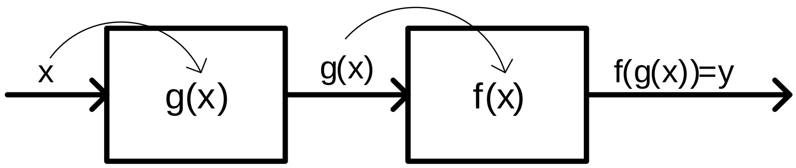
Multivariable Differentiation – The Chain Rule

In single variable calculus the chain rule is used to differentiate composite functions. The composition of functions is a very common occurrence in any practical system and can be graphically depicted as shown below.



As such, the chain rule is one of the most widely used rules in differential calculus. Therefore, in this lesson, we extend the chain rule for multivariable functions. Before we start, it's a good idea to recall the different notations used with the chain rule for single variable functions.

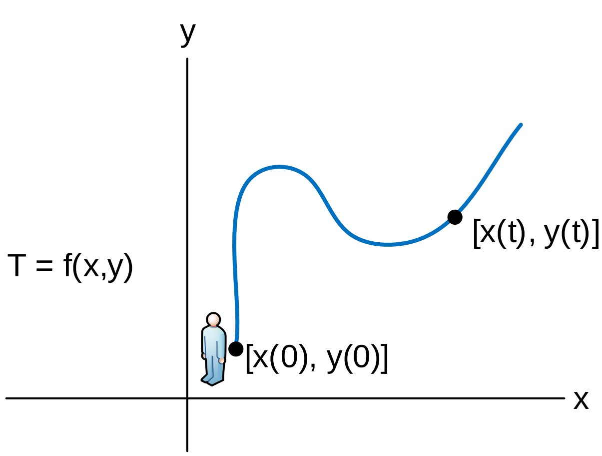
Lagrange Notation	***Leibniz Notation***
$\big(f(g(x))\big)' = f'(g(x))g'(x)$	$\dfrac{df}{dx} = \dfrac{df}{dg}\cdot\dfrac{dg}{dx}$

Although the Lagrange notation was mostly used in single variable calculus, the Leibniz notation is generally used for multivariable functions.

Multivariable Chain Rule

The multivariable chain rule can be expressed in many different forms depending on the number of variables in each of the functions involved in the composition. Nonetheless, there is a general rule that can be stated that is applicable to all cases. However, before stating this rule, we will derive a formula for a specific case. We will rely on the definition of differentials that we learned in the previous lesson.

As a visual example, imagine an area where the temperature varies as a function of the position and is represented as $T = f(x, y)$. Now let's assume that a person, with the ability to measure the temperature, is traveling along this surface, and that their position can be represented parametrically as $(x(t), y(t))$.



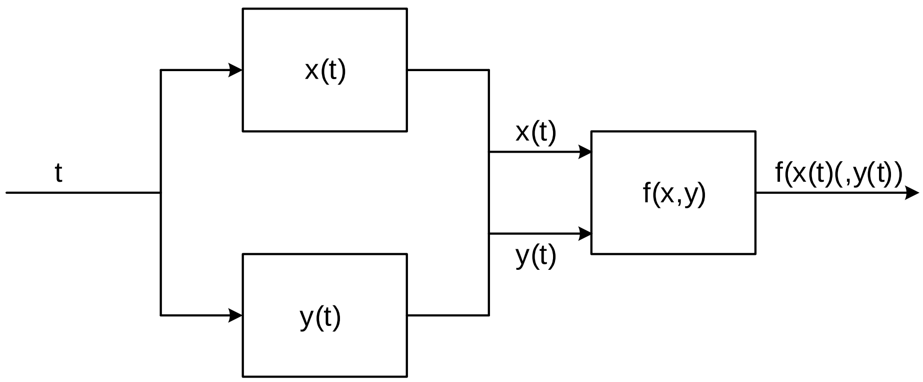
The question we would like to answer is: "how does the temperature change as a function of time as the person travels along their path?", i.e. $\frac{dT}{dt} = \frac{df}{dt}$. Recall, in the previous lesson we defined the differential of f, df, as follows:

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

Multiplying through by $1/dt$ we have

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

This expression represents the derivative of a composition of functions, where the outer function is a two variable function and the inner functions are of single variables. A diagram for this case, similar to the one above, is shown below.



Example 1: Find the derivative with respect to t of the function

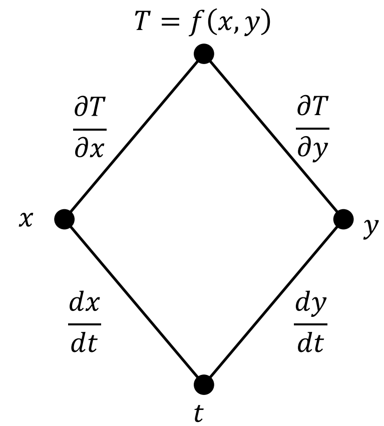
$$w(x, y) = 1 + x^2 - y^2$$

along the circular path $x(t) = cos(t), y(t) = sin(t)$.

Solution: Using the chain rule derived above we have

$$\begin{aligned}\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} \\ &= 2x(-sin(t)) - 2y(cos(t)) \\ &= -2\,cos(t)\,(sin(t)) - 2\,sin(t)\,(cos(t)) \\ &= -4\,cos(t)\,sin(t) \\ &= -2\,sin(2t)\end{aligned}$$

A convenient way to represent the composition of multivariable functions that also helps to derive the chain rule for all cases is referred to as a *tree diagram* and is shown below for the specific case above.



$$\frac{dT}{dt} = \frac{\partial T}{\partial x}\frac{dx}{dt} + \frac{\partial T}{\partial y}\frac{dy}{dt}$$

The dependent variable is placed at the top, the intermediate variables are placed in the middle, and independent variable is placed at the bottom. The corresponding derivatives are places along the lines joining the vertices. The final chain rule formula is computed by multiplying the derivatives along each path from the top of the tree to bottom and adding.

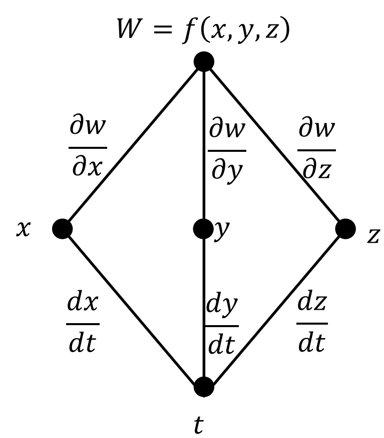
Before we state the general formula let's look at another example with three variables.

Example 2: Find the derivative with respect to t of the function

$$w(x, y, z) = \sqrt{x^2 + y^2 + z^2}$$

Where, $x(t) = cos(t), y(t) = sin(t)$, and $z(t) = t$

Solution: Let's draw the tree diagram first.



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

Computation is then as follows:

$$\begin{aligned}\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} \\ &= \frac{x}{\sqrt{x^2+y^2+z^2}}(-sin(t)) + \frac{y}{\sqrt{x^2+y^2+z^2}}(cos(t)) + \frac{z}{\sqrt{x^2+y^2+z^2}}(1) \\ &= -\cancel{\frac{cos(t)\,sin(t)}{\sqrt{x^2+y^2+z^2}}} + \cancel{\frac{sin(t)\,cos(t)}{\sqrt{x^2+y^2+z^2}}} + \frac{t}{\sqrt{x^2+y^2+z^2}} \\ &= \frac{t}{\sqrt{cos^2(t)+sin^2(t)+t^2}} \\ &= \frac{t}{\sqrt{1+t^2}}\end{aligned}$$

In all of the previous examples, the intermediate variables were functions of a single variable, e.g. t. The general form for the multivariate chain rule includes the cases where the intermediate variables are functions of more than one variable. The general rule is stated below.

Multivariable Chain Rule
Let $f(x_1,...,x_n)$ be a differentiable function of n variables. Suppose that each of the variables, $x_1,...,x_n$, is a differentiable function of m independent variables, $t_1,...,t_m$. Then for $k = 1,\dots,m$ $$\frac{\partial f}{\partial t_k} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_k} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_k} + \dots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_k}$$ Note: Since x_i is assumed to be a function of more than one variable, the partial derivative notation is required, $\frac{\partial x_i}{\partial t_k}$. If $m = 1$ then $\frac{dx_i}{dt}$ could be used.

Example 3: Find $\frac{dz}{dt}$ by using the chain rule. Check your answer by expressing z as a function of t and then differentiating.

$$z = x\,sin(y)\,, x = e^t, y = \pi t$$

Solution: Since the intermediate variables are functions of a single variable, then the chain rule is identical to the example from above. Using the multivariate chain rule, we have

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \\ &= sin(y)\,(e^t) + x\,cos(y)\,(\pi) \\ &= e^t sin(\pi t) + \pi e^t\,cos(\pi t) \\ &= e^t(sin(\pi t) + \pi\,cos(\pi t))\end{aligned}$$

Next, we express z as a function of t and use single variable differentiation to verify.

$$z = e^t \sin(\pi t)$$

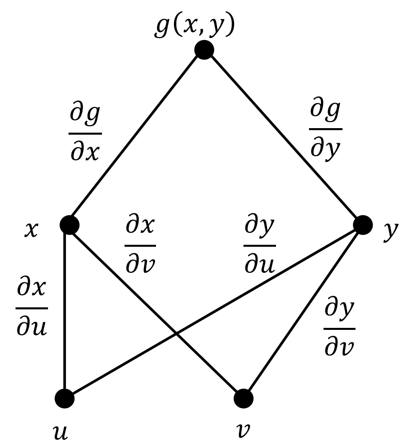
Using the product rule, we have

$$\frac{dz}{dt} = e^t \sin(\pi t) + \pi e^t \cos(\pi t)$$

Example 4: Compute $\frac{\partial g}{\partial u}$ and $\frac{\partial g}{\partial v}$ at the point $(u, v) = (0,1)$.

$$g = x^2 - y^2, x = e^u \cos(v), y = e^u \sin(v)$$

Solution: In this case the intermediate variables are function of more than one variable. To help we can draw the diagram as shown below.



For $\frac{\partial g}{\partial u}$ we proceed along all paths that lead from g to u, multiplying derivatives along the way and then summing the outputs from all paths.

$$\begin{aligned}\frac{\partial g}{\partial u} &= \frac{\partial g}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial u} \\ &= 2x(e^u \cos(v)) - 2y(e^u \sin(v)) \\ &= 2e^u(x\cos(v) - y\sin(v)) \\ &= 2e^u(e^u \cos(v)\cos(v) - e^u \sin(v)\sin(v)) \\ &= 2e^{2u}(\cos^2(v) - \sin^2(v)) \\ &= 2e^{2u}\cos(2v)\end{aligned}$$

Evaluating we have,

$$\left.\frac{\partial g}{\partial u}\right|_{(0,1)} = 2e^{2(0)}\cos\big(2(1)\big) = 2\cos(2)$$

Similarly, for $\frac{\partial g}{\partial v}$

$$\frac{\partial g}{\partial v} = \frac{\partial g}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial v}$$
$$= 2x(-e^u \sin(v)) - 2y(e^u \cos(v))$$
$$= -2e^u(x \sin(v) + y \cos(v))$$
$$= 2e^u(e^u \sin(v) \sin(v) + e^u \cos(v) \cos(v))$$
$$= 2e^{2u}(\sin^2(v) + \cos^2(v))$$
$$= 2e^{2u}$$

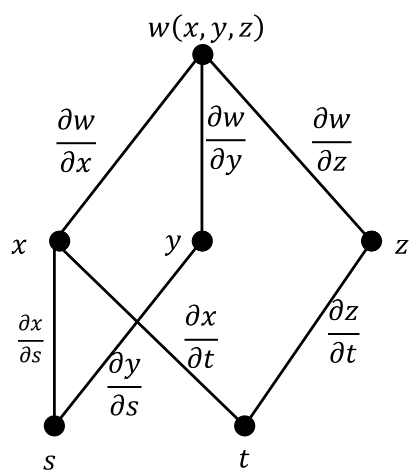
Evaluating we have,

$$\left.\frac{\partial g}{\partial v}\right|_{(0,1)} = 2e^{2(0)} = 2$$

Example 5: Find $\frac{\partial w}{\partial s}$ and $\frac{\partial w}{\partial t}$ by using the appropriate chain rule.

$$w = xy \sin(z^2), x = s - t, y = s^2, z = t^2$$

Solution: The tree diagram for this case is shown below.



Based on the diagram the partials are written as follows:

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s}$$
$$= y \sin(z^2)(1) + x \sin(z^2)(2s)$$
$$= \sin(z^2)(y + 2xs)$$
$$= \sin(t^4)\left(s^2 + 2s(s - t)\right)$$
$$= \sin(t^4)(3s^2 - 2st)$$

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t}$$
$$= y \sin(z^2)(-1) + x \sin(z^2)(2t)$$
$$= \sin(z^2)(2tx - y)$$
$$= \sin(t^4)(2t(s - t) - s^2)$$
$$= \sin(t^4)(2ts - 2t^2 - s^2)$$

Example 6: Suppose $J = f(x, y, z, w)$, where , $x = x(r, s, t), y = y(r, t), z = z(r, s)$ and $w = w(s, t)$. Use the chain rule to find $\frac{\partial J}{\partial r}, \frac{\partial J}{\partial s}$, and $\frac{\partial J}{\partial t}$.

Solution: Instead of drawing a tree diagram we'll develop the expression from the general form of the multivariable chain rule, shown below.

$$\frac{\partial f}{\partial t_k} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_k} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_k} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_k}$$

In our case $k = 3$, i.e. r, s, and t, however not all intermediate variables are a function of all three. For example, y is a function of r and t only. To illustrate this point we will initially write the expression with all possible combinations and then show which terms vanish.

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

In this case since w is not a function of r, $\frac{\partial w}{\partial r} = 0$. Therefore, the term vanishes, and we have

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

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

In this case since y is not a function of s, $\frac{\partial y}{\partial s} = 0$. Therefore, the term vanishes, and we have

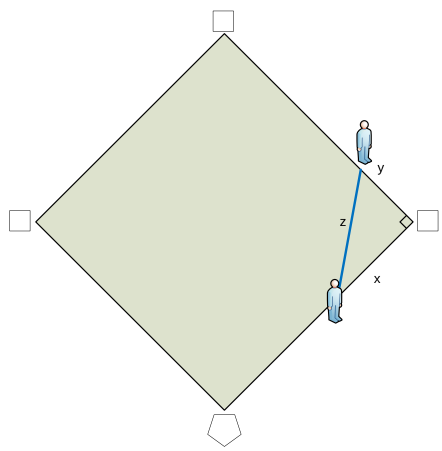
$$\boxed{\frac{\partial J}{\partial s} = \frac{\partial J}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial J}{\partial z}\frac{\partial z}{\partial s} + \frac{\partial J}{\partial w}\frac{\partial w}{\partial s}}$$

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

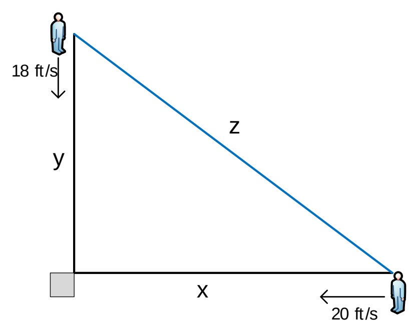
In this case since z is not a function of t, $\frac{\partial z}{\partial t} = 0$. Therefore, the term vanishes, and we have

$$\boxed{\frac{\partial J}{\partial t} = \frac{\partial J}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial J}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial J}{\partial w}\frac{\partial w}{\partial t}}$$

Example 7: A baseball player hits the ball and then runs down the first base line at $20\ ft/s$. The first baseman fields the ball and then runs toward first base along the second base line at $18\ ft/s$ as shown in the figure below. Determine how fast the distance between the two players is changing when the hitter is $8\ ft$ from the first base and the first baseman is $6\ ft$ from the first base.



Solution: Since the 1st base is at a right angle from the two paths we can redraw the scenario as shown below.



Now we can use the Pythagorean Theorem to write z as a function of x and y.

$$z = f(x, y) = \sqrt{x^2 + y^2}$$

The distances of the players, x and y, can be written as functions of time.

$$x(t) = 8 - 20t \qquad\qquad y(t) = 6 - 18t$$

The goal is to find the rate of change of the distance, z, between the players, i.e. $\frac{dz}{dt}$. Using the multivariable chain rule, we have

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \\ &= \frac{x}{\sqrt{x^2 + y^2}}(-20) + \frac{y}{\sqrt{x^2 + y^2}}(-18) \\ &= -\frac{(20x + 18y)}{\sqrt{x^2 + y^2}}\end{aligned}$$

Finally, we evaluate the derivative at $x = 8$ and $y = 6$.

$$\left.\frac{dz}{dt}\right|_{x=8,y=6} = -\frac{(20\cdot 8 + 18\cdot 6)}{\sqrt{64+36}} = -26.8\ ft/s$$

This problem is similar to the types of problems we encountered in calculus 1 when studying what we called *related rates*. Let's redo the problem using that method. Using the calculus 1 methodology we start by writing the with the Pythagorean Theorem as follows

$$z^2(t) = x^2(t) + y^2(t)$$

Next, we use implicit differentiation to differentiate the entire equation with respect to t.

$$\frac{d}{dt}\left(z^2(t)\right) = \frac{d}{dt}\left(x^2(t)\right) + \frac{d}{dt}\left(y^2(t)\right)$$

$$2z\frac{dz}{dt} = 2x\frac{dx}{dt} + 2y\frac{dy}{dt}$$

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

$$\frac{dz}{dt} = \frac{x(-20) + y(-18)}{\sqrt{x^2+y^2}}$$

$$\frac{dz}{dt} = -\frac{(20x+18y)}{\sqrt{x^2+y^2}}$$

Which is identical to the expression from above!

Implicit Differentiation

Example 7 leads us into our final topic for this lesson; *implicit differentiation for multivariable functions*. In single variable calculus, when presented with an implicit function for which we wanted to find dy/dx, we treated the variable y as a function of x and differentiated the entire equation with respect to x. Let's look at an example below.

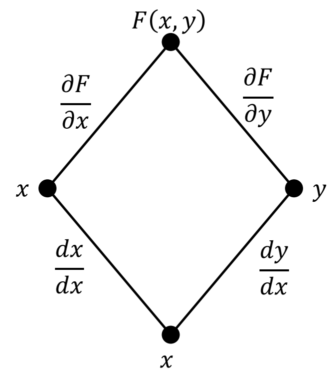
Example 8: Using implicit differentiation from single variable calculus find dy/dx for the following implicit function.

$$3x^2 - 9y^3 = 20$$

Solution: We differentiate the entire equation with respect to x. Since we assume y is a function of x we use the chain rule for the second term.

$$\frac{d}{dx}(3x^2) - \frac{d}{dx}(9y^3) = \frac{d}{dx}(20)$$
$$6x - 27y^2\frac{dy}{dx} = 0$$
$$\frac{dy}{dx} = \frac{2x}{9y^2}$$

The downside of this method is that, even after the derivatives are computed, it requires a fair amount of algebraic manipulation to arrive at the formula for dy/dx. We can reduce the workload by first rewriting the equation as a multivariable function of the form: $F(x, y) = 0$, and then use the multivariable chain rule. To start we assume that the equation defines y as a function of x, i.e. we can write $y = f(x)$. Additionally, x is a function of x and so we can consider x and y as intermediate variables of $F(x, y)$. The tree diagram is shown below.



Using the multivariable chain rule on $F(x, y) = 0$ we find

$$\frac{\partial}{\partial x}\big(F(x, y)\big) = \frac{\partial}{\partial x}(0)$$
$$\frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0$$

Where, $\frac{dx}{dx} = 1$. Using the alternate notation of $\frac{\partial F}{\partial x} = F_x$ and $\frac{\partial F}{\partial y} = F_y$, we can easily solve for $\frac{dy}{dx}$.

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

To illustrate let's redo example 8 using this method. We start by rewriting the equation as

$$3x^2 - 9y^3 - 20 = 0$$

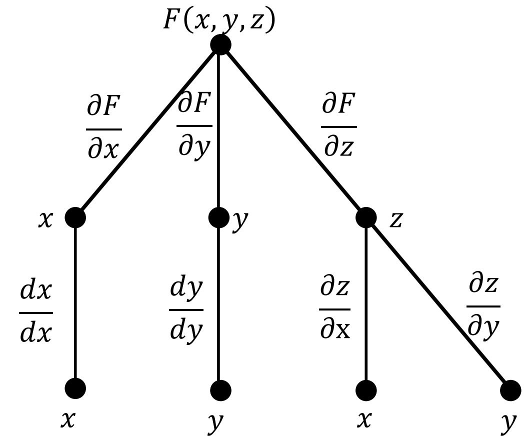
Next, we compute the partials.

$$F_x = 6x \qquad\qquad F_y = -27y^2$$

Direct substitution into the formula above yields the same answer as before.

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\left(\frac{6x}{-27y^2}\right) = \frac{2x}{9y^2}$$

We can derive similar formulas for functions of three variables. Suppose we are given a function which we can write in the form, $F(x, y, z) = 0$. The tree diagram is then drawn assuming z as a function of x and y, x as a function of x, and y as a function of y.



In this case, using the multivariable chain rule for $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ on $F(x, y, z) = 0$ we find.

$$\frac{\partial}{\partial x}\big(F(x,y,z)\big) = \frac{\partial}{\partial x}(0)$$
$$\frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0$$
$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}$$

$$\frac{\partial}{\partial y}\big(F(x,y,z)\big) = \frac{\partial}{\partial y}(0)$$
$$\frac{\partial F}{\partial y}\frac{dy}{dy} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial y} = 0$$
$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}$$

Multivariable Implicit Differentiation
Suppose we have the equation $F(x, y) = 0$, and that $F(x, y)$ is differentiable. Then $$\frac{dy}{dx} = -\frac{F_x}{F_y}$$ Provided $F_y \neq 0$
Suppose we have the equation $F(x, y, z) = 0$, and that $F(x, y, z)$ is differentiable. Then $$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}$$ *and* $$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}$$ Provided $F_z \neq 0$

Example 9: Calculate the partial derivatives using implicit differentiation.

a. $\frac{\partial z}{\partial x}$, $x^2y + y^2z + xz^2 = 10$ *b.* $\frac{\partial z}{\partial y}$, $e^{xy} + sin(xz) + y = 0$

c. $\frac{\partial w}{\partial y}$, $\frac{1}{w^2+x^2} + \frac{1}{w^2+y^2} = 1 \; at \; (x, y, w) = (1, 1, 1)$

Solution:

a. We start by rewriting the equation as follows:

$$F(x, y, z) = x^2y + y^2z + xz^2 - 10 = 0$$

Then, using the formulas developed above we have

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z} = -\frac{2xy + z^2}{y^2 + 2xz}$$

b. In this case we have

$$F(x, y, z) = e^{xy} + sin(xz) + y = 0$$

$$\frac{\partial z}{\partial y} = -\frac{F_y}{F_z} = -\frac{xe^{xy} + 1}{xcos(xz)}$$

c. In this case we are asked to evaluate the partial at $(x, y, w) = (1, 1, 1)$

$$F(x, y, w) = \frac{1}{w^2 + x^2} + \frac{1}{w^2 + y^2} - 1 = 0$$

$$\frac{\partial w}{\partial y} = -\frac{F_y}{F_w} = -\frac{\frac{-2y}{(w^2 + y^2)^2}}{\frac{-2w}{(w^2 + x^2)^2} + \frac{-2w}{(w^2 + y)^2}} = -\frac{\frac{y}{(w^2 + y^2)^2}}{\frac{w}{(w^2 + x^2)^2} + \frac{w}{(w^2 + y)^2}}$$

Evaluating at $(x, y, w) = (1, 1, 1)$ we have

$$\left.\frac{\partial w}{\partial y}\right|_{(1,1,1)} = -\frac{\frac{1}{(1^2 + 1^2)^2}}{\frac{1}{(1^2 + 1^2)^2} + \frac{1}{(1 + 1)^2}} = -\frac{1}{2}$$

Final Summary for Multivariable Differentiation – The Chain Rule

Multivariable Chain Rule

Let $f(x_1, \ldots, x_n)$ be a differentiable function of n variables. Suppose that each of the variables, $x_1, \ldots, x_n$, is a differentiable function of m independent variables, $t_1, \ldots, t_m$.
Then for $k = 1, \ldots, m$

$$\frac{\partial f}{\partial t_k} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_k} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_k} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_k}$$

Note: Since x_i is assumed to be a function of more than one variable, the partial derivative notation is required, $\frac{\partial x_i}{\partial t_k}$. If $m = 1$ then $\frac{dx_i}{dt}$ could be used.

Multivariable Implicit Differentiation

Suppose we have the equation $F(x, y) = 0$, and that $F(x, y)$ is differentiable. Then

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

Provided $F_y \neq 0$

Suppose we have the equation $F(x, y, z) = 0$, and that $F(x, y, z)$ is differentiable. Then

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

Provided $F_z \neq 0$

By: ferrantetutoring