution. Looking at each derivative in turn, the first has the form (13) and needs the chain rule; the second has the form (14) and needs the cyclic rule; the last needs only the reciprocal rule. Using these, the equation is transformed into

$$\frac{\partial U/\partial V}{\partial p/\partial V} - T \frac{\partial p/\partial T}{\partial p/\partial V} + \frac{p}{\partial p/\partial V} = 0.$$

The subscripts are unnecessary, if it is known that T and V are the independent variables; however there is no harm in including them and removing the common denominator, which gives finally

$$\left(\frac{\partial U}{\partial V}\right)_T - T \left(\frac{\partial p}{\partial T}\right)_V + p = 0$$

as the form the law takes when referred to the variables V and T.

5. Additional rules.* For the sake of completeness, we add two more rules which will enable you to make even uncommon selections of independent variables.

To state these last two rules, we need a determinant called the **Jacobian**. We give the notation and definition for two functions $u(x,y)$ and $v(x,y)$:

(15)
$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} \qquad \text{(the Jacobian)};$$

for three functions of three variables, etc. the definition would be analogous.

(16a,b)
$$\left(\frac{\partial u}{\partial x}\right)_v \left(\frac{\partial v}{\partial y}\right)_x = \frac{\partial(u,v)}{\partial(x,y)}; \qquad \left(\frac{\partial u}{\partial x}\right)_v = \left(\frac{\partial u}{\partial x}\right)_y - \left(\frac{\partial u}{\partial v}\right)_x \left(\frac{\partial v}{\partial x}\right)_y$$
$$\textbf{Jacobian rule}$$

(17)
$$\left(\frac{\partial u}{\partial v}\right)_w = \frac{\partial(u,w)/\partial(x,y)}{\partial(v,w)/\partial(x,y)} \qquad \textbf{two-Jacobian rule}$$

We leave the proof of the Jacobian rule (16b) as a good exercise in the use of differentials; the form (16a) follows from it by applying the chain rule (9b) and the definition (15).

The two-Jacobian rule can be proved directly either with differentials or the standard chain rule for functions of several variables. It is the mother of all rules: the other four can be derived from it by making some of the variables equal to each other.

As in section 4, these new rules allow the remaining choices of independent variable:

(18)
$$\left(\frac{\partial u}{\partial x}\right)_v = \left(\frac{\partial u}{\partial x}\right)_y - \frac{(\partial u/\partial y)_x}{(\partial v/\partial y)_x} \left(\frac{\partial v}{\partial x}\right)_y, \qquad \text{by (16b)}$$

(19)
$$\left(\frac{\partial u}{\partial v}\right)_w = \frac{\partial(u,w)/\partial(x,y)}{\partial(v,w)/\partial(x,y)}, \qquad \text{by (17)}$$

Exercises: Section 2J

18.02 Notes and Exercises by A. Mattuck and Bjorn Poonen with the assistance of T.Shifrin and S. LeDuc