1.9 The Mean value theorem and criteria for differentiability

> *I turn with terror and horror from this lamentable scourge of continuous functions with no derivatives.*—Charles Hermite, in a letter to Thomas Stieltjes, 1893

In this section we discuss two applications of the mean value theorem. The first extends that theorem to functions of several variables, and the second gives a criterion for determining when a function is differentiable.

The mean value theorem for functions of several variables

The derivative measures the difference of the values of functions at different points. For functions of one variable, the mean value theorem (theorem 1.6.12) says that if $f : [a, b] \to \mathbb{R}$ is continuous, and f is differentiable on (a, b), then there exists $c \in (a, b)$ such that

$$f(b) - f(a) = f'(c)(b - a). \tag{1.9.1}$$

The analogous statement in several variables is the following.

Theorem 1.9.1 (Mean value theorem for functions of several variables). *Let $U \subset \mathbb{R}^n$ be open, $f : U \to \mathbb{R}$ be differentiable, and the segment $[\mathbf{a}, \mathbf{b}]$ joining $\mathbf{a}$ to $\mathbf{b}$ be contained in U. Then there exists $\mathbf{c} \in [\mathbf{a}, \mathbf{b}]$ such that*

$$f(\mathbf{b}) - f(\mathbf{a}) = [\mathbf{D}f(\mathbf{c})](\overrightarrow{\mathbf{b} - \mathbf{a}}). \tag{1.9.2}$$

Theorem 1.9.1: The segment $[\mathbf{a}, \mathbf{b}]$ is the image of the map

$$t \mapsto (1 - t)\mathbf{a} + t\mathbf{b},$$

for $0 \le t \le 1$.

Corollary 1.9.2. *If f is a function as defined in theorem 1.9.1, then*

$$|f(\mathbf{b}) - f(\mathbf{a})| \le \left(\sup_{\mathbf{c} \in [\mathbf{a}, \mathbf{b}]} \Big| [\mathbf{D}f(\mathbf{c})] \Big| \right) |\overrightarrow{\mathbf{b} - \mathbf{a}}|. \tag{1.9.3}$$

Why write equation 1.9.3, with the sup, rather than

$$|f(\mathbf{b}) - f(\mathbf{a})| \le |[\mathbf{D}f(\mathbf{c})]|\, |\overrightarrow{\mathbf{b} - \mathbf{a}}|,$$

which of course is also true? Using the sup means that we do not need to know the value of $\mathbf{c}$ in order to relate how fast f is changing to its derivative; we can run through all $\mathbf{c} \in [\mathbf{a}, \mathbf{b}]$ and choose the one where the derivative is greatest. This will be useful in section 2.8 when we discuss Lipschitz ratios.

Proof of corollary 1.9.2. This follows immediately from theorem 1.9.1 and proposition 1.4.11. □

Proof of theorem 1.9.1. As t varies from 0 to 1, the point $(1 - t)\mathbf{a} + t\mathbf{b}$ moves from $\mathbf{a}$ to $\mathbf{b}$. Consider the mapping $g(t) = f((1 - t)\mathbf{a} + t\mathbf{b})$. By the chain rule, g is differentiable, and by the one-variable mean value theorem, there exists t_0 such that

$$g(1) - g(0) = g'(t_0)(1 - 0) = g'(t_0). \tag{1.9.4}$$

Set $\mathbf{c} = (1 - t_0)\mathbf{a} + t_0\mathbf{b}$. By proposition 1.7.14, we can express $g'(t_0)$ in terms of the derivative of f:

$$\begin{aligned} g'(t_0) &= \lim_{s \to 0} \frac{g(t_0 + s) - g(t_0)}{s} \\ &= \lim_{s \to 0} \frac{f\big(\mathbf{c} + s(\mathbf{b} - \mathbf{a})\big) - f(\mathbf{c})}{s} = [\mathbf{D}f(\mathbf{c})](\overrightarrow{\mathbf{b} - \mathbf{a}}). \end{aligned} \tag{1.9.5}$$

So equation 1.9.4 reads

$$f(\mathbf{b}) - f(\mathbf{a}) = [\mathbf{D}f(\mathbf{c})](\overrightarrow{\mathbf{b}-\mathbf{a}}). \quad \square \tag{1.9.6}$$

Differentiability and pathological functions

Most often, the Jacobian matrix of a function is its derivative. But as we mentioned in section 1.7, there are exceptions. It is possible for all partial derivatives of f to exist, and even all directional derivatives, and yet for f not to be differentiable! In such a case the Jacobian matrix exists but does not represent the derivative.

Example 1.9.3 (A nondifferentiable function with Jacobian matrix). This happens even for the innocent-looking function

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x^2 y}{x^2 + y^2} \tag{1.9.7}$$

shown in figure 1.9.1. Actually, we should write this function

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \dfrac{x^2 y}{x^2 + y^2} & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ 0 & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{cases} \tag{1.9.8}$$

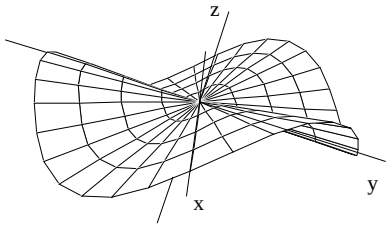

FIGURE 1.9.1.

The graph of f is made up of straight lines through the origin, so if you leave the origin in any direction, the directional derivative in that direction certainly exists. Both axes are among the lines making up the graph, so the directional derivatives in those directions are 0. But clearly there is no tangent plane to the graph at the origin.

"Vanishes" means "equals 0." "Identically" means "at every point."

You have probably learned to be suspicious of functions that are defined by different formulas for different values of the variable. In this case, the value at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is really natural, in the sense that as $\begin{pmatrix} x \\ y \end{pmatrix}$ approaches $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, the function f approaches 0. This is not one of those functions whose value takes a sudden jump; indeed, f is continuous everywhere. Away from the origin, this is obvious by corollary 1.5.30; away from the origin, f is a rational function whose denominator does not vanish. So we can compute both its partial derivatives at any point $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

That f is continuous at the origin requires a little checking, as follows. If $x^2 + y^2 = r^2$, then $|x| \leq r$ and $|y| \leq r$ so $|x^2 y| \leq r^3$. Therefore,

$$\left| f\begin{pmatrix} x \\ y \end{pmatrix} \right| \leq \frac{r^3}{r^2} = r, \quad \text{and} \quad \lim_{\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}} f\begin{pmatrix} x \\ y \end{pmatrix} = 0. \tag{1.9.9}$$

So f is continuous at the origin. Moreover, f vanishes identically on both axes, so both partial derivatives of f vanish at the origin.

So far, f looks perfectly civilized: it is continuous, and both partial derivatives exist everywhere. But consider the derivative in the direction of the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$: the directional derivative

$$\lim_{t \to 0} \frac{f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} + t\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) - f\begin{pmatrix} 0 \\ 0 \end{pmatrix}}{t} = \lim_{t \to 0} \frac{t^3}{2t^3} = \frac{1}{2}. \tag{1.9.10}$$