14.5: The Chain Rule

Friday, November 25, 2022 2:48 PM

The Chain Rule

❖ **The Chain Rule**

From single variable calculus we learned that if y is a differentiable function of x, that is $y = f(x)$, and x is a differentiable function of t, that is $x = g(t)$, then the chain rule for functions of one variable states that: $\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$ (Leibniz notation).

variable states

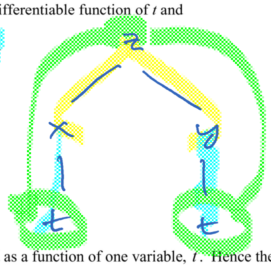
Another way we learned this was: If $F(x) = f(g(x))$ then $F'(x) = f'(g(x))g'(x)$

We can extend this idea to functions in two variables:

<u>**The Chain Rule (Case 1)**</u>: Suppose that $z = f(x, y)$ is a differentiable function of *x* and *y*, where $x = g(t)$ and $y = h(t)$ are both differentiable functions of *t*. Then *z* is a differentiable function of *t* and

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

The proof of this is in the book in case you are interested.



Note that $z = f(g(t), h(t))$, which means z can be expressed as a function of one variable, t. Hence the notation $\frac{dz}{dt}$. 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.

Other ways to write the multivariate chain rule are:

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

$$\frac{df}{dt} = f_x\frac{dx}{dt} + f_y\frac{dy}{dt}$$

$$\frac{df}{dt} = f_x x'(t) + f_y y'(t)$$

Example 1: Suppose that $z = x^2 y$, $x = t^2$ and $y = t^3$

a) Use the chain rule to find $\frac{dz}{dt}$.

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}$$
$$= 2\cancel{x}y \cdot 2t + \cancel{x}^2 \, 3t^2$$
$$= 2t^2 t^3 \cdot 2t + (t^2)^2 \cdot 3t^2$$

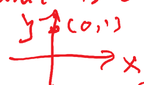
b) Can we find $\frac{dz}{dt}$ an easier way?

$$= 4t^6 + 3t^6 = 7t^6$$

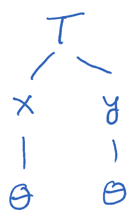
$z = x^2 y$
$= (t^2)^2 t^3$ $\Rightarrow z = t^7$
$\Rightarrow \frac{dz}{dt} = 7t^6$

The method in (b) is not convenient as the number of variables increases!

Example 2: Suppose you have a parametric curve C, defined by $x = \cos\theta$ and $y = \sin\theta$. Where θ represents time and $T(x, y) = \sqrt{xy + y}$ gives the temperature in the xy-plane. Find the rate of change of temperature with respect to time as we move along C through $(x, y) = (0,1)$. $\Leftarrow$ what is θ



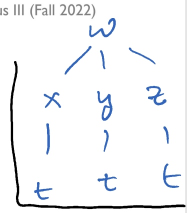
$\theta = \pi/2$



$$\frac{dT}{d\theta} = \frac{\partial T}{\partial x}\frac{dx}{d\theta} + \frac{\partial T}{\partial y}\frac{dy}{d\theta}$$
$$= \frac{y}{2\sqrt{xy+y}}(-\sin\theta) + \frac{x+1}{2\sqrt{xy+y}}\cos\theta$$
$$= \frac{-\sin^2\theta + \cos^2\theta + \cos\theta}{2\sqrt{\sin\theta\cos\theta + \sin\theta}}$$
$\Big|_{\theta = \frac{\pi}{2}}$
$$= \frac{-1 + 0 + 0}{2\sqrt{1(0) + 1}}$$
$$= -\frac{1}{2} \text{ (Rate of change @ } (0,1)\text{)}.$$

The same ideas hold in higher dimensional spaces.

Example 3: Suppose that $w = x^3y^2z$, $x = t^2$, $y = e^t$ and $z = \ln t$. Find $\frac{dw}{dt}$.



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

$$= 3x^2y^2z \cdot 2t + 2x^3yz\, e^t + x^3y^2\frac{1}{t}$$

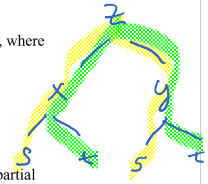
$$= 3(t^2)^2(e^t)^2 \cdot \ln t(2t) + 2(t^2)^3 e^t \ln t\, e^t + (t^2)^3(e^t)^2 \cdot \frac{1}{t}$$

$$= 6t^5e^{2t}\ln t + 2t^6e^{2t}\ln t + t^5e^{2t}$$

What if x and y are multivariable functions?

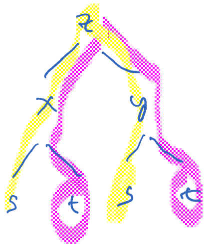
The Chain Rule (Case 2): Suppose that $z = f(x,y)$ is a differentiable function of x and y, where $x = g(s,t)$ and $y = h(s,t)$ are both differentiable functions of s and t. Then

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \quad \text{and} \quad \frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}$$



Note that $z = f(g(s,t), h(s,t))$, is a function of s and t. Here we have two first order partial derivatives.

Example 4: Find the first partial derivatives of $z = e^{2x}\sin(3y)$ where $x = st - t^2$ and $y = \sqrt{s^2+t^2}$.



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

$$= 2e^{2x}\sin(y)\, t + 3e^{2x}\cos(3y)\frac{s}{\sqrt{s^2+t^2}}$$

correction at 22:52 in video

$$= e^{2(st-t^2)}\left[2t\sin(y) + \frac{3\cos\left(3\sqrt{s^2+t^2}\right)s}{\sqrt{s^2+t^2}}\right]$$

correction at 24:10 in video

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

$$= 2e^{2x}\sin(y)(s-2t) + 3e^{2x}\cos(3y)\frac{t}{\sqrt{s^2+t^2}}$$

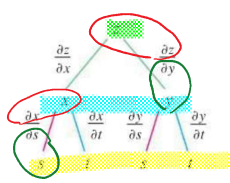
correction at 27:30 in video.

$$= e^{2(st-t^2)}\left[2\sin(y)(s-2t) + \frac{3\cos\left(3\sqrt{s^2+t^2}\right)\cdot t}{\sqrt{s^2+t^2}}\right]$$

14.5: The Multivariate Chain Rule
Math 163: Calculus III (Fall 2022)

Case 2 of the Chain Rule contains three types of variables: s and t are **independent** variables, x and y are called **intermediate** variables, and z is the **dependent** variable.

To remember the chain rule, it's helpful to draw a **tree diagram**. We start at the top with the function itself and then branch out from that point. The first set of branches is for the intermediate variables in the function. From each of these endpoints we put down a further set of branches that gives independent variables. 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.

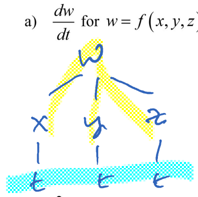


To use this to get the chain rule we start at the bottom and for each branch that ends with the variable we want to take the derivative with respect to (s in this case) we move up the tree until we hit the top multiplying the derivatives that we see along that set of branches. Once we've done this for 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 usually put the derivatives in the tree. They are always an assumed part of the tree.

Example 5: Use a tree diagram to write down the chain rule for the given derivatives.

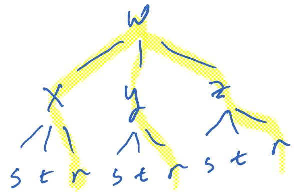
a) $\frac{dw}{dt}$ for $w = f(x,y,z)$ where $x = g(t)$, $y = h(t)$ and $z = p(t)$



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

b) $\frac{\partial w}{\partial r}$ for $w = f(x,y,z)$ where $x = g(s,t,r)$, $y = h(s,t,r)$ and $z = p(s,t,r)$

b) $\frac{\partial w}{\partial r}$ for $w = f(x,y,z)$ where $x = g(s,t,r)$, $y = h(s,t,r)$ and $z = p(s,t,r)$



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

So, provided we can write down the tree diagram (easy), the chain rule can be applied to most any situation we run across.

8

The Chain Rule (General Case): Suppose that $z = f(x_1, x_2, \ldots, x_n)$ is a differentiable function of the n variables $x_1, x_2, \ldots, x_n$ and each x_j is a differentiable function of m variables $t_1, t_2, \ldots, t_m$. hence

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

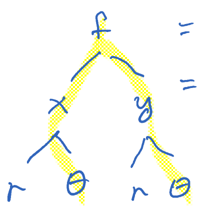
for each $i = 1, 2, \ldots, m$

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

$\boxed{\text{1st}}$

$$\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta}$$
$$= f_x(-r\sin\theta) + f_y\, r\cos\theta$$
$$= r(f_y\cos\theta - f_x\sin\theta)$$



$\boxed{\text{2nd}}$

$$\frac{\partial^2 f}{\partial \theta^2} = \frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial \theta}\right) = \frac{\partial}{\partial \theta}\left(r f_y\cos\theta - r f_x\sin\theta\right)$$
$$= \frac{\partial}{\partial \theta} f_y\,(r\cdot -\sin\theta) - \frac{\partial}{\partial \theta} f_x\, r\cos\theta\,)$$
$$= -r\sin\theta\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial y}\right) - r\cos\theta\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right)$$
$$= -r\sin\theta\left[\frac{\partial^2 f}{\partial x\,\partial y}\cdot\frac{\partial x}{\partial \theta} + \frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial \theta}\right]$$
$$- r\cos\theta\left[\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial \theta} + \frac{\partial^2 f}{\partial y\,\partial x}\frac{\partial y}{\partial \theta}\right]$$

$r \quad \theta \quad r \quad \theta$

$$= -r\sin\theta\left[f_{yx}\,x_\theta + f_{yy}\,y_\theta\right] - r\cos\theta\left[f_{xx}\,x_\theta + f_{xy}\,y_\theta\right]$$

$$= -r\sin\theta\left[f_{yx}\,r(-\sin\theta) + f_{yy}\,r\cos\theta\right]$$
$$- r\cos\theta\left[f_{xx}(-r\sin\theta) + f_{xy}\,r\cos\theta\right]$$

$$= r^2\left[f_{yx}\sin^2\theta - f_{yy}\sin\theta\cos\theta + f_{xx}\sin\theta\cos\theta - f_{xy}\cos^2\theta\right]$$

$$= r^2\left[f_{xy}\left(\sin^2\theta - \cos^2\theta\right) + \sin\theta\cos\theta\left(f_{xx} - f_{yy}\right)\right]$$

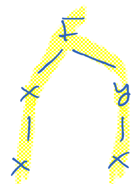
$$\frac{\partial^2 f}{\partial\theta^2} = r^2\left[-f_{xy}\cos(2\theta) + \tfrac{1}{2}\sin(2\theta)\left(f_{xx} - f_{yy}\right)\right]$$

History: The symbol used to denote partial derivatives is ∂. One of the first known uses of this symbol in mathematics is by Marquis de Condorcet from 1770, who used it for partial differences. The modern partial derivative notation was created by Adrien-Marie Legendre (1786), although he later abandoned it; Carl Gustav Jacob Jacobi reintroduced the symbol in 1841.

❖ ***Implicit Differentiation***

The final topic in this section is to revisit 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 $y=f(x)$. In single variable calculus, we were asked to find $\frac{dy}{dx}$ 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 of $F(x,y)=0$ with respect to x. This will mean using the chain rule:



$F_x\frac{dx}{dx}+F_y\frac{dy}{dx}=0$ Note that $\frac{dx}{dx}=1$ so we will have

$F_x+F_y\frac{dy}{dx}=0$, solving for $\frac{dy}{dx}$:

$\frac{dy}{dx}=-\frac{F_x}{F_y}$ which can also be written as $\frac{dy}{dx}=-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}$

Example 7: Find $\frac{dy}{dx}$ for $x\cos(3y)+x^3y^5=3x-e^{xy}$.

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

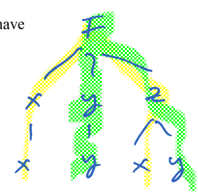
$\Rightarrow \frac{dy}{dx}=-\frac{F_x}{F_y}=-\frac{\cos(3y)+3x^2y^5-3+ye^{xy}}{-3x\sin(3y)+5x^3y^4+xe^{xy}}$

If we have $F(x,y,z)=0$, with $z=f(x,y)$ then we can find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ using the chain rule:

$F_x\frac{\partial x}{\partial x}+F_y\frac{\partial y}{\partial x}+F_z\frac{\partial z}{\partial x}=0$ Note that $\frac{\partial x}{\partial x}=1$ and $\frac{\partial y}{\partial x}=0$ so we will have

$F_x+F_z\frac{\partial z}{\partial x}=0$, solving for $\frac{\partial z}{\partial x}$:

$\frac{\partial z}{\partial x}=-\frac{F_x}{F_z}$ and similarly $\frac{\partial z}{\partial y}=-\frac{F_y}{F_z}$



Example 8: Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for $x^2 \sin(2y-5z) = 1 + y\cos(6zx)$

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

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

And

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