The Chain Rule and Implicit Function Theorems

1 The Chain Rule for Functions of Several Variables

First recall the Chain Rule for functions of one variable.

Chain Rule for Functions of One Variable. *Let f be differentiable at a and let g be differentiable at $f(a)$. Then $g \circ f$ is differentiable at a and $(g \circ f)'(a) = g'(f(a))f'(a)$.*

The reader should be familiar with the concept of composition for functions of one variable. For example the function $\arcsin(x^3)$ is the composition of the function $f(x) = x^3$ by the function $g(x) = \arcsin x$. Notice that the composition of g by f is a different function; namely, $(f \circ g)(x) = \arcsin^3 x$. We employ the Chain Rule to determine $\frac{d}{dx}\arcsin(x^3)$. Using the notation of the preceding example $g'(x) = \frac{1}{\sqrt{1-x^2}}$ and $f'(x) = 3x^2$. By the Chain Rule

$$\frac{d}{dx}\arcsin(x^3) = g'(x^3)f'(x) = \frac{3x^2}{\sqrt{1-(x^3)^2}}.$$

Similarly

$$\frac{d}{dx}\arcsin^3 x = f'(\arcsin x)g'(x) = (3\arcsin^2 x)\left(\frac{1}{\sqrt{1-x^2}}\right).$$

The Chain Rule for functions of several variables begins with a discussion of the type of functions to which the rule applies. The outer function g should be a function of several variables; $g : E \subset \mathbb{R}^k \to \mathbb{R}$. Consequently the range of the inner function must be $\mathbb{R}^n$ but in addition the inner function should be a function of several variables. Specifically let $F : D \subset \mathbb{R}^n \to \mathbb{R}^k$. When the range of a function is more than one dimensional, we introduce the so-called coordinate functions. (Recall studying $\mathbf{r}(t) = (x(t), y(t), z(t))$,) So for each $j = 1, 2, \ldots, k$ let $f_j : D \subset \mathbb{R}^n \to \mathbb{R}$ such that for each $P \in D$ we have $F(P) = (f_1(P), f_2(P), \ldots, f_k(P))$. Guided by our study of paths in $\mathbb{R}^n$, we will say that the function is continuous (differentiable) at P means that for each $j = 1, 2, \ldots, k$ the coordinate function, f_j is continuous (differentiable) at P. Finally the composition $g \circ F$ is obtained by replacing the j^{th} variable of g by the function f_j creating a new function $h = g \circ F$. Specifically $h(x_1, x_2, \ldots, x_n) = g(f_1(x_1, x_2, \ldots, x_n), f_2(x_1, x_2, \ldots, x_n), \ldots, f_k(x_1, x_2, \ldots, x_n))$. Then $h : D \subset \mathbb{R}^n \to \mathbb{R}$.

In the first example the outer function, g, will a function of two variables and each of these variables will be replaced by a function of one variable resulting in a function of one variable.

Example 1. *Let $g(x, y) = x^3 \arctan y$ Let $x = \sin t$ and let $y = t^2$. Then $h(t) = (\sin t)^3 \arctan t^2$.*

In the next example g will be a function of one variable and that one variable will be replaced by a function of three variables resulting in a function of three variables.

Example 2. *Let $g(x) = \ln x$ and let $x(u, v, w) = u^2 + vw$. Then $h(u, v, w) = \ln(u^2 + vw)$.*

In the third example, g is a function of three variable and each of these variables is replaced by a function of two variables. resulting in a function of two variables.

Example 3. *Let $g(u, v, w) = e^{u^2+v^2}\cos w$ and let $u(x, y) = x + y$, $v(x, y) = \ln(xy)$ and let $w(x, y) = \frac{x}{y}$. Then $h(x, y) = e^{(x+y)^2+(\ln(xy))^2}\cos(\frac{x}{y})$.*

The partial derivatives for all of the above examples can be computed my elementary techniques. But if all that is known is the partial derivatives of the outer function, standard techniques don't apply, but the Chain Rule does.

Chain Rule for Fucntions of Several Variables. *Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^k$ with coordinate functions $f_j : D \subset \mathbb{R}^n \to \mathbb{R}$ for $j = 1, 2, \ldots, k$ and let $g : E \subset \mathbb{R}^k \to \mathbb{R}$. Suppose F is differentiable at P_0, an interior point of D, (That is, f_j is differentiable at P_0 for each coordinate function f_j.), and that g is differentiable at $F(P_0) = (f_1(P_0), f_2(P_0), \ldots, f_k(P_0))$ an interior point of E. Then the function $h = g \circ F$ is differentiable at P_0 and for each $i = 1, 2, \ldots n$*

$$\frac{\partial h}{\partial x_i} = \frac{\partial g}{\partial y_1}\frac{\partial f_1}{\partial x_i} + \frac{\partial g}{\partial y_2}\frac{\partial f_2}{\partial x_i} + \cdots + \frac{\partial g}{\partial y_k}\frac{\partial f_k}{\partial x_i}.$$

It can be useful the think of the function h as

$$h(x_1, x_2, \ldots, x_n) = g(f_1(x_1, x_2, \ldots, x_n), f_2(x_1, x_2, \ldots, x_n), \ldots, f_k(x_1, x_2, \ldots, x_n).$$

Then the formula for a partial derivative of h can be interpreted as multiplying the partial derivative of g with respect to one of its variables by the partial derivative of the function replacing that variable with respect to the variable of differentiation and adding all such products together.

It is easy to overlook the first assertion of the theorem; namely, that if F is differentiable at P_0; that is, if each f_j is differentiable at P_0 and if g is differentiable at $F(P_0)$, then h is differentiable at P_0. Remember being able to compute the partial derivatives doesn't imply that the function is differentiable.

Example 4. *Let g, x and y be as in Example* **??**. *Then*

$$h'(t) = \frac{\partial g}{\partial x}x'(t) + \frac{\partial g}{\partial y}y'(t) = 3x^2 \arctan y \cos t + \frac{x^3}{1+y^2}2t = 3\sin^2 t \arctan(t^2) + \frac{\sin^3 t}{1+t^4}.$$

Example 5. *Let g and x be as in Example* **??**. *Then*

$$\begin{aligned} \frac{\partial h}{\partial u} &= g'(x)\frac{\partial x}{\partial u} = \frac{1}{x}2u = \frac{1}{u^2+vw}2u \\ \frac{\partial h}{\partial v} &= g'(x)\frac{\partial x}{\partial v} = \frac{1}{x}w = \frac{1}{u^2+vw}w \\ \textit{and } \frac{\partial h}{\partial w} &= g'(x)\frac{\partial x}{\partial w} = \frac{1}{x}u = \frac{1}{u^2+vw}u \end{aligned}$$

Example 6. *Let g, u, v, and w be as in Example* **??**. *Then*

$$\begin{aligned} \frac{\partial h}{\partial x} &= \frac{\partial g}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial g}{\partial w}\frac{\partial w}{\partial x} = (2ue^{u^2+v^2}\cos w)1 + (2ve^{u^2+v^2}\cos w)\frac{1}{x} + e^{u^2+v^2}(-\sin w)\frac{1}{y} \\ &= 2(x+y)e^{(x+y)^2+(\ln(xy))^2}\cos\big(\frac{x}{y}\big) + 2\ln(xy)e^{(x+y)^2+(\ln(xy))^2}\cos\big(\frac{x}{y}\big)\frac{1}{x} + e^{(x+y)^2+(\ln(xy))^2}\Big(-\sin\big(\frac{x}{y}\big)\Big)\frac{1}{y} \\ \frac{\partial h}{\partial y} &= \frac{\partial g}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial y} + \frac{\partial g}{\partial w}\frac{\partial w}{\partial y} = (2ue^{u^2+v^2}\cos w)1 + (2ve^{u^2+v^2}\cos w)\frac{1}{y} + e^{u^2+v^2}(-\sin w)(-\frac{x}{y^2}) \\ &= 2(x+y)e^{(x+y)^2+(\ln(xy))^2}\cos\big(\frac{x}{y}\big) + 2\ln(xy)e^{(x+y)^2+(\ln(xy))^2}\cos\big(\frac{x}{y}\big)\frac{1}{y} \\ &\qquad + e^{(x+y)^2(\ln(xy))^2}\Big(-\sin\big(\frac{x}{y}\big)\Big)(-\frac{x}{y^2}) \end{aligned}$$

Each of the formulas derived in each of the above examples can be verified by computing the corresponding derivative directly from the formula for h.

2 Implicit Function Theorems

Several of the problems in the text pertain to the Implicit Function Theorem. The theorem give conditions under which it is possible to solve an equation of the form $F(x, y) = 0$ for y as a function of x.

Implicit Function Theorem I. *Let $F : D \subset \mathbb{R}^2 \to \mathbb{R}$ and let (x_0, y_0) be an interior point of D with $F(x_0, y_0) = 0$. Suppose both first order partial derivatives of F exist in D and are continuous at (x_0, y_0) with $F_y(x_0, y_0) \neq 0$. Then there is an interval $I \subset \mathbb{R}$ with x_0 an interior point of I and a function $\phi : I \to \mathbb{R}$ such that ϕ is differentiable on I, $\phi(x_0) = y_0$ and $F(x, \phi(x)) = 0$ for each $x \in I$.*

The Chain Rule for Functions of Several Variables is then used to compute a formula for $\phi'(x)$. Because $F(x, \phi(x)) = 0$ for all $x \in I$, the derivative of the function $h(x) = F(x, \phi(x))$ is 0. But this derivative can also be computed using the Chain Rule for Functions of Several Variables.

$$0 = F_1(x, \phi(x)) + F_2(x, \phi(x))\phi'(x).$$

Solving this equation for $\phi'(x)$ yields

$$\phi'(x) = \frac{-F_1(x, \phi(x))}{F_2(x, \phi(x))}.$$

It is also possible to "solve" an equation of the form $F(x, y, z)$ for z as a function of z and y.

Implicit Function Theorem II. *Let $F : D \subset \mathbb{R}^3 \to \mathbb{R}$ and let (x_0, y_0, z_0) be an interior point of D with $F(x_0, y_0, z_0) = 0$. Suppose all first order partial derivatives of F exist in D and are continuous at (x_0, y_0, z_0) with $F_z(x_0, y_0, z_0) \neq 0$. Then there is a rectangle $R \subset \mathbb{R}^2$ with (x_0, y_0) an interior point of R and a function $\phi : R \to \mathbb{R}$ such that ϕ has continuous first order partial derivatives on R, $\phi(x_0, y_0) = z_0$ and $F\big((x, y, \phi(x, y)\big) = 0$ for each $(x, y) \in R$.*

As before the Chain Rule for Functions of Several Variables is then used to derive a formulas for $\phi_1(x, y)$ and for $\phi_2(x, y)$. Because $F\big((x, y), \phi(x, y)\big) = 0$ for all $(x, y) \in R$, the partial derivatives of the function $h(x, y) = F\big((x, y, \phi(x, y)\big)$ are 0. But these partial derivatives can also be computed using the Chain Rule for Functions of Several Variables. First compute the partial derivative with respect to x.

$$0 = F_1\big((x, y, \phi(x, y)\big) + F_3\big((x, y, \phi(x, y)\big)\phi_1(x, y).$$

Solving this equation for $\phi_1(x, y)$ yields

$$\phi_1(x, y) = \frac{-F_1\big((x, y, \phi(x, y)\big)}{F_3\big((x, y, \phi(x, y)\big)}.$$

As an exercise, you are encouraged to show that

$$\phi_2(x, y) = \frac{-F_2\big((x, y, \phi(x, y)\big)}{F_3\big((x, y, \phi(x, y)\big)}.$$

These types of applications of the Chain Rule for Functions of Several Variables are typical of its use. It isn't used to compute specific partial derivatives. The Chain Rule from first semester calculus is sufficient for that purpose.