Lagrange Multipliers

We will give the argument for why Lagrange multipliers work later. Here, we'll look at where and how to use them. Lagrange multipliers are used to solve constrained optimization problems. That is, suppose you have a function, say $f(x, y)$, for which you want to find the maximum or minimum value. But, you are not allowed to consider all (x, y) while you look for this value. Instead, the (x, y) you can consider are constrained to lie on some curve or surface. There are lots of examples of this in science, engineering and economics, for example, optimizing some utility function under budget constraints.

Lagrange multipliers problem:
Minimize (or maximize) $w = f(x, y, z)$ constrained by $g(x, y, z) = c$.

Lagrange multipliers solution:
Local minima (or maxima) must occur at a *critical point*. This is a point where $\boldsymbol{\nabla} f = \lambda \boldsymbol{\nabla} g$, and $g(x, y, z) = c$.

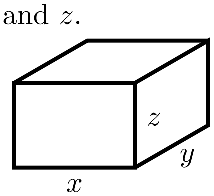
Example: Making a box using a minimum amount of material.

A box is made of cardboard with double thick sides, a triple thick bottom, single thick front and back and no top. It's volume is fixed at 3.

What dimensions use the least amount of cardboard?

Answer: We did this problem once before by solving for z in terms of x and y and substituting for it. That led to an unconstrained optimization problem in x and y. Here we will do it as a constrained problem. It is important to be able to do this because eliminating one variable is not always easy.

The box shown has dimensions x, y, and z.



The area of one side $= yz$. There are two double thick sides $\Rightarrow$ cardboard used $= 4yz$.

The area of the front (and back) $= xz$. It is single thick $\Rightarrow$ cardboard used $= 2xz$.

The area of the bottom $= xy$. It is triple thick $\Rightarrow$ cardboard used $= 3xy$.

Thus, the total cardboard used is

$$w = f(x, y, z) = 4yz + 2xz + 3xy.$$

The fixed volume acts as the constraint. It forces a relation between x, y and z so they can't all be varied independently. The constraint is

$$V = xyz = 3.$$

Our first job is to set up the equations to look for critical points. $\boldsymbol{\nabla} f = \langle 2z + 3y, 4z + 3x, 4y + 2x \rangle$ and $\boldsymbol{\nabla} V = \langle yz, xz, xy \rangle$.

The Lagrange multiplier equations are then

$$\begin{aligned} & \boldsymbol{\nabla} f = \lambda \boldsymbol{\nabla} V, \text{ and } V = 3 \\ \Leftrightarrow \quad & \langle 2z + 3y,\, 4z + 3x,\, 4y + 2x \rangle = \lambda \langle yz,\, xz,\, xy \rangle, \quad xyz = 3 \end{aligned}$$

Next we solve these equations for critical points. We do this by solving for λ in each equation (we call this *solving symmetrically*).

$\frac{2z+3y}{yz} = \lambda \quad \frac{4z+3x}{xz} = \lambda, \quad \frac{4y+2x}{xy} = \lambda, \quad xyz = 3 \qquad \Rightarrow \frac{2}{y} + \frac{3}{z} = \frac{4}{x} + \frac{3}{z} = \frac{4}{x} + \frac{2}{y}$

$\Rightarrow \frac{2}{y} = \frac{4}{x} \Rightarrow x = 2y$ and $\frac{3}{z} = \frac{2}{y} \Rightarrow z = \frac{3}{2}y$

Now, $xyz = 3 \Rightarrow 3y^3 = 3 \Rightarrow y = 1$

Answer: $x = 2, \quad y = 1, \quad z = \frac{3}{2}, \quad w = 18.$

Sphere example:

Minimize $w = y$ constrained to $x^2 + y^2 + z^2 = 1$.

Answer: $\nabla f = \langle 0, 1, 0\rangle$, $\nabla g = \langle 2x, 2y, 2z\rangle$

$\nabla f = \lambda \nabla g \Rightarrow \langle 0, 1, 0\rangle = \lambda\langle 2x, 2y, 2z\rangle \Rightarrow x = z = 0.$

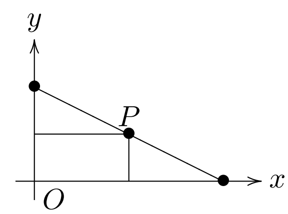
Constraint $\Rightarrow y = \pm 1$. (Gives the minimum and maximum respectively).

Example: (**checking the boundary**)

A rectangle in the plane is placed in the first quadrant so that one corner O is at the origin and the two sides adjacent to O are on the axes. The corner P opposite O is on the curve $x + 2y = 1$. Using Lagrange multipliers find for which point P the rectangle has maximum area. Say how you know this point gives the maximum.

Answer: We need some names

$$g(x, y) = x + 2y = 1 = \text{the constraint} \quad \text{and} \quad f(x, y) = xy = \text{the area.}$$



The gradients are: $\nabla g = \hat{\mathbf{i}} + 2\hat{\mathbf{j}}, \quad \nabla f = y\,\hat{\mathbf{i}} + x\,\hat{\mathbf{j}}$.

Lagrange multipliers: $\Rightarrow y = \lambda, \quad x = 2\lambda, \quad x + 2y = 1.$

The first two equations $\Rightarrow x = 2y$;

Combine this with the third equation $\Rightarrow 4y = 1$.

$\Rightarrow y = 1/4, \quad x = 1/2 \Rightarrow P = (1/2, 1/4).$

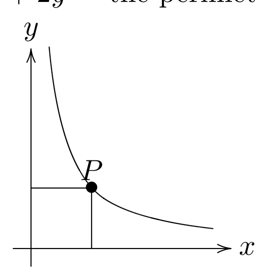
We know this is a maximum because the maximum occurs either at a critical point or on the boundary. In this case, the boundary points are on the axes at (1,0) and (0,1/2), which gives a rectangle with area = 0.

Example: (**boundary at ∞**)

A rectangle in the plane is placed in the first quadrant so that one corner O is at the origin and the two sides adjacent to O are on the axes. The corner P opposite O is on the curve $xy = 1$. Using Lagrange multipliers find for which point P the rectangle has minimum perimeter. Say how you know this point gives the minimum.

Answer: Let $g(x, y) = xy = 1$ = the constraint and $f(x, y) = 2x + 2y$ = the perimeter.

Gradients: $\nabla g = y\,\hat{\mathbf{i}} + x\,\hat{\mathbf{j}}, \quad \nabla f = 2\hat{\mathbf{i}} + 2\hat{\mathbf{j}}$.



Lagrange multipliers: $\Rightarrow$
$$2 = \lambda y$$
$$2 = \lambda x$$
$$xy = 1$$

The first two equations $\Rightarrow x = y$;

Combine this with the third equation $\Rightarrow x^2 = 1$.

$\Rightarrow x = 1, \quad x = 1 \Rightarrow P = (1, 1).$

We know this is a minimum because the minimum occurs either at a critical point or on the boundary. In this case the boundary points are infinitely far out on the axes which gives a rectangle with perimeter = ∞.

Proof of Lagrange Multipliers

Here we will give two arguments, one geometric and one analytic for why Lagrange multipliers work.

Critical points

For the function $w = f(x, y, z)$ constrained by $g(x, y, z) = c$ (c a constant) the critical points are defined as those points, which satisfy the constraint and where ∇f is parallel to ∇g. In equations:

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) \quad \textit{and} \quad g(x, y, z) = c.$$

Statement of Lagrange multipliers

For the constrained system local maxima and minima (collectively extrema) occur at the critical points.

Geometric proof for Lagrange

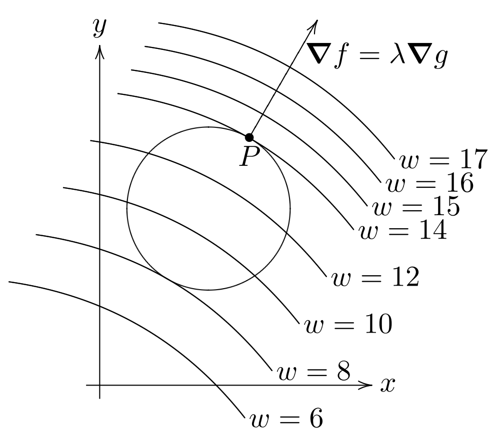
(We only consider the two dimensional case, $w = f(x, y)$ with constraint $g(x, y) = c$.)

For concreteness, we've drawn the constraint curve, $g(x, y) = c$, as a circle and some level curves for $w = f(x, y) = c$ with explicit (made up) values. Geometrically, we are looking for the point on the circle where w takes its maximum or minimum values.

Now, start at the level curve with $w = 17$, which has no points on the circle. So, clearly, the maximum value of w on the constraint circle is less than 17. Move down the level curves until they first touch the circle when $w = 14$. Call the point where the first touch P. It is clear that P gives a local maximum for w on $g = c$, because if you move away from P in either direction on the circle you'll be on a level curve with a smaller value.

Since the circle is a level curve for g, we know ∇g is perpendicular to it. We also know ∇f is perpendicular to the level curve $w = 14$, since the curves themselves are tangent, these two gradients must be parallel.

Likewise, if you keep moving down the level curves, the last one to touch the circle will give a local minimum and the same argument will apply.



Analytic proof for Lagrange (in three dimensions)

Suppose f has a local maximum at P on the constraint surface.

Let $\mathbf{r}(t) = \langle x(t), y(t), z(t) \rangle$ be an arbitrary parametrized curve which lies on the constraint surface and has $(x(0), y(0), z(0)) = P$. Finally, let $h(t) = f(x(t), y(t), z(t))$. The setup guarantees that $h(t)$ has a maximum at $t = 0$.

Taking a derivative using the chain rule in vector form gives

$$h'(t) = \boldsymbol{\nabla} f|_{\mathbf{r}(t)} \cdot \mathbf{r}'(t).$$

Since $t = 0$ is a local maximum, we have

$$h'(0) = \boldsymbol{\nabla} f|_P \cdot \mathbf{r}'(0) = 0.$$

Thus, $\boldsymbol{\nabla} f|_P$ is perpendicular to any curve on the constraint surface through P.

This implies $\boldsymbol{\nabla} f|_P$ is perpendicular to the surface. Since $\boldsymbol{\nabla} g|_P$ is also perpendicular to the surface we have proved $\boldsymbol{\nabla} f|_P$ is parallel to $\boldsymbol{\nabla} g|_P$. QED

Non-independent Variables

1. We give a worked example here. A fuller explanation will be given in the next session. Let

$$w = x^3y^2 + x^2y^3 + y$$

and assume x and y satisfy the relation

$$x^2 + y^2 = 1.$$

We consider x to be the independent variable, then, because y depends on x we have w is ultimately a function of the single variable x.

a) Compute $\frac{dw}{dx}$ using implicit differentiation.

b) Compute $\frac{dw}{dx}$ using total differentials.

<u>**Answer:**</u>

a) Implicit differentiation means remembering that y is a function of x, e.g., $\frac{dy^2}{dx} = 2y\frac{dy}{dx}$. Thus,

$$\frac{dw}{dx} = 3x^2y^2 + 2x^3y\frac{dy}{dx} + 2xy^3 + 3x^2y^2\frac{dy}{dx} + \frac{dy}{dx}.$$

Now we differentiate the constraint to find $\frac{dy}{dx}$.

$$x^2 + y^2 = 1 \;\Rightarrow\; 2x + 2y\frac{dy}{dx} = 0 \;\Rightarrow\; \frac{dy}{dx} = -\frac{x}{y}.$$

Substituting this in the equation for $\frac{dw}{dx}$ gives

$$\frac{dw}{dx} = 3x^2y^2 - 2x^3y\frac{x}{y} + 2xy^3 - 3x^2y^2\frac{x}{y} - \frac{x}{y} = 3x^2y^2 - 2x^4 + 2xy^3 - 3x^3y - \frac{x}{y}.$$

b) Taking total differentials of both w and the constraint equation gives

$$\begin{aligned} dw &= 3x^2y^2\,dx + 2x^3y\,dy + 2xy^3\,dx + 3x^2y^2\,dy + dy \\ &= (3x^2y^2 + 2xy^3)\,dx + (2x^3y + 3x^2y^2 + 1)\,dy \end{aligned}$$

$$2x\,dx + 2y\,dy = 0.$$

We can solve the second equation for dy and substitute in the equation for dw.

$$\begin{aligned} dy &= -\frac{x}{y}\,dx \;\Rightarrow \\ dw &= (3x^2y^2 + 2xy^3)\,dx + (2x^3y + 3x^2y^2 + 1)\left(-\frac{x}{y}\right)\,dx \\ &= (3x^2y^2 - 2x^4 + 2xy^3 - 3x^3y - \frac{x}{y})\,dx \end{aligned}$$

Thus,

$$\frac{dw}{dx} = 3x^2y^2 - 2x^4 + 2xy^3 - 3x^3y - \frac{x}{y}.$$

Non-independent Variables

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

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

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

$$\begin{aligned}
\left(\frac{\partial f}{\partial x}\right)_y &= f_x + f_z\left(\frac{\partial z}{\partial x}\right)_y = y^2z^4 + 8xy^2z^3;\\
\left(\frac{\partial f}{\partial y}\right)_x &= f_y + f_z\left(\frac{\partial z}{\partial y}\right)_x = 2xyz^4 + 12xy^2z^3;\\
\left(\frac{\partial f}{\partial z}\right)_x &= f_y\left(\frac{\partial y}{\partial z}\right)_x + f_z = \tfrac{2}{3}xyz^4 + 4xy^2z^3.
\end{aligned}$$

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

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

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

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:

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

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 $\left(\frac{\partial w}{\partial r}\right)_\theta$ 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,

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

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

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

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

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

(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

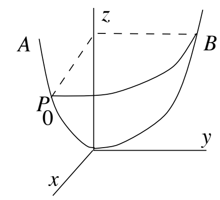
$$\begin{aligned} w &= x^2 + y^2 + z^2 = x^2 + (z - x^2) + z^2 \\ &= z + z^2; \\ \frac{\partial w}{\partial x} &= 0. \end{aligned}$$

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

$$w = f(x, y, z, t), \qquad \text{where } xy = zt, \tag{2}$$

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

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

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

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)
$$dw = f_x dx + f_y dy + f_z dz,$$

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 $\left(\frac{\partial w}{\partial y}\right)_{x,t}$, where $w = x^3y - z^2t$ 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)
$$dw = 3x^2y\,dx + x^3dy - 2zt\,dz - z^2dt, \qquad y\,dx + x\,dy = z\,dt + t\,dz.$$

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

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

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

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

(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:*

$$A\,dx + B\,dy + \ldots = A_1\,dx + B_1\,dy + \ldots \qquad \Leftrightarrow \qquad A = A_1,\ B = B_1, \ldots ; \tag{5}$$

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

$$dw = 2x\,dx - z\,dy - y\,dz + 2t\,dt, \qquad x\,dy + y\,dx = z\,dt + t\,dz, \qquad 2z\,dz = dx + 2y\,dy. \tag{6}$$

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

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

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

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

Problems: Lagrange Multipliers

1. Find the maximum and minimum values of $f(x, y) = x^2 + x + 2y^2$ on the unit circle.

Answer: The objective function is $f(x, y)$. The constraint is $g(x, y) = x^2 + y^2 = 1$.

Lagrange equations:	$f_x = \lambda g_x$	$\Leftrightarrow$	$2x + 1 = \lambda 2x$
	$f_y = \lambda g_y$	$\Leftrightarrow$	$4y = \lambda 2y$
Constraint:			$x^2 + y^2 = 1$

The second equation shows $y = 0$ or $\lambda = 2$.

$\lambda = 2 \Rightarrow x = 1/2,\ y = \pm\sqrt{3}/2$.

$y = 0 \Rightarrow x = \pm 1$.

Thus, the critical points are $(1/2, \sqrt{3}/2)$, $(1/2, -\sqrt{3}/2)$, $(1, 0)$, and $(-1, 0)$.

$f(1/2, \pm\sqrt{3/2}) = 9/4$ (maximum).

$f(1, 0) = 2$ (neither min. nor max).

$f(-1, 0) = 0$ (minimum).

2. Find the minimum and maximum values of $f(x, y) = x^2 - xy + y^2$ on the quarter circle $x^2 + y^2 = 1$, $x, y \geq 0$.

Answer: The constraint function here is $g(x, y) = x^2 + y^2 = 1$. The maximum and minimum values of $f(x, y)$ will occur where $\nabla f = \lambda \nabla g$ or at endpoints of the quarter circle.

$$\nabla f = \langle 2x - y, -x + 2y \rangle \quad \text{and} \quad \nabla g = \langle 2x, 2y \rangle.$$

Setting $\nabla f = \lambda \nabla g$, we get $2x - y = \lambda \cdot 2x$ and $-x + 2y = \lambda \cdot 2y$.

Solving for λ and setting the results equal to each other gives us:

$$\begin{aligned} \frac{2x - y}{2x} &= \frac{-x + 2y}{2y} \\ 2xy - y^2 &= -x^2 + 2xy \\ x^2 &= y^2. \end{aligned}$$

Because we're constrained to $x^2 + y^2 = 1$ with x and y non-negative, we conclude that $x = y = \frac{1}{\sqrt{2}}$.

Thus, the extreme points of $f(x, y)$ will be at $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, $(1, 0)$, or $(0, 1)$.

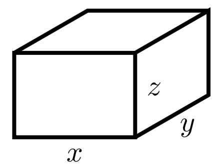
$$f(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = \frac{1}{2} \quad \text{is the minimum value of } f \text{ on this quarter circle.}$$

$$f(1, 0) = f(0, 1) = 1 \quad \text{are the maximal values of } f \text{ on this quarter circle.}$$

Lagrange multipliers

1. In an open-top wooden drawer, the two sides and back cost \$2/sq. ft., the bottom \$1/sq. ft. and the front \$4/sq. ft. Using Lagrange multipliers find the dimensions of the drawer with the largest capacity that can be made for \$72.

Answer: The box shown has dimensions x, y, and z.



The area of each side $= yz$; the area of the front (and back) $= xz$; the area of the bottom $= xy$. Thus, the cost of the wood is

$$C(x, y, z) = 2(2yz + xz) + xy + 4xz = 4yz + 6xz + xy = 72.$$

This is our constraint. We are trying to maximize the volume

$$V = xyz.$$

The Lagrange multiplier equations are then

$$\begin{aligned} & \boldsymbol{\nabla} V = \lambda \boldsymbol{\nabla} C, \text{ and } C = 72 \\ \Leftrightarrow \quad & \langle yz,\, xz,\, xy \rangle = \lambda \langle 6z + y,\, 4z + x,\, 4y + 6x \rangle, \quad 4yz + 6xz + xy = 72. \end{aligned}$$

We solve for the critical points by isolating $1/\lambda$.

$$\frac{1}{\lambda} = \frac{6}{y} + \frac{1}{z} = \frac{4}{x} + \frac{1}{z} = \frac{4}{x} + \frac{6}{y}$$

Comparing the third and fourth terms gives $\dfrac{1}{z} = \dfrac{6}{y} \;\Rightarrow\; y = 6z$.
Likewise the second and fourth terms give $x = 4z$.
Substituting this in the constraint gives $72z^2 = 72 \;\Rightarrow\; z = 1$. Thus,

$$z = 1,\; x = 4,\; y = 6.$$

Problems: Non-independent Variables

1. Find the total differential for $w = zxe^y + xe^z + ye^z$.

Answer:

$$\begin{aligned} dw &= ze^y\,dx + zxe^y\,dy + xe^y\,dz + e^z\,dx + xe^z\,dz + e^z\,dy + ye^z\,dz \\ &= (ze^y + e^z)dx + (zxe^y + e^z)dy + (xe^y + xe^z + ye^z)dz. \end{aligned}$$

2. With w as above, suppose we have $x = t$, $y = t^2$ and $z = t^3$. Write dw in terms of dt.

Answer: Here $dx = dt$, $dy = 2t\,dt$ and $dz = 3t^2\,dt$. We do not substitute for x, y and z because it does not greatly simplify the expression for dw and because in practice those values may be given or easily calculated from t.

$$dw = (ze^y + e^z)dt + (zxe^y + e^z)2t\,dt + (xe^y + xe^z + ye^z)3t^2\,dt.$$

3. Now suppose w is as above and $x^2y + y^2x = 1$. Assuming x is the independent variable, find $\frac{\partial w}{\partial x}$.

Answer: The constraint $x^2y + y^2x = 1$ becomes $(2xy + y^2)dx + (x^2 + 2xy)dy = 0$. Solving for dy in terms of x, y and dx we get $dy = \frac{2xy + y^2}{x^2 + 2xy}\,dx$.

Using the equation for dw from (1) gives:

$$\begin{aligned} dw &= (ze^y + e^z)dx + (zxe^y + e^z)dy + (xe^y + xe^z + ye^z)dz \\ &= (0 + e^0)dx + (0 + e^0)\left(\frac{2xy + y^2}{x^2 + 2xy}\,dx\right) + (xe^y + xe^z + ye^z)dz \\ &= dx + \frac{2xy + y^2}{x^2 + 2xy}\,dx + (xe^y + xe^z + ye^z)dz \\ &= \frac{x^2 + 4xy + y^2}{x^2 + 2xy}\,dx + (xe^y + xe^z + ye^z)dz. \end{aligned}$$

Thus, $\frac{\partial w}{\partial x} = \frac{x^2 + 4xy + y^2}{x^2 + 2xy}$.

Problems: The Chain Rule with Constraints

Suppose $w = u^3 - uv^2$, $u = xy$ and $v = u + x$.

1. Find $\left(\dfrac{\partial w}{\partial u}\right)_x$ and $\left(\dfrac{\partial w}{\partial x}\right)_u$ using the chain rule.

Answer: In finding $\left(\frac{\partial w}{\partial u}\right)_x$ we assume v and y are functions of u and that x is a constant.

$$\begin{aligned}\left(\frac{\partial w}{\partial u}\right)_x &= (3u^2 - v^2) - u \cdot 2v\left(\frac{\partial v}{\partial u}\right)_x \\ &= 3u^2 - v^2 - 2uv.\end{aligned}$$

Similarly,

$$\begin{aligned}\left(\frac{\partial w}{\partial x}\right)_u &= 0 - u \cdot 2v\left(\frac{\partial v}{\partial x}\right)_u \\ &= -2uv.\end{aligned}$$

2. Find $\left(\dfrac{\partial w}{\partial u}\right)_x$ and $\left(\dfrac{\partial w}{\partial x}\right)_u$ using differentials.

Answer: We can compute:

$$dw = (3u^2 - v^2)du - 2uv\,dv; \qquad du = x\,dy + y\,dx; \qquad dv = du + dx.$$

We're interested in the independent variables u and x so we substitute $dv = du + dx$ to get:

$$dw = (3u^2 - v^2)du - 2uv(du + dx) = (3u^2 - v^2 - 2uv)du - 2uv\,dx.$$

Using the fact that $dw = \left(\dfrac{\partial w}{\partial u}\right)_x du + \left(\dfrac{\partial w}{\partial x}\right)_u dx$, we get the expected answer.

Note that we did not need the variable y or the equation $u = xy$ in these calculations!

Chain rule with constraints

1. Let $P = (1, 2, 3)$ and assume $f(x, y, z)$ is a differentiable function with $\nabla f = \mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$ at P. Also assume that x, y and z satisfy the relation $x^3 - y^2 + z = 0$.

Take x and y to be the independent variables and let $g(x, y) = f(x, y, z(x, y))$. Find ∇g at the point (1,2).

<u>**Answer:**</u> Since f and g are the same, we have $df = dg$. The reason for using two symbols is that f is formally a function of x, y and z and g is formally a function of just x and y.

The gradient gives us the derivatives of f, so at P we have

$$df = dx - 2\,dy + 3\,dz.$$

The constraint gives us

$$3x^2\,dx - 2y\,dy + dz = 0 \;\Rightarrow\; dz = -3x^2\,dx + 2y\,dy.$$

At the point $(1, 2)$ this gives $dz = -3\,dx + 4\,dy$. Substituting this in the equation for df at P gives

$$df = dx - 2\,dy + 3(-3\,dx + 4\,dy) = -8\,dx + 10\,dy.$$

Having written df in terms of dx and dy we have found dg at (1,2). Thus $\dfrac{\partial g}{\partial x} = -8$ and $\dfrac{\partial g}{\partial y} = 10 \;\Rightarrow\; \nabla g = \langle -8, 10 \rangle$ at the point (1,2).