FUNCTIONS OF SEVERAL VARIABLES

1 Limits and Continuity

We begin with a review of the concepts of limits and continuity for real-valued functions of one variable. Recall that the definition of the limit of such functions is as follows.

Definition 1.1. *Let $f : D \subset \mathbb{R} \to \mathbb{R}$ and let $a \in \mathbb{R}$. Then $\lim_{x\to a} f(x) = L$ means that for each $\epsilon > 0$ there is a $\delta > 0$ if $0 < |x - a| < \delta$, then such that $|f(x) - L| < \epsilon$.*

The two fundamental specific limits results which follow easily from the definition are:

$$(1) \text{ If } c \in \mathbb{R}, \text{ then } \lim_{x\to a} c = c \text{ and } (2) \lim_{x\to a} x = a \text{ for any } a \in \mathbb{R}.$$

The basic facts used to compute limits are contained in the following theorem.

Theorem 1.1. Basic Limit Theorem: *Let $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$. Then*

1. $\lim_{x\to a} f(x) + g(x) = L + M$
2. $\lim_{x\to a} f(x)g(x) = LM$ *and*
3. $\lim_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{L}{M}$ *provided $M \neq 0$.*

Moreover, if $\lim_{x\to a} f(x) = L = \lim_{x\to a} g(x)$ and if $f(x) \leq h(x) \leq g(x)$, then $\lim_{x\to a} h(x) = L$. (This assertion is commonly called the Sandwich Theorem.)

From these assertions it was proved that for any rational function R (Recall that a rational function is the quotient of two polynomials.) $\lim_{x\to a} R(x) = R(a)$ for any $a \in D_R$. (Recall that the domain of a rational function is the set of numbers where the polynomial in the denominator is not 0.)

Continuity was defined taking a hint from the above result.

Definition 1.2. *Let $f : D \subset \mathbb{R} \to \mathbb{R}$ and let $a \in D$. Then f is continuous at L means $\lim_{\substack{x\to a\\ x\in D}} f(x) = f(a)$.*

By the comment preceding Definition 1.2 all rational functions are continuous at each number in their domains. The same is true of all of the trigonometric functions, the logarithmic functions, the exponential functions and the inverse trigonometric functions.

From the Basic Limit Theorems and the definition of continuity the Basic Continuity Theorem follows immediately.

Theorem 1.2. Basic Continuity Theorem: *Let $f, g : D \subset \mathbb{R} \to \mathbb{R}$ and let $a \in D$. Suppose f and g are continuous at a. Then $f + g$ and fg are continuous at a as is $\dfrac{f}{g}$ provided $g(a) \neq 0$.*

Neither Theorem 1.1 nor Theorem 1.2 deal with the most important method of combining two functions; namely, the composition of two functions. The definition of that concept is recalled next.

Definition 1.3. *Let $f : D \subset \mathbb{R} \to \mathbb{R}$ and $g : E \subset \mathbb{R} \to \mathbb{R}$. Then the composition of f by g is denoted by $g \circ f$ and defined by $g \circ f(x) = g(f(x))$.*

How limits and continuity are related to composition is explained in the following two theorems.

Theorem 1.3. Limit Composition Theorem *Let f and g be as in Definition 1.3 with $a \in D$ and $L \in E$. Suppose $\lim_{x\to a} f(x) = L$ and suppose g is continuous at L. Then $\lim_{x\to a} g(f(x)) = g(L)$.*

The conclusion of this assertion can also be written as $\lim_{x\to a} g(f(x)) = g(\lim_{x\to a} f(x))$ which can then be remembered as, "the limit of a continuous function is the continuous function of the limit." An immediate consequence of this theorem is the following corollary.

Theorem 1.4. Continuity Composition Theorem: *Let f and g be as in Definition 1.3 with $a \in D$ and $f(a) \in E$. Suppose f is continuous at a and g is continuous at $f(a)$. Then $g \circ f$ is continuous at a.*

One remembers this assertion as, "the composition of two continuous functions is continuous."

This completes our review of the single variable situation. Now we take up the subjects of Limits and Continuity for real-valued functions of several variables.

Definition 1.4. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$, let $P_0 \in \mathbb{R}^n$ and let $L \in \mathbb{R}$. Then $\lim_{\substack{P\to P_0\\ P\in D}} f(P) = L$ means that the distance, for each $\epsilon > 0$ there is a $\delta > 0$ such that if $P \in D$ and if $0 < |P - P_0| < \delta$, then $|f(P) - L| < \epsilon$.*

Note that the first use of vertical lines denotes absolute value while the second denotes distance between two points in $\mathbb{R}^n$.

To begin computing limits we first need some specific results similar to those for functions of one variable. The basic principle is that if a function of one variable is considered as a function of more than one variable, then the limit of the function is computed by taking the limit of the function with respect to its only variable. One specific case of this principle is stated below. Other cases are left to the reader's imagination.

Theorem 1.5. *Let $h : E \subset \mathbb{R} \to \mathbb{R}$ and set $f(x, y) = h(x)$. Suppose $\lim_{x\to a} h(x) = L$. Then $\lim_{(x,y)\to(a,b)} f(x, y) = L$ for any $b \in \mathbb{R}$.*

So for example $\lim_{(x,y)\to(2,9)} x^2 = \lim_{x\to 2} x^2 = 4$ and $\lim_{(x,y)\to(2,9)} \sqrt{y} = \lim_{y\to 9} \sqrt{y} = 3$.

Essentially all examples of functions of several variables we will encounter are constructed from functions of one variable by addition, multiplication, division and composition. So the following Basic Limit Theorem will permit us to compute limits.

Theorem 1.6. Basic Limit Theorem: *Let $f, g : D \subset \mathbb{R}^n \to \mathbb{R}$. Suppose $\lim_{\substack{P\to P_0\\ P\in D}} f(P) = L$ and $\lim_{\substack{P\to P_0\\ P\in D}} g(P) = M$. Then*

1. $\lim_{\substack{P\to P_0\\ P\in D}} f(P) + g(P) = L + M$
2. $\lim_{\substack{P\to P_0\\ P\in D}} f(P)g(P) = LM$ *and*
3. $\lim_{\substack{P\to P_0\\ P\in D}} \dfrac{f(P)}{g(P)} = \dfrac{L}{M}$ *provided $M \neq 0$.*

Moreover, if $\lim_{\substack{P\to P_0\\ P\in D}} f(P) = L = \lim_{\substack{P\to P_0\\ P\in D}} g(P)$ and if $f(P) \le h(P) \le g(P)$, then $\lim_{\substack{P\to P_0\\ P\in D}} h(P) = L$. (The Sandwich Theorem for functions of several variables.)

Examples: Page 921; 8, 14, 18.

Doing problem 8 and others on Page 921 requires the use of the following assertion analogous to the one variable version above.

Theorem 1.7. Limit Composition Theorem: *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ and let $h : E \subset \mathbb{R} \to \mathbb{R}$. Suppose $\lim_{\substack{P\to P_0\\ P\in D}} f(P) = L$ and suppose h is continuous at L. Then $\lim_{\substack{P\to P_0\\ P\in D}} h(f(P)) = h(L)$.*

A comment similar to the one following Theorem 1.3 applies here.

Additional examples: Page 921; 3, 10.

Having defined the limit concept for functions of several variables, the notion of continuity for such functions is defined in a fashion analogous to the one variable situation.

Definition 1.5. *Let $f : D \subset \mathbb{R}^n$ and let $P_0 \in D$. Then f is continuous at P_0 means $\lim_{\substack{P\to P_0\\ P\in D}} f(P) = f(P_0)$.*

Note that contrary to the limit definition, we require that P_0 be in the domain of the function. The purpose is to guarantee that $f(P_0)$ is defined. Also according to definition of limit for functions of several variables, the only values of P that are allowed are those that are in D. We emphasize that here by adding $P \in D$ in the limit statement.

The Basic Limit Theorem and the Limit Composition Theorem yield the next two theorems.

Theorem 1.8. Basic Continuity Theorem: *Let f, $g : D \subset \mathbb{R}^n \to \mathbb{R}$ and let $P_0 \in D$. Suppose f and g are continuous at P_0. Then $f + g$ and fg are continuous at P_0 as is $\dfrac{f}{g}$ provided $g(P_0) \neq 0$.*

Examples: Page 922: 28, 29(b), 30, 31(a) 33, 34

Theorem 1.9. Continuity Composition Theorem: *Let $f : D \subset \mathbb{R}^n$, let $P_0 \in D$ and let $h : E \subset \mathbb{R} \to \mathbb{R}$. Suppose f is continuous at P_0 and that h is continuous at $f(P_0)$. Then the function $h \circ f$ defined by $(h \circ f)(P) = h(f(P))$ is continuous at P_0.*

Examples: Page 922; 27, 29(a), 31(b), 32.

2 Differentiation

We begin this section by reviewing the concept of differentiation for functions of one variable.

Definition 2.1. *Let $f : D \subset \mathbb{R} \to \mathbb{R}$ and let a be an interior point of D. (A point $a \in D$ is an interior point of D means there is an $r > 0$ such that $(a - r, a + r) \subset D$.) Then f is differentiable at a means there is a number, denoted by $f'(a)$ such that*

$$\lim_{h \to 0} \frac{f(a + h) - f(a)}{h} = f'(a)$$

or equivalently

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = f'(a)$$

exists. The number $f'(a)$ is called the derivative of f at a.

Geometrically the derivative of a function at a is interpreted as the slope of the line tangent to the graph of f at the point $(a, f(a))$. Not every function is differentiable at every number in its domain even if that function is continuous. For example $f(x) = |x|$ is not differentiable at 0 but f is continuous at 0. However we do have the following theorem.

Theorem 2.1. *If f is differentiable at a, then f is continuous at a.*

Extending the definition of differentiability in its present form to functions of several variables is not possible because the definition involves division and dividing by a vector or by a point in n dimensional space is not possible. To carry out the extension an equivalent definition is developed that involves division by a distance. The limit statement can be rewritten as

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} - f'(a) = 0 \text{ or } \lim_{x \to a} \frac{f(x) - f(a) - (x - a)f'(a)}{x - a} = 0.$$

One final modification is still necessary.

$$\lim_{x \to a} \frac{f(x) - f(a) - (x - a)f'(a)}{|x - a|} = 0.$$

So the following definition is equivalent to the original one.

Definition 2.2. *Let $f : D \subset \mathbb{R} \to \mathbb{R}$ and let a be an interior point of D. Then f is differentiable at a means there is a number, $f'(a)$, such that*

$$\lim_{x \to a} \frac{f(x) - f(a) - (x - a)f'(a)}{|x - a|} = 0.$$

One way to interpret this expression is that $f(x) - f(a) - (x - a)f'(a)$ tends to 0 faster than $|x - a|$ tends to 0 and consequently $f(x)$ is approximately equal to $f(a) + (x - a)f'(a)$. The equation $y = f(a) + (x - a)f'(a)$ is an equation of the line tangent to the graph of f at the point $(a, f(a))$. So $f(x)$ is approximated very well by its tangent line. This observation is the bases for linear approximation.

Using this form of the definition as a model it is possible to construct a definition of differentiability for functions of several variables. What goes in the denominator is fairly easy to see; namely, $|P - P_0|$. Similarly the first two term in the numerator would become $f(P) - f(P_0)$. But what should replace the term $(x - a)f'(a)$? First we note that it must be a number. One of the factors will be $(P_0 - P)$ or better yet $(\overrightarrow{P_0P})$ — a vector. Consequently the other must also be a vector and the product will be the dot product. With these observation the definition of differentiability for functions of several variable is as follows.

Definition 2.3. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ and let P_0 be an interior point of D. (A point $P_0 \in D$ is an interior point of D means there is an $r > 0$ such that $\{P \in \mathbb{R}^n; |P - P_0| < r\} \subset D$.) Then f is differentiable at P_0 means there is a vector, denoted by $f'(P_0)$ for now, such that*

$$\lim_{P \to P_0} \frac{f(P) - f(P_0) - (\overrightarrow{P_0P}) \cdot f'(P_0)}{|P - P_0|} = 0.$$

For functions of two variables the definition becomes the following.

Definition 2.4. *Let $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ and let (x_0, y_0) be an interior point of D. Then f is differentiable at (x_0, y_0) means there are two numbers, $f_1(x_0, y_0)$ and $f_2(x_0, y_0)$ such that*

$$\lim_{(x,y) \to (x_0,y_0)} \frac{f(x, y) - f(x_0, y_0) - (x - x_0)f_1(x_0, y_0) - (y - y_0)f_2(x_0, y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0. \tag{1}$$

The vector $f_1(x_0, y_0)\mathbf{i} + f_2(x_0, y_0)\mathbf{j}$ or the pair $(f_1(x_0, y_0), f_2(x_0, y_0))$ is called the derivative of f at the point (x_0, y_0).

Interpret this definition as requiring that the graph of f have a tangent plane at the point $(x_0, y_0, f(x_0, y_0))$. In fact it is easy to get an equation for this tangent plane. It is

$$z = f(x_0, y_0) + f_1(x_0, y_0)(x - x_0) + f_2(x_0, y_0)(y - y_0).$$

A vector normal to this plane is $f_1(x_0, y_0)\mathbf{i} + f_2(x_0, y_0)\mathbf{j} - \mathbf{k}$. The two numbers, $f_1(x_0, y_0)$ and $f_2(x_0, y_0)$, are computed using techniques learned for computing derivatives of functions of one variable. To find $f_1(x_0, y_0)$ let $y = y_0$ in equation (1) of Definition 2.4. We get

$$\lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0) - (x - x_0)f_1(x_0, y_0)}{|x - x_0|} = 0.$$

Comparing this statement to Definition 2.2 we see that $f_1(x_0, y_0)$ is the derivative of the function $h(x) = f(x, y_0)$ at x_0. For example suppose $f(x, y) = x^2y + xy^3$. Then $h(x) = x^2y_0 + xy_0^3$. Differentiating h with respect to x; that is, treating y_0 as a constant, we get that $f_1(x_0, y_0) = 2x_0y_0 + y_0^3$ or more generally for each (x, y) we have $f_1(x, y) = 2xy + y^3$. Notice that this equation is obtained by differentiating the formula for $f(x, y)$ with respect to x treating y as if it were a constant. In a similar fashion $f_2(x, y)$ is obtained by differentiating the formula for $f(x, y)$ with respect to y treating x as a constant. So for the preceding example we get $f_2(x, y) = x^2 + 3xy^2$.

We call $f_1(x, y)$ the first partial derivative of f with respect to x (or with respect to the first variable) and $f_2(x, y)$ the first partial derivative of f with respect to y (or with respect to the second variable). As the word, "first" indicates, there are "second", "third" etc. order partial derivatives as well. We will discuss them later. When it is clear that we are dealing with first order partial derivatives the word "first" is often omitted. The following notation is also used to denote partial derivatives.

$$f_1(x, y) = \frac{\partial f}{\partial x}(x, y) = \frac{\partial}{\partial x} f(x, y) = f_x(x, y).$$

In all three expressions the same symbol x is use for two different purposes. First to denote the variable of differentiation and second as the first coordinate of a point in $\mathbb{R}^2$. Strictly speaking such a dual use of one symbol is improper, but this abuse is so common as to be acceptable. As one would expect there is analogous notation for $f_2(x, y)$. Consequently, $f'((x_0, y_0)) = \big(\frac{\partial}{\partial x} f(x_0, y_0)\big)\mathbf{i} + \big(\frac{\partial}{\partial y} f(x_0, y_0)\big)\mathbf{j}$. This vector is called the gradient of f at (x_0, y_0) and denoted by $\operatorname{grad} f(x_0, y_0)$ or $\nabla f(x_0, y_0)$. If $f'(x_0, y_0)$ exists, it is $\operatorname{grad} f(x_0, y_0)$.

The situation for functions of more than two variables is analogous. In the general case, the derivative is a vector in n space and it is computed by finding all of the first order partial derivatives of f at P_0. The derivative of f at P_0 is also called the gradient of f at P_0 and denoted by $\operatorname{grad} f(P_0)$ or $\nabla f(P_0)$. As above, if f has a derivative at P_0, then it is $\operatorname{grad} f(P_0)$.

As in the case of functions of one variable, differentiability implies continuity.

Theorem 2.2. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ and let P_0 be an interior point of D. Suppose f is differentiable at P_0. Then f is continuous at P_0.*

PROOF: First write

$$\begin{aligned} f(P) - f(P_0) &= f(P) - f(P_0) - (P - P_0) \cdot \operatorname{grad} f(P_0) + (P - P_0) \cdot \operatorname{grad} f(P_0) \\ &= \frac{f(P) - f(P_0) - (P - P_0) \cdot \operatorname{grad} f(P_0)}{|P - P_0|}|P - P_0| + (P - P_0) \cdot \operatorname{grad} f(P_0) \end{aligned}$$

Since both terms on the right hand side have limit 0 as $P \to P_0$,

$$\lim_{P \to P_0} f(P) - f(P_0) = 0; \text{ that is, } \lim_{P \to P_0} f(P) = f(P_0).$$

The converse of the preceding theorem is not true since the converse of the analogous theorem for functions of one variable is not true. Recall that $f(x) = |x|$ is continuous but isn't differentiable at 0.

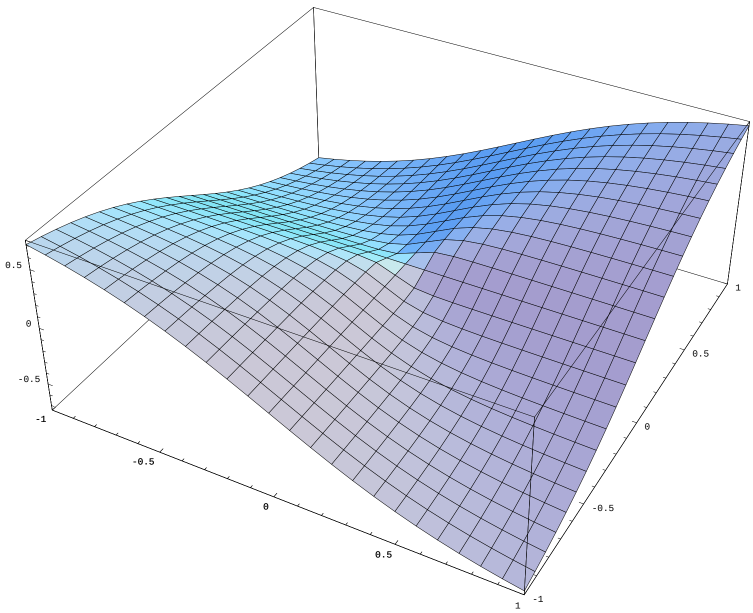
The analogy between differentiation for functions of one variable and for functions of several variable is not a total one. For functions of one variable if the derivative, $f'(x)$, can be computed, then f is differentiable at x. The corresponding assertion for functions of two variables is false which stands to reason after considering for a moment what it takes to compute the derivative, grad $f(x, y)$, of a function of two variable. To find $f_1(x_0, y_0)$ one need only know the values of the function, f, at points on the line $y = y_0$ and to find $f_2(x_0, y_0)$ one need only know the values of f at points on the line $x = x_0$. Consequently, the values of f at points not on these two lines play no role in determining the derivative of f. However these values certainly are taken into account when determining whether or not f is differentiable at (x_0, y_0); that is, if the graph of f has a tangent plane at the point (x_0, y_0). For example let

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \text{ or } y = 0 \\ 1 & \text{otherwise .} \end{cases}$$

Since f is 0 on the two coordinate axes, $f_1(0, 0) = 0 = f_2(0, 0)$ but f is not continuous at $(0, 0)$ and by the preceding theorem, f can't be differentiable at $(0, 0)$. You might suspect that if f is continuous at (x_0, y_0) and the first order partial derivatives exist there, then f is differentiable at (x_0, y_0) but that conjecture is false as the following example shows. Let

$$f(x, y) = \begin{cases} \dfrac{xy}{\sqrt{x^2 + y^2}} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

The graph of f is pictured below.



Again since f is 0 on the two coordinate axes, $f_1(0, 0) = 0 = f_2(0, 0)$. So if f were differentiable at $(0, 0)$, we would have that

$$\lim_{(x,y) \to (0,0)} \frac{f(x, y)}{\sqrt{x^2 + y^2}} = 0; \text{ that is, } \lim_{(x,y) \to (0,0)} \frac{xy}{x^2 + y^2} = 0.$$

But if the limit is computed along the path $y = x$, we get $\lim_{x \to 0} \dfrac{x^2}{2x^2} = \dfrac{1}{2}$.

The natural question to ask then is under what conditions can we conclude that f is differentiable at (x, y). The answer is contained in the following theorem.

Theorem 2.3. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ and let P_0 be an interior point of D. Suppose all of the first order partial derivatives of f exist in a ball about P_0 and are continuous at P_0. Then f is differentiable at P_0.*

For example let $f(x,y) = \sqrt{y^2 - x^2} = (y^2 - x^2)^{1/2}$. Then $f_1(x,y) = -x(y^2 - x^2)^{-1/2}$ and $f_2(x,y) = y(y^2 - x^2)^{-1/2}$. These two functions are continuous in the region consisting of that part of $\mathbb{R}^2$ above the graph of $y = |x|$ together with that part of $\mathbb{R}^2$ below the graph of $y = -|x|$. According to the theorem, f is differentiable on this region.