14.5 The Chain Rule

The Chain Rule

We know that the Chain Rule for functions of a single variable gives the rule for differentiating a composite function: If $y = f(x)$ and $x = g(t)$, where f and g are differentiable functions, then y is indirectly a differentiable function of t and

$$\boxed{1} \qquad \frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$$

For functions of more than one variable, the Chain Rule has several versions, each of them giving a rule for differentiating a composite function.

The Chain Rule

The first version (Theorem 2) deals with the case where $z = f(x, y)$ and each of the variables x and y is, in turn, a function of a variable t.

This means that z is indirectly a function of t, $z = f(g(t), h(t))$, and the Chain Rule gives a formula for differentiating z as a function of t. We assume that f is differentiable.

The Chain Rule

We know that this is the case when f_x and f_y are continuous.

2 The Chain Rule (Case 1) 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

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

Since we often write $\partial z/\partial x$ in place of $\partial f/\partial x$, we can rewrite the Chain Rule in the form

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

Example 1

If $z = x^2y + 3xy^4$, where $x = \sin 2t$ and $y = \cos t$, find dz/dt when $t = 0$.

Solution:

The Chain Rule gives

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

$$= (2xy + 3y^4)(2\cos 2t) + (x^2 + 12xy^3)(-\sin t)$$

It's not necessary to substitute the expressions for x and y in terms of t.

Example 1 – *Solution*

We simply observe that when t = 0, we have x = sin 0 = 0 and y = cos 0 = 1.

Therefore

$$\left.\frac{dz}{dt}\right|_{t=0} = (0 + 3)(2\cos 0) + (0 + 0)(-\sin 0) = 6$$

The Chain Rule

We now consider the situation where $z = f(x, y)$ but each of x and y is a function of two variables s and t:
$x = g(s, t)$, $y = h(s, t)$.

Then z is indirectly a function of s and t and we wish to find $\partial z/\partial s$ and $\partial z/\partial t$.

We know that in computing $\partial z/\partial t$ we hold s fixed and compute the ordinary derivative of z with respect to t.

Therefore we can apply Theorem 2 to obtain

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

The Chain Rule

A similar argument holds for $\partial z/\partial s$ and so we have proved the following version of the Chain Rule.

3 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 differentiable functions of s and t. Then

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

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.

The Chain Rule

Notice that Theorem 3 has one term for each intermediate variable and each of these terms resembles the one-dimensional Chain Rule in Equation 1.

To remember the Chain Rule, it's helpful to draw the **tree diagram** in Figure 2.

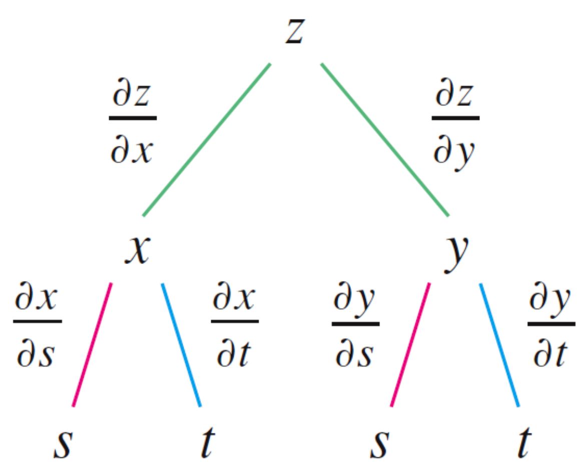

Figure 2

The Chain Rule

We draw branches from the dependent variable *z* to the intermediate variables *x* and *y* to indicate that *z* is a function of *x* and *y*. Then we draw branches from *x* and *y* to the independent variables *s* and *t*.

On each branch we write the corresponding partial derivative. To find $\partial z/\partial s$, we find the product of the partial derivatives along each path from *z* to *s* and then add these products:

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

The Chain Rule

Similarly, we find $\partial z/\partial t$ by using the paths from z to t.

Now we consider the general situation in which a dependent variable u is a function of n intermediate variables $x_1, \ldots, x_n$, each of which is, in turn, a function of m independent variables $t_1, \ldots, t_m$.

Notice that there are n terms, one for each intermediate variable. The proof is similar to that of Case 1.

The Chain Rule

4 The Chain Rule (General Version) Suppose that u is a differentiable function of the n variables $x_1, x_2, \ldots, x_n$ and each x_j is a differentiable function of the m variables $t_1, t_2, \ldots, t_m$. Then u is a function of $t_1, t_2, \ldots, t_m$ and

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

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

Implicit Differentiation

The Chain Rule can be used to give a more complete description of the process of implicit differentiation.

We suppose that an equation of the form $F(x, y) = 0$ defines y implicitly as a differentiable function of x, that is, $y = f(x)$, where $F(x, f(x)) = 0$ for all x in the domain of f.

If F is differentiable, we can apply Case 1 of the Chain Rule to differentiate both sides of the equation $F(x, y) = 0$ with respect to x.

Since both x and y are functions of x, we obtain

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

Implicit Differentiation

But $dx/dx = 1$, so if $\partial F/\partial x \neq 0$ we solve for dy/dx and obtain

6

$$\frac{dy}{dx} = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial y}} = -\frac{F_x}{F_y}$$

To derive this equation we assumed that $F(x, y) = 0$ defines y implicitly as a function of x.

Implicit Differentiation

The **Implicit Function Theorem**, proved in advanced calculus, gives conditions under which this assumption is valid: it states that if F is defined on a disk containing (a, b), where $F(a, b) = 0$, $F_y(a, b) \neq 0$, and F_x and F_y are continuous on the disk, then the equation $F(x, y) = 0$ defines y as a function of x near the point (a, b) and the derivative of this function is given by Equation 6.

Example 8

Find y' if $x^3 + y^3 = 6xy$.

Solution:

The given equation can be written as

$$F(x, y) = x^3 + y^3 - 6xy = 0$$

so Equation 6 gives

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

$$= -\frac{3x^2 - 6y}{3y^2 - 6x} = -\frac{x^2 - 2y}{y^2 - 2x}$$

Implicit Differentiation

Now we suppose that z is given implicitly as a function $z = f(x, y)$ by an equation of the form $F(x, y, z) = 0$.

This means that $F(x, y, f(x, y)) = 0$ for all (x, y) in the domain of f. If F and f are differentiable, then we can use the Chain Rule to differentiate the equation $F(x, y, z) = 0$ as follows:

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

Implicit Differentiation

But $\dfrac{\partial}{\partial x}(x) = 1$ and $\dfrac{\partial}{\partial x}(y) = 0$

so this equation becomes

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

If $\partial F/\partial z \neq 0$, we solve for $\partial z/\partial x$ and obtain the first formula in Equations 7.

The formula for $\partial z/\partial y$ is obtained in a similar manner.

Implicit Differentiation

7 $$\frac{\partial z}{\partial x} = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial z}} \qquad \frac{\partial z}{\partial y} = -\frac{\dfrac{\partial F}{\partial y}}{\dfrac{\partial F}{\partial z}}$$

Again, a version of the **Implicit Function Theorem** stipulates conditions under which our assumption is valid: if F is defined within a sphere containing (a, b, c), where $F(a, b, c) = 0$, $F_z(a, b, c) \neq 0$, and F_x, F_y, and F_z are continuous inside the sphere, then the equation $F(x, y, z) = 0$ defines z as a function of x and y near the point (a, b, c) and this function is differentiable, with partial derivatives given by (7).