# THE CHAIN RULE FOR FUNCTIONS OF SEVERAL VARIABLES AND THE IMPLICIT FUNCTION THEOREM

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 use the Chain Rule to determine $\dfrac{d}{dx}\arcsin(x^3)$. Employing the notation used above $g'(x) = \dfrac{1}{\sqrt{1-x^2}}$ 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

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

To discuss the Chair Rule for functions of several variables in the most concise fashion, we first define differentiability for a function whose domain a subset of $\mathbb{R}^n$ and whose range is a subset of $\mathbb{R}^k$. For our purposes $n$ and $k$ will be restricted to 1,2, and 3. The notation used for such a function is $F : D \subset \mathbb{R}^n \to \mathbb{R}^k$. Observe that if $F$ is such a function and if $P = (x_1, x_2, \ldots, x_n) \in D$, then $F(P)$ is a point in $k$ dimensional space; that is, $F(P) = (y_1, y_2, \ldots, y_k)$. The values of the numbers $y_1, y_2, \ldots, y_k$ depend on $P$; that is, they are each a function of $P$. Consequently they are denoted by $y_1 = f_1(P)$, $y_2 = f_2(P), \ldots, y_k = f_k(P)$. (These subscripts mustn't be confused with partial derivatives.) For example

$$F(x, y, z) = (y\sin(xz), (\arctan xz)e^{xy}).$$

Here $F : \mathbb{R}^3 \to \mathbb{R}^2$.

**Definition 1.** *Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^k$ and let $P_0$ be an interior point of $D$. For each $P \in D$ let $F(P) = (f_1(P), f_2(P), \ldots, f_k(P))$. Then $F$ is differentiable at $P_0$ means $f_1, f_2, \ldots, f_k$ are differentiable at $P_0$. The derivative of $F$ is denoted by $DF$ and is defined to be the following $k \times n$ matrix*

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(P_0) & \frac{\partial f_1}{\partial x_2}(P_0) & \ldots & \frac{\partial f_1}{\partial x_n}(P_0) \\ \frac{\partial f_2}{\partial x_1}(P_0) & \frac{\partial f_2}{\partial x_2}(P_0) & \ldots & \frac{\partial f_2}{\partial x_n}(P_0) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_k}{\partial x_1}(P_0) & \frac{\partial f_k}{\partial x_2}(P_0) & \ldots & \frac{\partial f_k}{\partial x_n}(P_0) \end{bmatrix}$$

Notice that the first row of the derivative matrix is the derivative of the first coordinate function of $F$ at $P_0$, the second row is the derivative of the second coordinate function of $F$ at $P_0$, etc. The reader should be able to see that if $F : \mathbb{R}^3 \to \mathbb{R}^2$, then the derivative matrix would be a $2 \times 3$ matrix; that is, one with two rows and three columns. For example if $F(x, y, z) = (x\sin y + e^{yz}z^2, x^3\ln z - yz)$, then the general derivative matrix for $F$ would be

$$\begin{bmatrix} \sin y & x\cos y + e^{yz}z^3 & e^{yz}yz^2 + eyz2z \\ 3x^2\ln z & -z & x^3z^{-1} - y \end{bmatrix}$$

and the derivative matrix at the point $(-1, 0, 1)$ is

$$\begin{bmatrix} 0 & 0 & 2 \\ 0 & -1 & -1 \end{bmatrix}$$

The derivative matrix plays an important role in the statement of the Chain Rule, the most important of the differentiation rules. The Chain Rule deals with the differentiation of the composition of two functions. So next we define composition for functions From $\mathbb{R}^n$ to $\mathbb{R}^k$..

**Definition 2.** *Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^k$ and let $G : E \subset \mathbb{R}^k \to \mathbb{R}^m$. Then the composition of $G$ by $F$ is the function denoted by $G \circ F$ and defined by $(G \circ F)(P) = G(F(P))$ for each $P \in D$ for which $F(P) \in E$.*

For our purposes we will work only with values for $n$, $k$ and $m$ being one of the integers 1, 2 or 3.

**Chain Rule for Fucntions of Several Variables.** *Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^k$ and let $G : E \subset \mathbb{R}^k \to \mathbb{R}^m$. Suppose $F$ is differentiable at $P_0$, an interior point of $D$, and that $G$ is differentiable at $F(P_0)$, an interior point of $E$. Then $G \circ F$ is differentiable at $P_0$ and $D(G \circ F)(P_0) = DG(F(P_0))DF(P_0)$; that is, the derivative matrix of $G \circ F$ at $P_0$ is the matrix product of the derivative matrix of $G$ at $F(P_0)$ times the derivative matrix of $G$ at $P_0$.*

Note that the statement of this theorem is a direct analogue of the Chain Rule for functions of one variable. The proof of this theorem is essentially the same as the proof for the one variable case once the algebra of matrices is developed.

We will now see how to use the Chain Rule to obtain formulas for the partial derivatives of composite functions that you will find in the text. For example suppose that $F : D \subset \mathbb{R}^2 \to \mathbb{R}^3$ and $G : E \subset \mathbb{R}^3 \to \mathbb{R}$. Write $F(x, y) = (u(x, y), v(x, y), w(x, y))$. Then

$$DF = \begin{bmatrix} u_1(x,y) & u_2(x,y) \\ v_1(x,y) & v_2(x,y) \\ w_1(x,y) & w_2(x,y) \end{bmatrix}$$

and $DG(u, v, w) = \begin{bmatrix} G_1(u,v,w) & G_2(u,v,w) & G_3(u,v,w) \end{bmatrix}$. Let $H = G \circ F$. Then

$$\begin{aligned} DH(x,y) =& DG(F(x,y))DF(x,y) \\ =& \begin{bmatrix} G_1(F(x,y)) & G_2(F(x,y)) & G_3(F(x,y)) \end{bmatrix} \begin{bmatrix} u_1(x,y) & u_2(x,y) \\ v_1(x,y) & v_2(x,y) \\ w_1(x,y) & w_2(x,y) \end{bmatrix} \\ =& \begin{bmatrix} G_1(F(x,y))u_1(x,y) + G_2(F(x,y))v_1(x,y) + G_3(F(x,y)w_1(x,y) \\ G_1(F(x,y))u_2(x,y) + G_2(F(x,y))v_2(x,y) + G_3(F(x,y)w_2(x,y) \end{bmatrix}. \end{aligned}$$

Consequently

$$H_1(x,y) = G_1(F(x,y))u_1(x,y) + G_2(F(x,y))v_1(x,y) + G_3(F(x,y)w_1(x,y)$$

and

$$H_2(x,y) = G_1(F(x,y))u_2(x,y) + G_2(F(x,y))v_2(x,y) + G_3(F(x,y)w_2(x,y).$$

The general formula for partial derivatives of compositions is obtained in a similar fashion.

# 1 Implicit Function Theorem

In Section 2.6 the technique of implicit differentiation was investigated for finding the derivative of a function defined implicitly by an equation in two variables such as $x^3 - xy^2 + y^3 = 1$. In the language of functions of several variables, such equations can be written as $F(x, y) = 0$. In Section 2.6 it was **assumed** that the equation could be "solved" for $y$ as a function of $x$. Then a technique, based on the Chain Rule, was developed to find $\dfrac{dy}{dx}$. Here we state a theorem that gives conditions on the partial derivatives of $F$ when the equation $F(x, y) = 0$ can be "solved"; that is, when there is a function $y = \phi(x)$ such that $F(x, \phi(x)) = 0$ for every $x$ in the domain of $\phi$. In addition it gives a formula for the derivative of the function $\phi$. Several of the problems in the text pertain to this theorem.

**Implicit Function Theorem.** *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 and are continuous in $D$ 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 $\phi$ is differentiable on $I$, $\phi(x_0) = y_0$ and $F(x, \phi(x)) = 0$ for each $x \in I$. Moreover*

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

The advanced part of this theorem is proving the existence of the function $\phi$. From there the Chain Rule for Functions of Several Variables is 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

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

There are several other versions of this theorem. As an example we state one where the challenge is to solve an equation of the form $F(x, y, z) = 0$ for $z$ in terms of $x$ and $y$.

**Theorem 1.** *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 and are continuous in $D$ with $F_3(x_0, y_0, z_0) \neq 0$. 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 $\phi$ is differentiable on $R$, $\phi(x_0, y_0) = z_0$ and $F(x, y, \phi(x, y)) = 0$ for all $(x, y) \in R$. Moreover*

$$\phi_1(x, y) = -\frac{F_1((x, y, \phi(x, y))}{F_3(x, y, \phi(x, y))} \text{ and } \phi_2(x, y) = -\frac{F_2((x, y, \phi(x, y))}{F_3(x, y, \phi(x, y))}.$$