Advanced Calculus

Professor David Wagner

[1]Department of Mathematics
University of Houston

September 18

Definition

Definition

Suppose $f : D \subset \mathbb{R}^n \to \mathbb{R}^k$ and suppose a is an interior point of D. If the limit

$$\lim_{h \to 0} \frac{f(x_1, \cdots, x_i + h, \cdots, x_n) - f(x_1, \cdots, x_i, \cdots, x_n)}{h}$$

exists, we call this limit the *partial derivative* of f with respect to x_i. Several notations are used for this:

$$\frac{\partial f}{\partial x_i}, \quad f_{x_i}, \quad f_i, \quad \partial_{x_i} f, \quad \partial_i f.$$

Example

Let $f(x, y, z) = \frac{e^{3x}\sin(xy)}{1+5y-7z}$. Then

$$\partial_x f = \partial_1 f = \frac{\partial f}{\partial x} = \frac{3e^3 x \sin(xy) + e^{3x} y \cos(xy)}{1 + 5y - 7x}, \tag{1}$$

$$\partial_y f = \partial_2 f = \frac{\partial f}{\partial y} = \frac{e^{3x} y \cos(xy) - 5e^{3x}\sin(xy)}{(1 + 5y - 7x)^2} \tag{2}$$

$$\partial_z f = \partial_3 f = \frac{\partial f}{\partial z} = \frac{7e^{3x}\sin(xy)}{(1 + 5y - 7x)^2} \tag{3}$$

Example

We have already seen that the function $f(x, y) = \frac{xy}{x^2+y^2}$ has no limit as $(x, y) \to 0$. But

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{h0}{h^2 + 0} = 0,$$

and similarly $\frac{\partial f}{\partial y}(0,0) = 0$.

Thus, the existence of first order partial derivatives does not give us continuity of f.

We need a definition of derivative that gives us the rate of change of $f(x)$ in any direction. We could try a straight-forward generalization of the $1 - D$ derivative:

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

where we let x and h denote vectors. This has the obvious difficulty that we can't divide by the vector h.

Instead we start with the relation

$$f(x+h) - f(x) - f'(x)h = \epsilon(h)h,$$

where $\epsilon(h) \to 0$ as $h \to 0$. With two changes, this makes perfect sense if we allow x and h to be vectors in $\mathbb{R}^n$, and we take f and ϵ to be functions on $D \subset \mathbb{R}^n$ to $\mathbb{R}$. One change is that we replace $\epsilon(h)h$ with $\epsilon(h)\,\|h\|$. The other change is that we replace $f'(x)h$ by either $Df(x)(h)$, or $\nabla f(x) \cdot h$. With the first choice, $Df(x)$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}$ (so that $Df(x)$ maps h to $\mathbb{R}$). The second choice is to identify a vector "$\nabla f(x)$" so that $\nabla f(x) \cdot h = Df(x)(h)$. As it turns out, for real valued functions of several variables, it is simpler to work with ∇f.

Definition of "differentiable" and "gradient"

Definition

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$, and let $\mathbf{x}$ be an interior point of $\mathbb{R}$. We say that f is *differentiable* at $\mathbf{x}$ if there is a vector $\mathbf{v} \in \mathbb{R}^n$ such that

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \mathbf{v} \cdot \mathbf{x} + \epsilon(\mathbf{h}) \|\mathbf{h}\| .$$

In this case, $\mathbf{v}$ uniquely determined and is called the *gradient* of f at $\mathbf{x}$, $\nabla f(\mathbf{x})$.

Equivalently, we could divide this equation by $\|\mathbf{h}\|$ to get:

$$\frac{f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - \nabla f(\mathbf{x}) \cdot \mathbf{h}}{\|\mathbf{h}\|} = \epsilon(\mathbf{h}) \to 0 \text{ as } \|\mathbf{h}\| \to 0.$$

Either characterization tells us that the (usually) nonlinear function of $\mathbf{h}$ given by $g(\mathbf{h}) = f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})$ is very well approximated by the linear function of $\mathbf{h}$, $L(\mathbf{h}) = \nabla f(\mathbf{x}) \cdot \mathbf{h}$. The difference is $\epsilon(\mathbf{h}) \|\mathbf{h}\|$, not $C \|\mathbf{h}\|$.

This relation also characterizes the meaning of a tangent plane ($n = 2$) or tangent hyperplane ($n > 2$). To see this let $n = 2$. The graph of f, $z = f(x + h_1, y + h_2)$ intersects the graph of $z = f(x, y) + \nabla f(x, y) \cdot \mathbf{h}$ at $(x, y, f(x, y))$. The two graphs stay very close to each other as long as $\mathbf{h} = (h_1, h_2)$ is small: the vertical distance between the graphs at $(x + h_1, y + h_2)$, is $\epsilon(\mathbf{h}) \|\mathbf{h}\|$.

Now let $\mathbf{h} = he_i$, where e_i is the i^{th} standard unit vector ($\{e_1, \cdots, e_n\}$ form an orthonormal basis for $\mathbb{R}^n$). Then

$$\lim_{h \to 0} \frac{f(x_1, \cdots, x_i + h, \cdots, x_n) - f(x_1, \cdots, x_n)}{h} = \nabla f(\mathbf{x}) \cdot e_i.$$

The left hand side is $\frac{\partial f}{\partial x_i}(\mathbf{x})$. The right side is the i^{th} component of $\nabla f(\mathbf{x})$. We conclude:

Theorem

If f is differentiable at $\mathbf{x}$, then all first order partial derivatives $\frac{\partial f}{\partial x_i}(\mathbf{x})$, $i = 1, \cdots, n$ exist and equal the components of $\nabla f(\mathbf{x})$.

For a different perspective, consider expanding $\nabla f(\mathbf{x})$ in the orthonormal basis $(e_1, \cdots, e_n)$:

$$\nabla f(\mathbf{x}) = \sum_{i=1}^{n} (\nabla f(\mathbf{x}) \cdot e_i)\, e_i = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}) \cdot e_i.$$

Example, Theorem

Let $f(x, y, z) = x^2 + 2y^2 + 3z^2$. Then $\nabla f(x, y, z) = (2x, 4y, 6z)$.
$\nabla f(1, -2, 5) = (2, -8, 30)$

Theorem

If f is differentiable at $\mathbf{x}$, then f is continuous at $\mathbf{x}$.

Proof.

$$f(\mathbf{y}) - f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) + \epsilon(\mathbf{y} - \mathbf{x}) \|\mathbf{y} - \mathbf{x}\| \to 0 \text{ as } \mathbf{y} \to \mathbf{x}.$$

□

Definition

If f and its first order partial derivatives are continuous in an open set U, we say that f is continuously differentiable, or of class C^1, on U.

Theorem

If f is continuously differentiable on an open set U, then f is differentiable at each $\mathbf{x} \in U$.

Proof

Proof: For simplicity, let $n = 2$. Fix $\mathbf{a} = (x, y) \in U$, and let $\epsilon > 0$. For $\mathbf{h} = (h_1, h_2)$ near $\mathbf{0}$,

$$f(\mathbf{a}+\mathbf{h})-f(\mathbf{x}) = f(x+h_1, y+h_2)-f(x, y+h_2)+f(x, y+h_2)-f(x, y).$$

By the MVT there are numbers c_1 between 0 and h_1, and c_2 between 0 and h_2 such that:

$$\begin{aligned} f(x+h_1, y+h_2) - f(x, y+h_2) &= \frac{\partial f}{\partial x}(x+c_1, y+h_2)\, h_1 \\ f(x, y+h_2) - f(x, y) &= \frac{\partial f}{\partial y}(x, y+c_2)\, h_2. \end{aligned}$$

Proof continued

Then

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - \nabla f \cdot \mathbf{h} = \left(\frac{\partial f}{\partial x}(x + c_1, y + h_2) - \frac{\partial f}{\partial x}(\mathbf{a}) \right) h_1 + \left(\frac{\partial f}{\partial x}(x, y + c_2) - \frac{\partial f}{\partial y}(\mathbf{a}) \right) h_2.$$

Since $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are continuous on U, and c_i is between 0 and h_i, the factors of h_1 and h_2 tend to 0 as $\mathbf{h} \to 0$. So the right hand side is $\epsilon(\mathbf{h}) \|\mathbf{h}\|$. □

Directional Derivatives

e

Definition

Let D be an open subset of $\mathbb{R}^n$, and let $f : D \to \mathbb{R}$. Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^n$. We say that f has a *directional derivative* at $\mathbf{x} \in D$ in the direction $\mathbf{u}$, if

$$\lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{u}) - f(\mathbf{x})}{h} \text{ exists.}$$

We denote this limit by $(D_{\mathbf{u}} f)(\mathbf{x})$.

The partial derivatives of f are examples of directional derivatives (when the coordinate system is orthonormal). Note that the existence or value of a directional derivative does not depend on a choice of coordinates.

Directional Derivatives

Theorem

If f is differentiable at $\mathbf{x}$, then for each unit vector $\mathbf{u}$, $(D_{\mathbf{u}}f)(\mathbf{x})$ exists and equals $\nabla f(\mathbf{x}) \cdot \mathbf{u}$.

Proof.

$$\frac{f(\mathbf{x} + h\mathbf{u}) - f(\mathbf{x})}{h} = \nabla f(\mathbf{x}) \cdot \mathbf{u} + \epsilon(h\mathbf{u}).$$

In the limit as $h \to 0$, we obtain the stated result. □

Tangent plane

Let f be differentiable at $\mathbf{x}$. When $n = 2$ and $\mathbf{x}$ fixed, the graph of $z = f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot \mathbf{h}$ is a plane in $\mathbb{R}^3$ tangent to the graph of $z = f(\mathbf{x} + \mathbf{h})$ at $(\mathbf{h} = 0, z = f(\mathbf{x}))$. If $\mathbf{u}$ is a unit vector in $\mathbb{R}^2$, the line with direction vector $(\mathbf{u}, (D_{\mathbf{u}} f)(\mathbf{x}))$ through $(\mathbf{h} = \mathbf{0}, z = f(\mathbf{x}))$ is in the tangent plane because $(D_{\mathbf{u}} f)(\mathbf{x}) = \nabla f \cdot \mathbf{u}$.

Mean Value Theorem

Theorem

Suppose f is continuous on $[a, b]$ and differentiable on (a, b). Then there is at least one $c \in (a, b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Proof.

The line through $(a, f(a))$ and $(b, f(b))$ has slope $\frac{f(b)-f(a)}{b-a}$. Let

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)\,.$$

Then g is continuous on $[a, b]$ and differentiable on (a, b). $g(b) = g(a)$, so Rolle's Theorem states that there must be $c \in (a, b)$ such that $g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a} = 0$. □

Applications of MVT

Definition

We say that a function f is *increasing* (respectively *strictly increasing*) on an interval I, if $f(a) \leq f(b)$ (resp. $f(a) < f(b)$) whenever $a,\ b \in I$ and $a < b$. We define *decreasing* and *strictly decreasing* functions similarly.

Theorem (Interpretation of the derivative)

Suppose f is differentiable on the open interval I.

a *If $|f'(x)| \leq C$ for all $x \in I$, then $|f(b) - f(a)| \leq C\,|b - a|$ for all $a,\ b \in I$.*

b *If $f'(x) = 0$ for all $x \in I$, then f is constant on I.*

c *If $f'(x) \geq 0$ (resp. $f'(x) > 0$, $f'(x) \leq 0$, or $f'(x) < 0$) for all $x \in I$, then f is increasing (resp. strictly increasing, decreasing, strictly decreasing) on I.*

Proof

Proof.

Let $a,\ b \in I$. Since f is differentiable on I and $[a, b] \subset I$, f is continuous on I. Then the MVT gives us a point $c \in (a, b)$ such that $f(b) - f(a) = f'(c)(b - a)$. For a, $|f'(c)| \leq C$, so $|f(b) - f(a)| \leq C\,|b - a|$. Parts b, c are proved similarly. □

Remark

If all we know about f is that f is differentiable at a and $f'(a) > 0$, it does *not* follow that f is increasing in some neighborhood of a.

Generalized MVT

Theorem (Generalized MVT)

Suppose that f and g are continuous on $[a, b]$, differentiable on (a, b), and that $g'(x) \neq 0$ for all $x \in (a, b)$. Then there is $c \in (a, b)$ such that

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Proof of Generalized MVT

Proof

- Let $h(x) = (f(b) - f(a))\,(g(x) - g(a)) - (g(b) - g(a))\,(f(x) - f(a))$.
- Then h is continuous on $[a, b]$, differentiable on (a, b), and $h(a) = h(b) = 0$.
- By Rolle's Theorem there is $c \in (a, b)$ such that $h'(c) = (f(b) - f(a))\,g'(c) - (g(b) - g(a))\,f'(c) = 0$.
- Since g' is never 0 on (a, b), $g'(c) \neq 0$ and $g(b) - g(a) \neq 0$ (by MVT).
- Then dividing by $g'(c)\,(g(b) - g(a))$ gives the result.

Remark

If we use f and g to parameterize a curve: $y = f(t)$, $x = g(t)$, $t \in [a, b]$, then by the chain rule, $\frac{dy}{dx} = \frac{f'(t)}{g'(t)}$. In this case the Generalized MVT says that there is a tangent line to the parameterized curve, that is parallel to the secant line through $(a, f(a))$ and $(b, f(b))$.

Application of Generalized MVT

Theorem (L'Hôpital's Rule I)

Suppose f and g are differentiable functions on (a, b) and

$$\lim_{x \to a+} f(x) = \lim_{x \to a+} g(x) = 0.$$

If g' never vanishes on (a, b) and

$$\lim_{x \to a+} \frac{f'(x)}{g'(x)} = L,$$

then g never vanishes on (a, b) and

$$\lim_{x \to a+} \frac{f(x)}{g(x)} = L.$$