1.8.11 Show that if $f\begin{pmatrix} x \\ y \end{pmatrix} = \varphi\left(\frac{x+y}{x-y}\right)$ for some differentiable function $\varphi : \mathbb{R} \to \mathbb{R}$, then

$$xD_1 f + yD_2 f = 0.$$

1.8.12 True or false? Explain your answers.

a. If $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ is differentiable and $[\mathbf{Df}(\mathbf{0})]$ is not invertible, then there is no differentiable function $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^2$ such that

$$(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \mathbf{x}.$$

b. Differentiable functions have continuous partial derivatives.

1.8.13 Let $U \subset \operatorname{Mat}(n,n)$ be the set of matrices A such that $A + A^2$ is invertible. Compute the derivative of the map $F : U \to \operatorname{Mat}(n,n)$ given by $F(A) = (A + A^2)^{-1}$.

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 1.6.13) 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

$$f(b) - f(a) = f'(c)(b - a). \qquad 1.9.1$$

The analogous statement in several variables is the following.

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 $0 \le t \le 1$.

Theorem 1.9.1 (Mean value theorem for functions of several variables). *Let $U \subset \mathbb{R}^n$ be open, let $f : U \to \mathbb{R}$ be differentiable, and let the segment $[\mathbf{a}, \mathbf{b}]$ joining $\mathbf{a}$ to $\mathbf{b}$ be contained in U. Then there exists $\mathbf{c}_0 \in [\mathbf{a}, \mathbf{b}]$ such that*

$$f(\mathbf{b}) - f(\mathbf{a}) = [\mathbf{D}f(\mathbf{c}_0)](\overrightarrow{\mathbf{b} - \mathbf{a}}). \qquad 1.9.2$$

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

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

Why do we write inequality 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

$$g(1) - g(0) = g'(t_0)(1-0) = g'(t_0). \qquad 1.9.4$$

Set $\mathbf{c}_0 = (1-t_0)\mathbf{a} + t_0\mathbf{b}$. By Proposition 1.7.14, we can express $g'(t_0)$ in terms of the derivative of f:

$$\begin{aligned} g'(t_0) &= \lim_{s\to 0} \frac{g(t_0+s) - g(t_0)}{s} \\ &= \lim_{s\to 0} \frac{f\big(\mathbf{c}_0 + s(\overrightarrow{\mathbf{b}-\mathbf{a}})\big) - f(\mathbf{c}_0)}{s} = [\mathbf{D}f(\mathbf{c}_0)](\overrightarrow{\mathbf{b}-\mathbf{a}}). \end{aligned} \qquad 1.9.5$$

So equation 1.9.4 reads

$$f(\mathbf{b}) - f(\mathbf{a}) = [\mathbf{D}f(\mathbf{c}_0)](\overrightarrow{\mathbf{b}-\mathbf{a}}). \quad \square \qquad 1.9.6$$

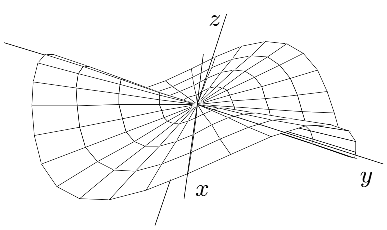

FIGURE 1.9.1.

The graph of the function f defined in equation 1.9.7 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.

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 (Nondifferentiable function with Jacobian matrix). This happens even for the innocent-looking function

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

shown in Figure 1.9.1. Actually, we should write this function as

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

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 1.5.31: 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 $|x| \le r$ and $|y| \le r$ so $|x^2 y| \le r^3$. Therefore,

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

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

"Identically" means "at every point."

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 $\begin{bmatrix}1\\1\end{bmatrix}$:

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

This is *not* what we get when we compute the same directional derivative by multiplying the Jacobian matrix of f by the vector $\begin{bmatrix}1\\1\end{bmatrix}$, as on the right side of equation 1.7.38:

$$\underbrace{\left[D_1 f\begin{pmatrix}0\\0\end{pmatrix}, D_2 f\begin{pmatrix}0\\0\end{pmatrix} \right]}_{\text{Jacobian matrix } [\mathbf{J}f(\mathbf{0})]} \begin{bmatrix}1\\1\end{bmatrix} = [0,0]\begin{bmatrix}1\\1\end{bmatrix} = 0. \qquad 1.9.11$$

Thus, by Proposition 1.7.14, f is not differentiable. △

If we change the function of Example 1.9.3, replacing the x^2y in the numerator of

$$\frac{x^2 y}{x^2+y^2}$$

by xy, then the resulting function, which we'll call g, will *not* be continuous at the origin. If $x = y$, then $g = 1/2$ no matter how close $\begin{pmatrix}x\\y\end{pmatrix}$ gets to the origin: we then have

$$g\begin{pmatrix}x\\x\end{pmatrix} = \frac{x^2}{2x^2} = \frac{1}{2}.$$

Things can get worse. The function we just discussed is continuous, but it is possible for all directional derivatives of a function to exist, and yet for the function not to be continuous, or even bounded in a neighborhood of $\mathbf{0}$. For instance, the function discussed in Example 1.5.24 is not continuous in a neighborhood of the origin; if we redefine it to be 0 at the origin, then all directional derivatives would exist everywhere, but the function would not be continuous. Exercise 1.9.2 provides another example. Thus knowing that a function has partial derivatives or directional derivatives does not tell you either that the function is differentiable or even that it is continuous.

Even knowing that a function is differentiable tells you less than you might think. The function in Example 1.9.4 has a positive derivative at x although it does not increase in a neighborhood of x!

Example 1.9.4 (A differentiable yet pathological function). Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \frac{x}{2} + x^2 \sin\frac{1}{x}, \qquad 1.9.12$$