The Chain Rule

What to know:

1. Be able to use tree diagrams to write the chain rule for functions, and use the chain rule to find derivatives.

2. Be able to compute second partial derivatives and know the notation for them, e.g.: $\frac{\partial^2 z}{\partial x^2}$ and $\frac{\partial^2 z}{\partial x \partial y}$.

In Calculus 1, we learned how to use the chain rule to take the derivative of the composition $f \circ g$ of two functions f and g: Recall that

$$(f \circ g)'(t) = f'(g(t))g'(t).$$

Another way to write this, if $z = f(x)$ and $x = g(t)$ we had

$$\frac{dz}{dt} = \frac{df}{dx}\frac{dg}{dt} = \frac{dz}{dx}\frac{dx}{dt}.$$

Now we will generalize that in higher dimensions:

Theorem 1. *Suppose we have a differentiable function $z = f(x, y)$ and $x = g(t)$, $y = h(t)$, where g and h are differentiable. Then the composition $z = f(g(t), h(t))$ is a differentiable function of t and we have*

$$\frac{dz}{dt}_{|t} = \frac{\partial f}{\partial x}_{|(g(t),h(t))} \frac{dg}{dt}_{|t} + \frac{\partial f}{\partial y}_{|(g(t),h(t))} \frac{dh}{dt}_{|t}$$

This is also frequently written as

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

which partially justifies the term "chain rule". Let's see an example:

Example 1. *Let $z = x^2y$, where $x = \sin(t)$, $y = \cos(3t)$. Find $\frac{dz}{dt}$.*

Solution. We have $\frac{dx}{dt} = \cos(t)$, $\frac{dy}{dt} = -3\sin(3t)$ and $\frac{\partial z}{\partial x} = 2xy$, $\frac{\partial z}{\partial y} = x^2$. So

$$\begin{aligned}\frac{dz}{dt} =& \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}\\ =& 2xy\cos(t) + x^2(-3\sin(3t)).\end{aligned}$$

However $x = \sin(t)$ and $y = \cos(3t)$, so

$$\frac{dz}{dt} = 2\sin(t)\cos(3t)\cos(t) + \sin^2(t)(-3\sin(3t)).$$

□

The above rule generalizes for the case when x and y depend on more than one variables. The simplest such case is the following:

Theorem 2. *Suppose we have a differentiable function $z = f(x, y)$ and $x = g(u, v)$, $y = h(u, v)$, where g and h are differentiable. Then the composition $z = f(g(u, v), h(u, v))$ is a differentiable function of u, v and we have*

$$\frac{\partial z}{\partial u}_{|(u,v)} = \frac{\partial f}{\partial x}_{|f(u,v)} \frac{\partial g}{\partial u}_{|(u,v)} + \frac{\partial f}{\partial y}_{|g(u,v)} \frac{\partial g}{\partial u}_{|(u,v)}$$

and

$$\frac{\partial z}{\partial v}_{|(u,v)} = \frac{\partial f}{\partial x}_{|f(u,v)} \frac{\partial g}{\partial v}_{|(u,v)} + \frac{\partial f}{\partial y}_{|g(u,v)} \frac{\partial g}{\partial v}_{|(u,v)}$$

Again, the above equations can be written as

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}$$

and

$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}.$$

Remark: Note the distinction between d and ∂ symbols: we use ∂ when our function depends on more than one variables, but d when our function depends on one variable only.

An easy way to remember the chain rule is by using a tree diagram:

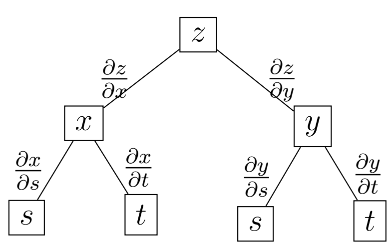
1. Under each function write the variables/functions it immediately depends upon. For example, if $z = z(x, y)$ and $x = x(s, t)$, $y = y(s, t)$ then under z we'd only put x and y, but not t because the dependency on t is not immediate.

2. Between a function/variable and each function/variable it depends on, we write the corresponding partial derivative. For example, between z and x, we write $\frac{\partial z}{\partial x}$.

3. If we want to find the derivative of z with respect to, say, t, we look at all the paths that take us from z to t. For each path we multiply all partials we find along the way with each other and we sum over all the paths. So we'd find, in this example:

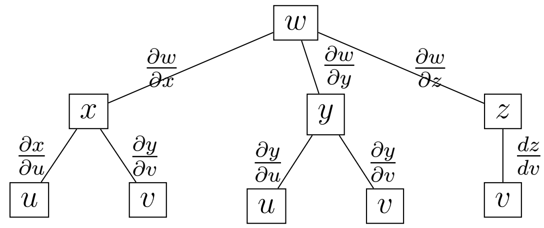
$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}$$

and the tree:



.

The trees can become arbitrarily large:



and from this tree we find, for example

$$\frac{\partial w}{\partial v} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial w}{\partial z}\frac{dz}{dv}.$$

Example 2. *Let $z = z(x, y)$, $x = x(r, \theta)$, $y = y(r, \theta)$. Find $\frac{\partial^2 z}{\partial r^2}$ (the second partial derivative with respect to r).*

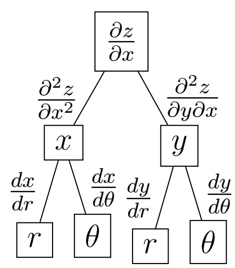
Solution. We apply the chain rule:

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

Then

$$\frac{\partial^2 z}{\partial r^2} = \frac{\partial}{\partial r}\left(\frac{\partial z}{\partial r}\right) = \frac{\partial}{\partial r}\left(\frac{\partial z}{\partial x}\right)\frac{\partial x}{\partial r} + \frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial r^2} + \frac{\partial}{\partial r}\left(\frac{\partial z}{\partial y}\right)\frac{\partial y}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial r^2}.$$

The problematic terms here are $\frac{\partial}{\partial r}\left(\frac{\partial z}{\partial x}\right)$ and $\frac{\partial}{\partial r}\left(\frac{\partial z}{\partial y}\right)$. To compute the first one, we write a tree diagram for $\frac{\partial z}{\partial x}$.



and so

$$\frac{\partial}{\partial r}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial x^2}\frac{dx}{dr} + \frac{\partial^2 z}{\partial y \partial x}\frac{dy}{dr}.$$

The second one is left as an exercise. □

*Implicit differentiation

We will look at level sets of a differentiable function, that is, sets of the form

$$S = \{(x, y, z) : F(x, y, z) = c\}.$$

A deep theorem of advanced calculus called the Implicit Function Theorem guarantees that if $p = (x_0, y_0, z_0) \in S$ and $\frac{\partial F}{\partial z}(x_0, y_0, z_0) \neq 0$ then in a neighborhood of p we can write $z = z(x, y)$ (that is, z can be written as a function of x, y).

Example 3. *The unit sphere can be thought of as the level set of the function*

$$F(x,y,z) = x^2 + y^2 + z^2,$$

and $p = (0,0,1)$ *satisfies* $\frac{\partial F}{\partial z}(0,0,1) \neq 0$ *and indeed near* p *we can write* $z = \sqrt{1-x^2-y^2}$*. At* $q = (0,1,0)$ *however,* $\frac{\partial F}{\partial z}(0,1,0) = 0$ *and we can't write* $z = z(x,y)$ *for* (x,y) *near* $(0,1)$*: a choice of root needs to be made.*

This theorem is of huge theoretical importance, though not so much of practical: we can't generally recover this function explicitly!

However, we can recover its partial derivatives using the chain rule.

$$\begin{aligned} F(x,y,z) = const. &\implies \frac{\partial F}{\partial x} = 0 \\ &\implies \frac{\partial F}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0 \\ &\implies \frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} \end{aligned}$$

and similarly

$$\frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}}$$

From the partial derivatives we can now use the formula

$$z = z(x_0,y_0) + \frac{\partial z}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial z}{\partial y}(x_0,y_0)(y-y_0)$$

to find the tangent plane to the surface at p and hence approximate the level set near p by something easy to work with - a plane!

Example 4. *Let* $F(x,y,z) = y^2z^3 + 3zy + 2xz + 2$*. Then if* S *is the level set where* $F = 0$*,* $(-6,2,1) \in S$*. Also,*

$$\partial_z F = 3z^2y^2 + 3y + 2x \implies \partial_z F(-6,2,1) = 6 \neq 0.$$

So, near (-6,2,1), $z = z(x,y)$*. Since* $\partial_x F = 2z$*, we find*

$$\frac{\partial z}{\partial x}(-6,2,1) = -1/3.$$

Similarly,

$$\frac{\partial z}{\partial y}(-6,2,1) = -7/6$$

and the tangent plane at (-6,2,1) becomes

$$z = 1 - \frac{1}{3}(x+6) - \frac{7}{6}(y-2).$$

In the next section we will see an easier way to find the tangent plane of a level set of a differentiable function.