2.3 Interpretations of Partial Derivatives

This is a fairly short section and is here so we can acknowledge that the two main interpretations of derivatives of functions of a single variable still hold for partial derivatives, with small modifications of course to account of the fact that we now have more than one variable.

The first interpretation we've already seen and is the more important of the two. As with functions of single variables partial derivatives represent the rates of change of the functions as the variables change. As we saw in the previous section, $f_x(x, y)$ represents the rate of change of the function $f(x, y)$ as we change x and hold y fixed while $f_y(x, y)$ represents the rate of change of $f(x, y)$ as we change y and hold x fixed.

Example 60

Determine if $f(x, y) = \frac{x^2}{y^3}$ is increasing or decreasing at $(2, 5)$,

(a) if we allow x to vary and hold y fixed.

(b) if we allow y to vary and hold x fixed.

Solution

(a) if we allow x to vary and hold y fixed.

In this case we will first need $f_x(x, y)$ and its value at the point.

$$f_x(x, y) = \frac{2x}{y^3} \qquad \Rightarrow \qquad f_x(2, 5) = \frac{4}{125} > 0$$

So, the partial derivative with respect to x is positive and so if we hold y fixed the function is increasing at $(2, 5)$ as we vary x.

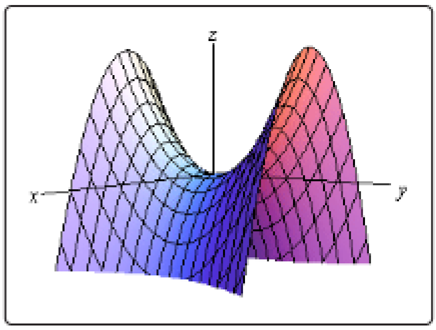
(b) if we allow y to vary and hold x fixed.

For this part we will need $f_y(x, y)$ and its value at the point.

$$f_y(x, y) = -\frac{3x^2}{y^4} \qquad \Rightarrow \qquad f_y(2, 5) = -\frac{12}{625} < 0$$

Here the partial derivative with respect to y is negative and so the function is decreasing at $(2, 5)$ as we vary y and hold x fixed.

Note that it is completely possible for a function to be increasing for a fixed y and decreasing for a fixed x at a point as this example has shown. To see a nice example of this take a look at the following graph.



This is a graph of a hyperbolic paraboloid and at the origin we can see that if we move in along the y-axis the graph is increasing and if we move along the x-axis the graph is decreasing. So it is completely possible to have a graph both increasing and decreasing at a point depending upon the direction that we move. We should never expect that the function will behave in exactly the same way at a point as each variable changes.

The next interpretation was one of the standard interpretations in a Calculus I class. We know from a Calculus I class that $f'(a)$ represents the slope of the tangent line to $y = f(x)$ at $x = a$. Well, $f_x(a, b)$ and $f_y(a, b)$ also represent the slopes of tangent lines. The difference here is the functions that they represent tangent lines to.

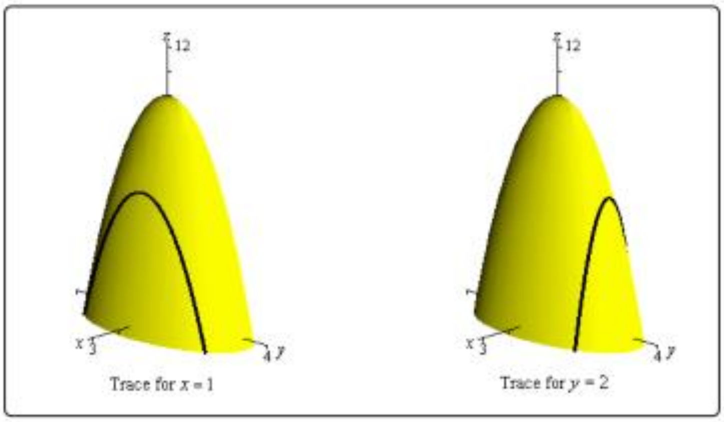
Partial derivatives are the slopes of traces. The partial derivative $f_x(a, b)$ is the slope of the trace of $f(x, y)$ for the plane $y = b$ at the point (a, b). Likewise the partial derivative $f_y(a, b)$ is the slope of the trace of $f(x, y)$ for the plane $x = a$ at the point (a, b).

Example 61

Find the slopes of the traces to $z = 10 - 4x^2 - y^2$ at the point $(1, 2)$.

Solution

We sketched the traces for the planes $x = 1$ and $y = 2$ in a previous section and these are the two traces for this point. For reference purposes here are the graphs of the traces.



Next, we'll need the two partial derivatives so we can get the slopes.

$$f_x(x,y) = -8x \qquad f_y(x,y) = -2y$$

To get the slopes all we need to do is evaluate the partial derivatives at the point in question.

$$f_x(1,2) = -8 \qquad f_y(1,2) = -4$$

So, the tangent line at $(1,2)$ for the trace to $z = 10 - 4x^2 - y^2$ for the plane $y = 2$ has a slope of -8. Also the tangent line at $(1,2)$ for the trace to $z = 10 - 4x^2 - y^2$ for the plane $x = 1$ has a slope of -4.

Finally, let's briefly talk about getting the equations of the tangent line. Recall that the equation of a line in 3-D space is given by a vector equation. Also, to get the equation we need a point on the line and a vector that is parallel to the line.

The point is easy. Since we know the x-y coordinates of the point all we need to do is plug this into the equation to get the point. So, the point will be,

$$\left(a, b, f(a,b)\right)$$

The parallel (or tangent) vector is also just as easy. We can write the equation of the surface as a vector function as follows,

$$\vec{r}(x,y) = \langle x, y, z \rangle = \left\langle x, y, f(x,y) \right\rangle$$

We know that if we have a vector function of one variable we can get a tangent vector by differentiating the vector function. The same will hold true here. If we differentiate with respect to x we will get a tangent vector to traces for the plane $y = b$ (*i.e.* for fixed y) and if we differentiate with respect to y we will get a tangent vector to traces for the plane $x = a$ (or fixed x).

So, here is the tangent vector for traces with fixed y.

$$\vec{r}_x(x, y) = \langle 1, 0, f_x(x, y) \rangle$$

We differentiated each component with respect to x. Therefore, the first component becomes a 1 and the second becomes a zero because we are treating y as a constant when we differentiate with respect to x. The third component is just the partial derivative of the function with respect to x.

For traces with fixed x the tangent vector is,

$$\vec{r}_y(x, y) = \langle 0, 1, f_y(x, y) \rangle$$

The equation for the tangent line to traces with fixed y is then,

$$\vec{r}(t) = \langle a, b, f(a, b) \rangle + t\langle 1, 0, f_x(a, b) \rangle$$

and the tangent line to traces with fixed x is,

$$\vec{r}(t) = \langle a, b, f(a, b) \rangle + t\langle 0, 1, f_y(a, b) \rangle$$

Example 62

Write down the vector equations of the tangent lines to the traces to $z = 10 - 4x^2 - y^2$ at the point $(1, 2)$.

Solution

There really isn't all that much to do with these other than plugging the values and function into the formulas above. We've already computed the derivatives and their values at $(1, 2)$ in the previous example and the point on each trace is,

$$(1, 2, f(1, 2)) = (1, 2, 2)$$

Here is the equation of the tangent line to the trace for the plane $y = 2$.

$$\vec{r}(t) = \langle 1, 2, 2 \rangle + t\langle 1, 0, -8 \rangle = \langle 1 + t, 2, 2 - 8t \rangle$$

Here is the equation of the tangent line to the trace for the plane $x = 1$.

$$\vec{r}(t) = \langle 1, 2, 2 \rangle + t\langle 0, 1, -4 \rangle = \langle 1, 2 + t, 2 - 4t \rangle$$

2.4 Higher Order Partial Derivatives

Just as we had higher order derivatives with functions of one variable we will also have higher order derivatives of functions of more than one variable. However, this time we will have more options since we do have more than one variable.

Consider the case of a function of two variables, $f(x, y)$ since both of the first order partial derivatives are also functions of x and y we could in turn differentiate each with respect to x or y. This means that for the case of a function of two variables there will be a total of four possible second order derivatives. Here they are and the notations that we'll use to denote them.

$$
\begin{aligned}
(f_x)_x &= f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} \\
(f_x)_y &= f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} \\
(f_y)_x &= f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} \\
(f_y)_y &= f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}
\end{aligned}
$$

The second and third second order partial derivatives are often called mixed partial derivatives since we are taking derivatives with respect to more than one variable. Note as well that the order that we take the derivatives in is given by the notation for each these. If we are using the subscripting notation, *e.g.* f_{xy}, then we will differentiate from left to right. In other words, in this case, we will differentiate first with respect to x and then with respect to y. With the fractional notation, *e.g.* $\frac{\partial^2 f}{\partial y \partial x}$, it is the opposite. In these cases we differentiate moving along the denominator from right to left. So, again, in this case we differentiate with respect to x first and then y.

Let's take a quick look at an example.

Example 63

Find all the second order derivatives for $f(x, y) = \cos(2x) - x^2 \mathbf{e}^{5y} + 3y^2$.

Solution

We'll first need the first order derivatives so here they are.

$$
\begin{aligned}
f_x(x, y) &= -2\sin(2x) - 2x\mathbf{e}^{5y} \\
f_y(x, y) &= -5x^2\mathbf{e}^{5y} + 6y
\end{aligned}
$$

Now, let's get the second order derivatives.

$$\begin{aligned} f_{xx} &= -4\cos(2x) - 2\mathbf{e}^{5y} \\ f_{xy} &= -10x\mathbf{e}^{5y} \\ f_{yx} &= -10x\mathbf{e}^{5y} \\ f_{yy} &= -25x^2\mathbf{e}^{5y} + 6 \end{aligned}$$

Notice that we dropped the (x, y) from the derivatives. This is fairly standard and we will be doing it most of the time from this point on. We will also be dropping it for the first order derivatives in most cases.

Now let's also notice that, in this case, $f_{xy} = f_{yx}$. This is not by coincidence. If the function is "nice enough" this will always be the case. So, what's "nice enough"? The following theorem tells us.

Clairaut's Theorem

Suppose that f is defined on a disk D that contains the point (a, b). If the functions f_{xy} and f_{yx} are continuous on this disk then,

$$f_{xy}(a, b) = f_{yx}(a, b)$$

Now, do not get too excited about the disk business and the fact that we gave the theorem for a specific point. In pretty much every example in this class if the two mixed second order partial derivatives are continuous then they will be equal.

Example 64

Verify Clairaut's Theorem for $f(x, y) = x\mathbf{e}^{-x^2y^2}$.

Solution

We'll first need the two first order derivatives.

$$\begin{aligned} f_x(x, y) &= \mathbf{e}^{-x^2y^2} - 2x^2y^2\mathbf{e}^{-x^2y^2} \\ f_y(x, y) &= -2yx^3\mathbf{e}^{-x^2y^2} \end{aligned}$$

Now, compute the two mixed second order partial derivatives.

$$\begin{aligned} f_{xy}(x, y) &= -2yx^2\mathbf{e}^{-x^2y^2} - 4x^2y\mathbf{e}^{-x^2y^2} + 4x^4y^3\mathbf{e}^{-x^2y^2} = -6x^2y\mathbf{e}^{-x^2y^2} + 4x^4y^3\mathbf{e}^{-x^2y^2} \\ f_{yx}(x, y) &= -6yx^2\mathbf{e}^{-x^2y^2} + 4y^3x^4\mathbf{e}^{-x^2y^2} \end{aligned}$$

Sure enough they are the same.

So far we have only looked at second order derivatives. There are, of course, higher order derivatives as well. Here are a couple of the third order partial derivatives of function of two variables.

$$f_{xyx} = (f_{xy})_x = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial^3 f}{\partial x \partial y \partial x}$$
$$f_{yxx} = (f_{yx})_x = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial^3 f}{\partial x^2 \partial y}$$

Notice as well that for both of these we differentiate once with respect to y and twice with respect to x. There is also another third order partial derivative in which we can do this, f_{xxy}. There is an extension to Clairaut's Theorem that says if all three of these are continuous then they should all be equal,

$$f_{xxy} = f_{xyx} = f_{yxx}$$

To this point we've only looked at functions of two variables, but everything that we've done to this point will work regardless of the number of variables that we've got in the function and there are natural extensions to Clairaut's theorem to all of these cases as well. For instance,

$$f_{xz}(x, y, z) = f_{zx}(x, y, z)$$

provided both of the derivatives are continuous.

In general, we can extend Clairaut's theorem to any function and mixed partial derivatives. The only requirement is that in each derivative we differentiate with respect to each variable the same number of times. In other words, provided we meet the continuity condition, the following will be equal

$$f_{ssrtsrr} = f_{trsrssr}$$

because in each case we differentiate with respect to t once, s three times and r three times.

Let's do a couple of examples with higher (well higher order than two anyway) order derivatives and functions of more than two variables.

Example 65

Find the indicated derivative for each of the following functions.

(a) Find f_{xxyzz} for $f(x,y,z) = z^3y^2\ln(x)$

(b) Find $\dfrac{\partial^3 f}{\partial y \partial x^2}$ for $f(x,y) = \mathbf{e}^{xy}$

Solution

(a) Find f_{xxyzz} for $f(x,y,z) = z^3y^2\ln(x)$

In this case remember that we differentiate from left to right. Here are the derivatives for this part.

$$f_x = \frac{z^3y^2}{x}$$
$$f_{xx} = -\frac{z^3y^2}{x^2}$$
$$f_{xxy} = -\frac{2z^3y}{x^2}$$
$$f_{xxyz} = -\frac{6z^2y}{x^2}$$
$$f_{xxyzz} = -\frac{12zy}{x^2}$$

(b) Find $\dfrac{\partial^3 f}{\partial y \partial x^2}$ for $f(x,y) = \mathbf{e}^{xy}$

Here we differentiate from right to left. Here are the derivatives for this function.

$$\frac{\partial f}{\partial x} = y\mathbf{e}^{xy}$$
$$\frac{\partial^2 f}{\partial x^2} = y^2\mathbf{e}^{xy}$$
$$\frac{\partial^3 f}{\partial y \partial x^2} = 2y\mathbf{e}^{xy} + xy^2\mathbf{e}^{xy}$$

2.5 Differentials

This is a very short section and is here simply to acknowledge that just like we had differentials for functions of one variable we also have them for functions of more than one variable. Also, as we've already seen in previous sections, when we move up to more than one variable things work pretty much the same, but there are some small differences.

Given the function $z = f(x, y)$ the differential dz or df is given by,

$$dz = f_x\,dx + f_y\,dy \qquad \text{or} \qquad df = f_x\,dx + f_y\,dy$$

There is a natural extension to functions of three or more variables. For instance, given the function $w = g(x, y, z)$ the differential is given by,

$$dw = g_x\,dx + g_y\,dy + g_z\,dz$$

Let's do a couple of quick examples.

Example 66

Compute the differentials for each of the following functions.

(a) $z = \mathbf{e}^{x^2+y^2}\tan(2x)$

(b) $u = \dfrac{t^3r^6}{s^2}$

Solution

(a) $z = \mathbf{e}^{x^2+y^2}\tan(2x)$

There really isn't a whole lot to these outside of some quick differentiation. Here is the differential for the function.

$$dz = \left(2x\mathbf{e}^{x^2+y^2}\tan(2x) + 2\mathbf{e}^{x^2+y^2}\sec^2(2x)\right)dx + 2y\mathbf{e}^{x^2+y^2}\tan(2x)\,dy$$

(b) $u = \dfrac{t^3r^6}{s^2}$

Here is the differential for this function.

$$du = \frac{3t^2r^6}{s^2}dt + \frac{6t^3r^5}{s^2}dr - \frac{2t^3r^6}{s^3}ds$$

Note that sometimes these differentials are called the **total differentials**.

2.6 Chain Rule

We've been using the standard chain rule for functions of one variable throughout the last couple of sections. It's now time to extend the chain rule out to more complicated situations. Before we actually do that let's first review the notation for the chain rule for functions of one variable.

The notation that's probably familiar to most people is the following.

$$F(x) = f(g(x)) \qquad F'(x) = f'(g(x))g'(x)$$

There is an alternate notation however that while probably not used much in Calculus I is more convenient at this point because it will match up with the notation that we are going to be using in this section. Here it is.

$$\text{If} \quad y = f(x) \quad \text{and} \quad x = g(t) \quad \text{then} \quad \frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$$

Notice that the derivative $\frac{dy}{dt}$ really does make sense here since if we were to plug in for x then y really would be a function of t. One way to remember this form of the chain rule is to note that if we think of the two derivatives on the right side as fractions the dx's will cancel to get the same derivative on both sides.

Okay, now that we've got that out of the way let's move into the more complicated chain rules that we are liable to run across in this course.

As with many topics in multivariable calculus, there are in fact many different formulas depending upon the number of variables that we're dealing with. So, let's start this discussion off with a function of two variables, $z = f(x, y)$. From this point there are still many different possibilities that we can look at. We will be looking at two distinct cases prior to generalizing the whole idea out.

Case 1 : $z = f(x, y)$, $x = g(t)$, $y = h(t)$ **and compute** $\frac{dz}{dt}$.

This case is analogous to the standard chain rule from Calculus I that we looked at above. In this case we are going to compute an ordinary derivative since z really would be a function of t only if we were to substitute in for x and y.

The chain rule for this case is,

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

So, basically what we're doing here is differentiating f with respect to each variable in it and then multiplying each of these by the derivative of that variable with respect to t. The final step is to then add all this up.

Let's take a look at a couple of examples.

Example 67

Compute $\frac{dz}{dt}$ for each of the following.

(a) $z = x\mathbf{e}^{xy}$, $x = t^2$, $y = t^{-1}$

(b) $z = x^2y^3 + y\cos(x)$, $x = \ln(t^2)$, $y = \sin(4t)$

Solution

(a) $z = x\mathbf{e}^{xy}$, $x = t^2$, $y = t^{-1}$

There really isn't all that much to do here other than using the formula.

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\\ &= \left(\mathbf{e}^{xy} + yx\mathbf{e}^{xy}\right)(2t) + x^2\mathbf{e}^{xy}\left(-t^{-2}\right)\\ &= 2t\left(\mathbf{e}^{xy} + yx\mathbf{e}^{xy}\right) - t^{-2}x^2\mathbf{e}^{xy}\end{aligned}$$

So, technically we've computed the derivative. However, we should probably go ahead and substitute in for x and y as well at this point since we've already got t's in the derivative. Doing this gives,

$$\frac{dz}{dt} = 2t\left(\mathbf{e}^{t} + t\mathbf{e}^{t}\right) - t^{-2}t^4\mathbf{e}^{t} = 2t\mathbf{e}^{t} + t^2\mathbf{e}^{t}$$

Note that in this case it might actually have been easier to just substitute in for x and y in the original function and just compute the derivative as we normally would. For comparison's sake let's do that.

$$z = t^2\mathbf{e}^{t} \qquad \Rightarrow \qquad \frac{dz}{dt} = 2t\mathbf{e}^{t} + t^2\mathbf{e}^{t}$$

The same result for less work. Note however, that often it will actually be more work to do the substitution first.

(b) $z = x^2y^3 + y\cos(x)$, $x = \ln(t^2)$, $y = \sin(4t)$

Okay, in this case it would almost definitely be more work to do the substitution first so we'll use the chain rule first and then substitute.

$$\begin{aligned}\frac{dz}{dt} &= \left(2xy^3 - y\sin(x)\right)\left(\frac{2}{t}\right) + \left(3x^2y^2 + \cos(x)\right)\left(4\cos(4t)\right)\\ &= \frac{4\sin^3(4t)\ln t^2 - 2\sin(4t)\sin\left(\ln t^2\right)}{t} + 4\cos(4t)\left(3\sin^2(4t)\left[\ln t^2\right]^2 + \cos\left(\ln t^2\right)\right)\end{aligned}$$

Note that sometimes, because of the significant mess of the final answer, we will only simplify the first step a little and leave the answer in terms of x, y, and t. This is dependent upon the situation, class and instructor however so be careful about not substituting in for without first talking to your instructor.

Now, there is a special case that we should take a quick look at before moving on to the next case. Let's suppose that we have the following situation,

$$z = f(x, y) \qquad y = g(x)$$

In this case the chain rule for $\frac{dz}{dx}$ becomes,

$$\frac{dz}{dx} = \frac{\partial f}{\partial x}\frac{dx}{dx} + \frac{\partial f}{\partial y}\frac{dy}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}$$

In the first term we are using the fact that,

$$\frac{dx}{dx} = \frac{d}{dx}(x) = 1$$

Let's take a quick look at an example.

Example 68

$\frac{dz}{dx}$ for $z = x\ln(xy) + y^3$, $y = \cos(x^2 + 1)$

Solution

We'll just plug into the formula.

$$\begin{aligned}
\frac{dz}{dx} &= \left(\ln(xy) + x\frac{y}{xy}\right) + \left(x\frac{x}{xy} + 3y^2\right)\left(-2x\sin\left(x^2+1\right)\right) \\
&= \ln\left(x\cos\left(x^2+1\right)\right) + 1 - 2x\sin\left(x^2+1\right)\left(\frac{x}{\cos\left(x^2+1\right)} + 3\cos^2\left(x^2+1\right)\right) \\
&= \ln\left(x\cos\left(x^2+1\right)\right) + 1 - 2x^2\tan\left(x^2+1\right) - 6x\sin\left(x^2+1\right)\cos^2\left(x^2+1\right)
\end{aligned}$$

Now let's take a look at the second case.

Case 2 : $z = f(x, y)$, $x = g(s, t)$, $y = h(s, t)$ **and compute** $\frac{\partial z}{\partial s}$ **and** $\frac{\partial z}{\partial t}$.

In this case if we were to substitute in for x and y we would get that z is a function of s and t and so it makes sense that we would be computing partial derivatives here and that there would be two of them.

Here is the chain rule for both of these cases.

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \qquad \frac{\partial z}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}$$

So, not surprisingly, these are very similar to the first case that we looked at. Here is a quick example of this kind of chain rule.

Example 69

Find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$ for $z = \mathbf{e}^{2r}\sin(3\theta)$, $r = st - t^2$, $\theta = \sqrt{s^2+t^2}$.

Solution

Here is the chain rule for $\frac{\partial z}{\partial s}$.

$$\begin{aligned}\frac{\partial z}{\partial s} &= \left(2\mathbf{e}^{2r}\sin(3\theta)\right)(t) + \left(3\mathbf{e}^{2r}\cos(3\theta)\right)\frac{s}{\sqrt{s^2+t^2}}\\ &= t\left(2\mathbf{e}^{2(st-t^2)}\sin\left(3\sqrt{s^2+t^2}\right)\right) + \frac{3s\mathbf{e}^{2(st-t^2)}\cos\left(3\sqrt{s^2+t^2}\right)}{\sqrt{s^2+t^2}}\end{aligned}$$

Now the chain rule for $\frac{\partial z}{\partial t}$.

$$\begin{aligned}\frac{\partial z}{\partial t} &= \left(2\mathbf{e}^{2r}\sin(3\theta)\right)(s-2t) + \left(3\mathbf{e}^{2r}\cos(3\theta)\right)\frac{t}{\sqrt{s^2+t^2}}\\ &= (s-2t)\left(2\mathbf{e}^{2(st-t^2)}\sin\left(3\sqrt{s^2+t^2}\right)\right) + \frac{3t\mathbf{e}^{2(st-t^2)}\cos\left(3\sqrt{s^2+t^2}\right)}{\sqrt{s^2+t^2}}\end{aligned}$$

Okay, now that we've seen a couple of cases for the chain rule let's see the general version of the chain rule.

Chain Rule

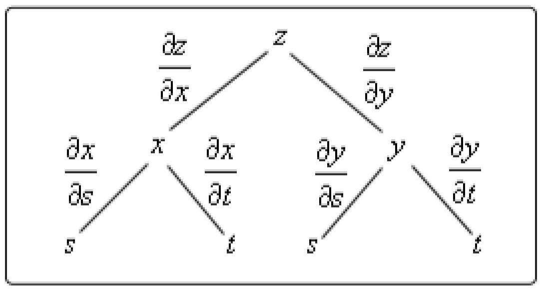
Suppose that z is a function of n variables, $x_1, x_2, \ldots, x_n$, and that each of these variables are in turn functions of m variables, $t_1, t_2, \ldots, t_m$. Then for any variable t_i, $i = 1, 2, \ldots, m$ we have the following,

$$\frac{\partial z}{\partial t_i} = \frac{\partial z}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial z}{\partial x_2}\frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial z}{\partial x_n}\frac{\partial x_n}{\partial t_i}$$

Wow. That's a lot to remember. There is actually an easier way to construct all the chain rules that we've discussed in the section or will look at in later examples. We can build up a **tree diagram** that will give us the chain rule for any situation. To see how these work let's go back and take a look at the chain rule for $\frac{\partial z}{\partial s}$ given that $z = f(x,y)$, $x = g(s,t)$, $y = h(s,t)$. We already know what this is, but it may help to illustrate the tree diagram if we already know the answer. For reference here is the chain rule for this case,

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

Here is the tree diagram for this case.



We start at the top with the function itself and the branch out from that point. The first set of branches is for the variables in the function. From each of these endpoints we put down a further set of branches that gives the variables that both x and y are a function of. We connect each letter with a line and each line represents a partial derivative as shown. Note that the letter in the numerator of the partial derivative is the upper "node" of the tree and the letter in the denominator of the partial derivative is the lower "node" of the tree.

To use this to get the chain rule we start at the bottom and for each branch that ends with the variable we want to take the derivative with respect to (s in this case) we move up the tree until we hit the top multiplying the derivatives that we see along that set of branches. Once we've done this for each branch that ends at s, we then add the results up to get the chain rule for that given situation.

Note that we don't always put the derivatives in the tree. Some of the trees get a little large/messy and so we won't put in the derivatives. Just remember what derivative should be on each branch and you'll be okay without actually writing them down.

Let's write down some chain rules.

Example 70

Use a tree diagram to write down the chain rule for the given derivatives.

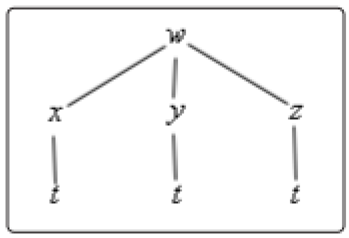
(a) $\frac{dw}{dt}$ for $w = f(x, y, z)$, $x = g_1(t)$, $y = g_2(t)$, and $z = g_3(t)$

(b) $\frac{\partial w}{\partial r}$ for $w = f(x, y, z)$, $x = g_1(s, t, r)$, $y = g_2(s, t, r)$, and $z = g_3(s, t, r)$

Solution

(a) $\frac{dw}{dt}$ for $w = f(x, y, z)$, $x = g_1(t)$, $y = g_2(t)$, and $z = g_3(t)$

So, we'll first need the tree diagram so let's get that.



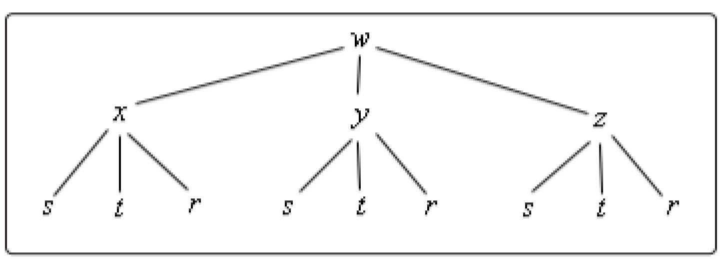
From this it looks like the chain rule for this case should be,

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

which is really just a natural extension to the two variable case that we saw above.

(b) $\frac{\partial w}{\partial r}$ for $w = f(x, y, z)$, $x = g_1(s, t, r)$, $y = g_2(s, t, r)$, and $z = g_3(s, t, r)$

Here is the tree diagram for this situation.



From this it looks like the derivative will be,

$$\frac{\partial w}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial r}$$

So, provided we can write down the tree diagram, and these aren't usually too bad to write down, we can do the chain rule for any set up that we might run across.

We've now seen how to take first derivatives of these more complicated situations, but what about higher order derivatives? How do we do those? It's probably easiest to see how to deal with these with an example.

Example 71

Compute $\frac{\partial^2 f}{\partial \theta^2}$ for $f(x, y)$ if $x = r\cos(\theta)$ and $y = r\sin(\theta)$.

Solution

We will need the first derivative before we can even think about finding the second derivative so let's get that. This situation falls into the second case that we looked at above so we don't need a new tree diagram. Here is the first derivative.

$$\begin{aligned}\frac{\partial f}{\partial \theta} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta} \\ &= -r\sin(\theta)\frac{\partial f}{\partial x} + r\cos(\theta)\frac{\partial f}{\partial y}\end{aligned}$$

Okay, now we know that the second derivative is,

$$\frac{\partial^2 f}{\partial \theta^2} = \frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial \theta}\right) = \frac{\partial}{\partial \theta}\left(-r\sin(\theta)\frac{\partial f}{\partial x} + r\cos(\theta)\frac{\partial f}{\partial y}\right)$$

The issue here is to correctly deal with this derivative. Since the two first order derivatives, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$, are both functions of x and y which are in turn functions of r and θ both of these terms are products. So, the using the product rule gives the following,

$$\frac{\partial^2 f}{\partial \theta^2} = -r\cos(\theta)\frac{\partial f}{\partial x} - r\sin(\theta)\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right) - r\sin(\theta)\frac{\partial f}{\partial y} + r\cos(\theta)\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial y}\right)$$

We now need to determine what $\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right)$ and $\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial y}\right)$ will be. These are both chain rule problems again since both of the derivatives are functions of x and y and we want to take the derivative with respect to θ.

Before we do these let's rewrite the first chain rule that we did above a little.

$$\frac{\partial}{\partial \theta}\Big(f\Big) = -r\sin(\theta)\frac{\partial}{\partial x}\Big(f\Big) + r\cos(\theta)\frac{\partial}{\partial y}\Big(f\Big) \tag{13.1}$$

Note that all we've done is change the notation for the derivative a little. With the first chain rule written in this way we can think of Equation 13.1 as a formula for differentiating any function of x and y with respect to θ provided we have $x = r\cos(\theta)$ and $y = r\sin(\theta)$.

This however is exactly what we need to do the two new derivatives we need above. Both of the first order partial derivatives, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$, are functions of x and y and $x = r\cos(\theta)$ and $y = r\sin(\theta)$ so we can use Equation 13.1 to compute these derivatives.

To do this we'll simply replace all the f's in Equation 13.1 with the first order partial derivative that we want to differentiate. At that point all we need to do is a little notational work and we'll get the formula that we're after.

Here is the use of Equation 13.1 to compute $\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right)$.

$$\begin{aligned}\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right) &= -r\sin(\theta)\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) + r\cos(\theta)\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\\ &= -r\sin(\theta)\frac{\partial^2 f}{\partial x^2} + r\cos(\theta)\frac{\partial^2 f}{\partial y\partial x}\end{aligned}$$

Here is the computation for $\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial y}\right)$.

$$\begin{aligned}\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial y}\right) &= -r\sin(\theta)\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) + r\cos(\theta)\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)\\ &= -r\sin(\theta)\frac{\partial^2 f}{\partial x\partial y} + r\cos(\theta)\frac{\partial^2 f}{\partial y^2}\end{aligned}$$

The final step is to plug these back into the second derivative and do some simplifying.

$$\begin{aligned}\frac{\partial^2 f}{\partial \theta^2} &= -r\cos(\theta)\frac{\partial f}{\partial x} - r\sin(\theta)\left(-r\sin(\theta)\frac{\partial^2 f}{\partial x^2} + r\cos(\theta)\frac{\partial^2 f}{\partial y\partial x}\right) -\\ &\qquad r\sin(\theta)\frac{\partial f}{\partial y} + r\cos(\theta)\left(-r\sin(\theta)\frac{\partial^2 f}{\partial x\partial y} + r\cos(\theta)\frac{\partial^2 f}{\partial y^2}\right)\\ &= -r\cos(\theta)\frac{\partial f}{\partial x} + r^2\sin^2(\theta)\frac{\partial^2 f}{\partial x^2} - r^2\sin(\theta)\cos(\theta)\frac{\partial^2 f}{\partial y\partial x} -\\ &\qquad r\sin(\theta)\frac{\partial f}{\partial y} - r^2\sin(\theta)\cos(\theta)\frac{\partial^2 f}{\partial x\partial y} + r^2\cos^2(\theta)\frac{\partial^2 f}{\partial y^2}\\ &= -r\cos(\theta)\frac{\partial f}{\partial x} - r\sin(\theta)\frac{\partial f}{\partial y} + r^2\sin^2(\theta)\frac{\partial^2 f}{\partial x^2} -\\ &\qquad 2r^2\sin(\theta)\cos(\theta)\frac{\partial^2 f}{\partial y\partial x} + r^2\cos^2(\theta)\frac{\partial^2 f}{\partial y^2}\end{aligned}$$

It's long and fairly messy but there it is.

The final topic in this section is a revisiting of implicit differentiation. With these forms of the chain rule implicit differentiation actually becomes a fairly simple process. Let's start out with the implicit differentiation that we saw in a Calculus I course.

We will start with a function in the form $F(x, y) = 0$ (if it's not in this form simply move everything to one side of the equal sign to get it into this form) where $y = y(x)$. In a Calculus I course we were then asked to compute $\frac{dy}{dx}$ and this was often a fairly messy process. Using the chain rule from this section however we can get a nice simple formula for doing this. We'll start by differentiating both sides with respect to x. This will mean using the chain rule on the left side and the right side will, of course, differentiate to zero. Here are the results of that.

$$F_x + F_y \frac{dy}{dx} = 0 \qquad \Rightarrow \qquad \frac{dy}{dx} = -\frac{F_x}{F_y}$$

As shown, all we need to do next is solve for $\frac{dy}{dx}$ and we've now got a very nice formula to use for implicit differentiation. Note as well that in order to simplify the formula we switched back to using the subscript notation for the derivatives.

Let's check out a quick example.

Example 72

Find $\dfrac{dy}{dx}$ for $x\cos(3y) + x^3y^5 = 3x - \mathbf{e}^{xy}$.

Solution

The first step is to get a zero on one side of the equal sign and that's easy enough to do.

$$x\cos(3y) + x^3y^5 - 3x + \mathbf{e}^{xy} = 0$$

Now, the function on the left is $F(x, y)$ in our formula so all we need to do is use the formula to find the derivative.

$$\frac{dy}{dx} = -\frac{\cos(3y) + 3x^2y^5 - 3 + y\mathbf{e}^{xy}}{-3x\sin(3y) + 5x^3y^4 + x\mathbf{e}^{xy}}$$

There we go. It would have taken much longer to do this using the old Calculus I way of doing this.

We can also do something similar to handle the types of implicit differentiation problems involving partial derivatives like those we saw when we first introduced partial derivatives. In these cases we will start off with a function in the form $F(x, y, z) = 0$ and assume that $z = f(x, y)$ and we want to find $\frac{\partial z}{\partial x}$ and/or $\frac{\partial z}{\partial y}$.

Let's start by trying to find $\frac{\partial z}{\partial x}$. We will differentiate both sides with respect to x and we'll need to remember that we're going to be treating y as a constant. Also, the left side will require the chain rule. Here is this derivative.

$$\frac{\partial F}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0$$

Now, we have the following,

$$\frac{\partial x}{\partial x} = 1 \qquad \text{and} \qquad \frac{\partial y}{\partial x} = 0$$

The first is because we are just differentiating x with respect to x and we know that is 1. The second is because we are treating the y as a constant and so it will differentiate to zero.

Plugging these in and solving for $\frac{\partial z}{\partial x}$ gives,

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}$$

A similar argument can be used to show that,

$$\frac{\partial z}{\partial y} = -\frac{F_y}{F_z}$$

As with the one variable case we switched to the subscripting notation for derivatives to simplify the formulas. Let's take a quick look at an example of this.

Example 73

Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for $x^2 \sin(2y - 5z) = 1 + y \cos(6zx)$.

Solution

This was one of the functions that we used the old implicit differentiation on back in the Partial Derivatives section. You might want to go back and see the difference between the two.

First let's get everything on one side.

$$x^2 \sin(2y - 5z) - 1 - y\cos(6zx) = 0$$

Now, the function on the left is $F(x, y, z)$ and so all that we need to do is use the formulas developed above to find the derivatives.

$$\frac{\partial z}{\partial x} = -\frac{2x\sin(2y-5z) + 6yz\sin(6zx)}{-5x^2\cos(2y-5z) + 6yx\sin(6zx)}$$

$$\frac{\partial z}{\partial y} = -\frac{2x^2\cos(2y-5z) - \cos(6zx)}{-5x^2\cos(2y-5z) + 6yx\sin(6zx)}$$

If you go back and compare these answers to those that we found the first time around you will notice that they might appear to be different. However, if you take into account the minus sign that sits in the front of our answers here you will see that they are in fact the same.