6 Functions of Several Variables II

From Rudin (1976).

Chapter 9

2/22: **5.** Prove that to every $A \in L(\mathbb{R}^n, \mathbb{R}^1)$ corresponds a unique $\mathbf{y} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{x} \cdot \mathbf{y}$. Prove also that $\|A\| = |\mathbf{y}|$. (Hint: Under certain conditions, equality holds in the Schwarz inequality.)

Proof. Let $\mathbf{y}$ be defined as follows.

$$\mathbf{y} = \begin{bmatrix} A\mathbf{e}_1 \\ \vdots \\ A\mathbf{e}_n \end{bmatrix}$$

Then if $\mathbf{x} = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n$, we have that

$$A\mathbf{x} = a_1 \cdot A\mathbf{e}_1 + \cdots + a_n \cdot A\mathbf{e}_n = \mathbf{x} \cdot \mathbf{y}$$

as desired.

Now suppose that $\mathbf{z}$ satisfies $A\mathbf{x} = \mathbf{x} \cdot \mathbf{z}$ for all $\mathbf{x} \in \mathbb{R}^n$. Then

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= \mathbf{x} \cdot \mathbf{z} \\ \mathbf{x} \cdot (\mathbf{y} - \mathbf{z}) &= 0 \end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^n$. In particular, if $\mathbf{x} = \mathbf{y} - \mathbf{z}$, then

$$\begin{aligned} 0 &= (\mathbf{y} - \mathbf{z}) \cdot (\mathbf{y} - \mathbf{z}) \\ &= \|\mathbf{y} - \mathbf{z}\|^2 \\ \mathbf{0} &= \mathbf{y} - \mathbf{z} \\ \mathbf{z} &= \mathbf{y} \end{aligned}$$

as desired.

We also have by the Cauchy-Schwarz inequality that

$$\|A\| = \sup_{\substack{\mathbf{x} \in \mathbb{R}^n \\ |\mathbf{x}|=1}} |A\mathbf{x}| = \sup_{\substack{\mathbf{x} \in \mathbb{R}^n \\ |\mathbf{x}|=1}} |\mathbf{x} \cdot \mathbf{y}| \leq \sup_{\substack{\mathbf{x} \in \mathbb{R}^n \\ |\mathbf{x}|=1}} |\mathbf{x}| \cdot |\mathbf{y}| = |\mathbf{y}|$$

Moreover, equality holds by noting that if $\mathbf{x} = \mathbf{y}/|\mathbf{y}|$, then

$$|\mathbf{x} \cdot \mathbf{y}| = \frac{|\mathbf{y}|^2}{|\mathbf{y}|} = |\mathbf{y}|$$

so the leftmost supremum is also at least $|\mathbf{y}|$, as desired. □

6. If

$$f(x,y) = \begin{cases} 0 & (x,y) = (0,0) \\ \frac{xy}{x^2+y^2} & (x,y) \neq (0,0) \end{cases}$$

prove that $(D_1 f)(x,y)$ and $(D_2 f)(x,y)$ exist at every point of $\mathbb{R}^2$, although f is not continuous at $(0,0)$.

Proof. Let $(x,y) \in \mathbb{R}^2$ be arbitrary. We divide into two cases ($(x,y) \neq (0,0)$ and $(x,y) = (0,0)$). If $(x,y) \neq (0,0)$, then as the sum, product, and quotient of linear (hence differentiable) functions, f is differentiable at (x,y). Therefore, by Theorem 9.17, $(D_1f)(x,y)$ and $(D_2f)(x,y)$ exist. On the other hand, if $(x,y)=(0,0)$, then

$$\begin{aligned}(D_jf)(0,0) &= \lim_{t\to 0}\frac{f(t\mathbf{e}_j)-f(0,0)}{t}\\ &= \lim_{t\to 0}\frac{\frac{t\cdot 0}{t^2+0^2}-0}{t}\\ &= \lim_{t\to 0}0\\ &= 0\end{aligned}$$

However, since

$$f(t,t)=\frac{t^2}{t^2+t^2}=\frac{1}{2}$$

for all t (notably arbitrarily small t), we find points $(x,y)\in\mathbb{R}^2$ arbitrarily close to $(0,0)$ for which $f(x,y)=1/2$. Therefore, f is not continuous at $(0,0)$. ☐

7. Suppose that f is a real-valued function defined in an open set $E\subset\mathbb{R}^n$, and that the partial derivatives $D_1f,\dots,D_nf$ are bounded on E. Prove that f is continuous in E. (Hint: Proceed as in the proof of Theorem 9.21.)

Proof. To prove that f is continuous at an arbitrary $\mathbf{x}\in E$, it will suffice to show that for all $\epsilon>0$, there exists a $\delta>0$ such that if $\mathbf{x}+\mathbf{h}\in E$ and $|\mathbf{h}|<\delta$, then $|f(\mathbf{x}+\mathbf{h})-f(\mathbf{x})|<\epsilon$. Let $\epsilon>0$ be arbitrary. Since each D_jf is bounded, we know that there exist $M_1,\dots,M_n$ such that $|D_jf|\leq M_j$ $(j=1,\dots,n)$. Define $M=\max M_j$. Choose $\delta=\epsilon/nM$. Let $\mathbf{h}$ be such that $\mathbf{x}+\mathbf{h}\in E$ and $|\mathbf{h}|<\delta$. Supposing $\mathbf{h}=\sum h_j\mathbf{e}_j$, put $\mathbf{v}_0=\mathbf{0}$ and $\mathbf{v}_k=h_1\mathbf{e}_1+\cdots+h_k\mathbf{e}_k$ $(k=1,\dots,n)$. Then, letting $\theta_j\in(0,1)$ for all j, we have that

$$\begin{aligned}|f(\mathbf{x}+\mathbf{h})-f(\mathbf{x})| &= \left|\sum_{j=1}^n[f(\mathbf{x}+\mathbf{v}_j)-f(\mathbf{x}+\mathbf{v}_{j-1})]\right|\\ &\leq \sum_{j=1}^n|f(\mathbf{x}+\mathbf{v}_j)-f(\mathbf{x}+\mathbf{v}_{j-1})|\\ &= \sum_{j=1}^n[h_j(D_jf)(\mathbf{x}+\mathbf{v}_{j-1}+\theta_jh_j\mathbf{e}_j)] && \text{MVT}\\ &\leq M\sum_{j=1}^n h_j\\ &\leq nM|\mathbf{h}|\\ &< \epsilon\end{aligned}$$

as desired. ☐

8. Suppose that f is a differentiable real function in an open set $E\subset\mathbb{R}^n$, and that f has a local maximum at a point $\mathbf{x}\in E$. Prove that $f'(\mathbf{x})=0$.

Proof. To prove that $f'(\mathbf{x})=0$, Theorem 9.17 tells us that it will suffice to show that $(D_jf)(\mathbf{x})=0$ for all $1\leq j\leq n$. Let j be an arbitrary natural number between 1 and n, inclusive. Define $U_j=(\mathbf{x}+\operatorname{span}(\mathbf{e}_j))\cap E$ and let $\mathbf{x}=(x_1,\dots,x_n)$. Then by definition, $(D_jf)(\mathbf{x})=(f|_{U_j})'(x_j)$ and $f|_{U_j}(x_j)$ is a local maximum of $f|_{U_j}$. It follows by Theorem 5.8 that

$$(D_jf)(\mathbf{x})=(f|_{U_j})'(x_j)=0$$

as desired. ☐

10. If f is a real function defined in a convex open set $E \subset \mathbb{R}^n$, such that $(D_1 f)(\mathbf{x}) = 0$ for every $\mathbf{x} \in E$, prove that $f(\mathbf{x})$ depends only on $x_2, \dots, x_n$. Show that the convexity of E can be replaced by a weaker condition, but that some condition is required. For example, if $n = 2$ and E is shaped like a horseshoe, the statement may be false.

Proof. To prove that $f(\mathbf{x})$ depends only on $x_2, \dots, x_n$, it will suffice to show that if $\mathbf{x}, \mathbf{y} \in E$ are such that $\mathbf{x} = (x, x_2, \dots, x_n)$ and $\mathbf{y} = (y, x_2, \dots, x_n)$, then $f(\mathbf{x}) = f(\mathbf{y})$. Let $\mathbf{x}, \mathbf{y} \in E$ be arbitrary points that satisfy the previous condition. Since E is convex, $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in E$ for all $\lambda \in (0, 1)$. Define $U = \{\lambda\mathbf{x} + (1-\lambda)\mathbf{y} : 0 \leq \lambda \leq 1\}$. Then since $(f|_U)'(t_1) = (D_1 f)(\mathbf{t}) = 0$ on U, Theorem 5.17b implies that $f|_U$ is constant. In particular, since $\mathbf{x}, \mathbf{y} \in U$, $f(\mathbf{x}) = f(\mathbf{y})$, as desired.

The following weaker condition will suffice: E open is a set such that $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in E$ for all $0 < \lambda < 1$ and $\mathbf{x}, \mathbf{y} \in E$ *that satisfy* $\mathbf{x} - \mathbf{y} \in \operatorname{span}(\mathbf{e}_1)$. Note that the above proof still works with this condition in place of convexity because the arbitrary $\mathbf{x}, \mathbf{y}$ to which we applied the definition of convexity satisfy $\mathbf{x} - \mathbf{y} \in \operatorname{span}(\mathbf{e}_1)$.

If we do not assert any condition, we may consider as a counterexample the subset of $\mathbb{R}^2$

$$E = ((-2, -1) \times (1, 2)) \cup ((1, 2) \times (1, 2))$$

and the function $f : E \to \mathbb{R}$ defined by

$$f(x, y) = \operatorname{sgn}(x) + y$$

Since $f(x, y) = -1 + y$ for $x \in (-2, -1) \times (1, 2)$ and $f(x, y) = 1 + y$ for $x \in (1, 2) \times (1, 2)$, clearly $(D_1 f)(\mathbf{x}) = 0$ for all $\mathbf{x} \in E$. Yet $f(-1.5, 0) = -1$ and $f(1.5, 0) = 1$. □

11. If f and g are differentiable real functions in $\mathbb{R}^n$, prove that

$$\nabla(fg) = f\nabla g + g\nabla f$$

and that

$$\nabla\left(\frac{1}{f}\right) = -\frac{\nabla f}{f^2}$$

wherever $f \neq 0$.

Proof. Since, by definition, partial derivatives only take limits over linear subsets of $\mathbb{R}^n$, i.e., ones that are isomorphic to $\mathbb{R}$, they behave like the one-dimensional derivatives of f restricted to such linear subsets. More precisely, $D_j f$ behaves like the derivative of $(f|_{N_r(\mathbf{x}) \cap \operatorname{span}(\mathbf{e}_j)})'$ where $\mathbf{x}$ is the point at which we're taking the derivative and r is taken to be small enough so that $N_r(\mathbf{x})$ is a subset of the open domain of f. In particular, this means that partial derivatives obey the usual sum, product, and quotient rules, as well as other rules carried over from the analysis of functions of a single variable. Thus, we have that for any $\mathbf{x} \in \mathbb{R}^n$,

$$\begin{aligned}
(\nabla(fg))(\mathbf{x}) &= \sum_{j=1}^{n} (D_j(fg))(\mathbf{x})\mathbf{e}_j \\
&= \sum_{j=1}^{n} (fD_j g + D_j f g)(\mathbf{x})\mathbf{e}_j \\
&= f(\mathbf{x}) \sum_{j=1}^{n} (D_j g)(\mathbf{x})\mathbf{e}_j + g(\mathbf{x}) \sum_{j=1}^{n} (D_j f)(\mathbf{x})\mathbf{e}_j \\
&= f(\mathbf{x})(\nabla g)(\mathbf{x}) + (\nabla f)(\mathbf{x})g(\mathbf{x}) \\
&= (f\nabla g + g\nabla f)(\mathbf{x})
\end{aligned}$$

and symmetrically for $\nabla(1/f)$. □

17. Let $\mathbf{f} = (f_1, f_2)$ be the mapping of $\mathbb{R}^2$ into $\mathbb{R}^2$ given by

$$f_1(x,y) = \mathrm{e}^x \cos y \qquad\qquad f_2(x,y) = \mathrm{e}^x \sin y$$

(a) What is the range of f?

Proof. The range of f is $\mathbb{R}^2 \setminus \{\mathbf{0}\}$.
Both trigonometric functions have range $[-1,1]$ and e^x has range $(0,\infty)$, so we can make e^x arbitrarily large and then take y such that such that the trigonometric function equals 1 or -1. However, since $\cos y = \pm 1$ when $\sin y = 0$ and vice versa, we can never achieve $(0,0)$. $\square$

(b) Show that the Jacobian of f is not zero at any point of $\mathbb{R}^2$. Thus, every point of $\mathbb{R}^2$ has a neighborhood in which f is one-to-one. Nevertheless, f is not one-to-one on $\mathbb{R}^2$.

Proof. Let $(x,y) \in \mathbb{R}^2$ be arbitrary. We divide into two cases ($y = n\pi$ and $y \neq n\pi$ $[n \in \mathbb{Z}]$). If $y = n\pi$, then since $(\sin y)' = \cos y$, $D_2 f_2 = \mathrm{e}^x \cos y$, $\mathrm{e}^x \neq 0$ for all x, and $\cos n\pi = \pm 1$ ($n \in \mathbb{Z}$), we have that $(D_2 f_2)(x,y) \neq 0$. If $y \neq n\pi$, then since $(\mathrm{e}^x)' = \mathrm{e}^x$ (Theorem 8.6b), $D_1 f_2 = \mathrm{e}^x \sin y$, $\mathrm{e}^x \neq 0$ for all x, and $\sin y \neq 0$ for all $y \neq n\pi$, we have that $(D_1 f_2)(x,y) \neq 0$.
It follows by Theorem 9.24 that every point of $\mathbb{R}^2$ has a neighborhood in which f is 1-1.
Since the trigonometric functions are periodic with period 2π, $f(0,0) = f(0,2\pi)$, for instance. $\square$

(c) Put $\mathbf{a} = (0, \pi/3)$, $\mathbf{b} = f(\mathbf{a})$, and let $\mathbf{g}$ be the continuous inverse of $\mathbf{f}$, defined in a neighborhood of $\mathbf{b}$, such that $\mathbf{g}(\mathbf{b}) = \mathbf{a}$. Find an explicit formula for $\mathbf{g}$, compute $\mathbf{f}'(\mathbf{a})$ and $\mathbf{g}'(\mathbf{b})$, and verify that

$$\mathbf{g}'(\mathbf{b}) = [\mathbf{f}'(\mathbf{g}(\mathbf{b}))]^{-1}$$

Proof. Let $U = \mathbb{R} \times (0, \pi/2)$. Then $V = f(U) = (0,\infty)^2$. Thus, we may define $\mathbf{g} : V \to U$ by

$$g_1(x,y) = \frac{1}{2}\ln\big(x^2 + y^2\big) \qquad\qquad g_2(x,y) = \tan^{-1}\left(\frac{y}{x}\right)$$

Under this definition, we can easily see that if $(x,y) \in U$, then

$$\begin{aligned}
g(f(x,y)) &= g(\mathrm{e}^x \cos y, \mathrm{e}^x \sin y) \\
&= \left(\frac{1}{2}\ln\left((\mathrm{e}^x \cos y)^2 + (\mathrm{e}^x \sin y)^2\right), \tan^{-1}\left(\frac{\mathrm{e}^x \sin y}{\mathrm{e}^x \cos y}\right)\right) \\
&= \left(\frac{1}{2}\ln\left(\mathrm{e}^{2x}(\cos^2 y + \sin^2 y)\right), \tan^{-1}(\tan y)\right) \\
&= (x,y)
\end{aligned}$$

as desired. We can also compute that

$$\begin{aligned}
\mathbf{f}'(\mathbf{a}) &= \begin{bmatrix} (D_1 f_1)(\mathbf{a}) & (D_2 f_1)(\mathbf{a}) \\ (D_1 f_2)(\mathbf{a}) & (D_2 f_2)(\mathbf{a}) \end{bmatrix} \\
&= \begin{bmatrix} \mathrm{e}^x \cos y & -\mathrm{e}^x \sin y \\ \mathrm{e}^x \sin y & \mathrm{e}^x \cos y \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}
\end{aligned}
\qquad\qquad
\begin{aligned}
\mathbf{g}'(\mathbf{b}) &= \begin{bmatrix} (D_1 f_1)(\mathbf{a}) & (D_2 f_1)(\mathbf{a}) \\ (D_1 f_2)(\mathbf{a}) & (D_2 f_2)(\mathbf{a}) \end{bmatrix} \\
&= \begin{bmatrix} \frac{x}{x^2+y^2} & \frac{y}{x^2+y^2} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}
\end{aligned}$$

It follows from the definition of matrix multiplication that

$$\mathbf{g}'(\mathbf{b})\mathbf{f}'(\mathbf{g}(\mathbf{b})) = \mathbf{g}'(\mathbf{b})\mathbf{f}'(\mathbf{a}) = I \qquad\qquad \mathbf{f}'(\mathbf{g}(\mathbf{b}))\mathbf{g}'(\mathbf{b}) = \mathbf{f}'(\mathbf{a})\mathbf{g}'(\mathbf{b}) = I$$

as desired. $\square$

(d) What are the images under $\mathbf{f}$ of lines parallel to the coordinate axes?

Proof. Let $U \subset \mathbb{R}^2$ denote a line parallel to the x-axis and a distance t away, and let $V \subset \mathbb{R}^2$ denote a line parallel to the y-axis and a distance t away. Formally,

$$U = \{(t, y) : y \in \mathbb{R}\} \qquad\qquad V = \{(x, t) : x \in \mathbb{R}\}$$

where $t \in \mathbb{R}$.

We first describe $\mathbf{f}(U)$. Since $x = t$ for all $(x, y) \in U$, $\mathrm{e}^x = \mathrm{e}^t$. Since y ranges over $\mathbb{R}$ as (x, y) ranges over U, the image of U under $\cos y$ is $[-1, 1]$. Similarly, the image of U under $\sin y$ is $[-1, 1]$. Thus, $\mathbf{f}(U) = (f_1(U), f_2(U))$ where the two components are given by

$$f_1(U) = [-\mathrm{e}^t, \mathrm{e}^t] \qquad\qquad f_2(U) = [-\mathrm{e}^t, \mathrm{e}^t]$$

We now describe $\mathbf{f}(V)$. Since x ranges over $\mathbb{R}$ as (x, y) ranges over V, the image of V under e^x is $(0, \infty)$. Since $y = t$ for all $(x, y) \in U$, $\cos y = 0$ if $t = \frac{\pi}{2} + n\pi$, $\cos y > 0$ if $t \in (-\frac{\pi}{2} + 2\pi n, \frac{\pi}{2} + 2\pi n)$, and $\cos y < 0$ if $t \in (\frac{\pi}{2} + 2\pi n, \frac{3\pi}{2} + 2\pi n)$ $(n \in \mathbb{Z})$. Similarly, $\sin y = 0$ if $t = n\pi$, $\sin y > 0$ if $t \in (2n\pi, (2n+1)\pi)$, and $\sin y < 0$ if $t \in ((2n-1)\pi, 2n\pi)$ $(n \in \mathbb{Z})$. Thus, $\mathbf{f}(V) = (f_1(V), f_2(V))$ where the two components are given by

$$f_1(V) = \begin{cases} \{0\} & t = \frac{\pi}{2} + n\pi \\ (0, \infty) & -\frac{\pi}{2} + 2\pi n < t < \frac{\pi}{2} + 2\pi n \\ (0, -\infty) & \frac{\pi}{2} + 2\pi n < t < \frac{3\pi}{2} + 2\pi n \end{cases} \quad f_2(V) = \begin{cases} \{0\} & t = n\pi \\ (0, \infty) & 2n\pi < t < (2n+1)\pi \\ (0, -\infty) & (2n-1)\pi < t < 2n\pi \end{cases}$$

where $n \in \mathbb{Z}$ in every occurrence. $\square$