Exercise Problem Sets 14

Dec. 24. 2021

Problem 1. Investigate the differentiability of

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

Solution. First we note that

$$f_x(0,0) = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h} = 0 \qquad \text{and} \qquad f_y(0,0) = \lim_{k\to 0} \frac{f(0,k)-f(0,0)}{k} = 0\,.$$

For $(x,y) \neq (0,0)$,

$$\frac{\big|f(x,y)-f(0,0)-f_x(0,0)x-f_y(0,0)y\big|}{\sqrt{x^2+y^2}} = \frac{xy}{x^2+y^2}$$

whose limit, as $(x,y) \to (0,0)$, does not exist. Therefore, f is not differentiable at $(0,0)$.

On the other hand, for $(x,y) \neq (0,0)$,

$$f_x(x,y) = \frac{y\sqrt{x^2+y^2} - \dfrac{x^2y}{\sqrt{x^2+y^2}}}{x^2+y^2} = \frac{y^3}{(x^2+y^2)^{\frac{3}{2}}}$$

and similarly, $f_y(x,y) = \dfrac{x^3}{(x^2+y^2)^{\frac{3}{2}}}$. Clearly f_x and f_y are continuous on $\mathbb{R}^2\backslash\{(0,0)\}$; thus f is differentiable on $\mathbb{R}^2\backslash\{(0,0)\}$. □

Problem 2. Investigate the differentiability of

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

Solution. For $x+y^2 \neq 0$,

$$f_x(x,y) = \frac{y(x+y^2)-xy}{(x+y^2)^2} = \frac{y^3}{(x+y^2)^2} \quad \text{and} \quad f_y(x,y) = \frac{x(x+y^2)-2xy^2}{(x+y^2)^2} = \frac{x^2-xy^2}{(x+y^2)^2}\,.$$

Clearly f_x and f_y are continuous on $\mathbb{R}^2\backslash\big\{(x,y)|x+y^2=0\big\}$; thus f is differentiable at point (x,y) satisfying $x+y^2 \neq 0$ (by Theorem 5.40 in the lecture note).

Now we consider the differentiability of f at (a,b) when $a+b^2=0$. First we note that

$$f_x(a,b) = \lim_{h\to 0} \frac{f(a+h,b)-f(a,b)}{h} = \lim_{h\to 0} \frac{(a+h)b}{h(a+h+b^2)} = \begin{cases} 0 & (a,b) = (0,0)\,, \\ \text{D.N.E.} & (a,b) \neq (0,0)\,; \end{cases}$$

thus f is not differentiable at (a,b) if $a+b^2=0$ and $(a,b) \neq (0,0)$ (because of Theorem 5.27 in the lecture note).

Finally we justify the differentiability of f at $(0,0)$. Note that

$$f_y(0,0) = \lim_{k\to 0} \frac{f(0,k) - f(0,0)}{k} = 0\,.$$

For $x = y^2$ with $y \neq 0$, we have

$$\frac{\big|f(x,y) - f(0,0) - f_x(0,0)x - f_y(0,0)y\big|}{\sqrt{x^2+y^2}} = \frac{|y^3|}{2y^2\sqrt{y^4+y^2}} = \frac{1}{2\sqrt{y^2+1}}$$

whose limit, as $y \to 0$, cannot be zero; thus

$$\lim_{(x,y)\to(0,0)} \frac{\big|f(x,y) - f(0,0) - f_x(0,0)x - f_y(0,0)y\big|}{\sqrt{x^2+y^2}} \neq 0\,.$$

Therefore, f is not differentiable at $(0,0)$. □

Problem 3. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\dfrac{1}{\sqrt{x^2+y^2}} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

Discuss the differentiability of f. Find $(\nabla f)(x,y)$ at points of differentiability.

Solution. If $(x,y) \neq (0,0)$, then

$$\begin{aligned} f_x(x,y) &= 2x\sin\frac{1}{\sqrt{x^2+y^2}} + (x^2+y^2)\cos\frac{1}{\sqrt{x^2+y^2}} \cdot \frac{-x}{(x^2+y^2)^{\frac{3}{2}}} \\ &= 2x\sin\frac{1}{\sqrt{x^2+y^2}} - \frac{1}{\sqrt{x^2+y^2}}\cos\frac{1}{\sqrt{x^2+y^2}} \end{aligned}$$

and similarly,

$$f_y(x,y) = 2y\sin\frac{1}{\sqrt{x^2+y^2}} - \frac{1}{\sqrt{x^2+y^2}}\cos\frac{1}{\sqrt{x^2+y^2}}\,.$$

Clearly f_x and f_y are continuous on $\mathbb{R}^2\backslash\{(0,0)\}$; thus f is differentiable at point $(x,y) \neq (0,0)$ (by Theorem 5.40 in the lecture note).

Now we justify the differentiability of f at $(0,0)$. First we compute $f_x(0,0)$ and $f_y(0,0)$ and find that

$$f_x(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} h\sin\frac{1}{|h|} = 0$$

and

$$f_y(0,0) = \lim_{k\to 0} \frac{f(0,k) - f(0,0)}{k} = \lim_{k\to 0} k\sin\frac{1}{|k|} = 0\,,$$

where the limits above are obtained by the Sandwich Lemma. For $(x,y) \neq (0,0)$, we have

$$\frac{\big|f(x,y) - f(0,0) - 0\cdot(x-0) - 0\cdot(y-0)\big|}{\sqrt{x^2+y^2}} = \sqrt{x^2+y^2}\sin\frac{1}{\sqrt{x^2+y^2}} \leqslant \sqrt{x^2+y^2}\,;$$

thus the Sandwich Lemma implies that

$$\lim_{(x,y)\to(0,0)} \frac{\big|f(x,y) - f(0,0) - 0\cdot(x-0) - 0\cdot(y-0)\big|}{\sqrt{x^2+y^2}} = 0\,.$$

Therefore, f is also differentiable at $(0,0)$; thus f is differentiable on $\mathbb{R}^2$. □

Problem 4. Let $X = \mathcal{M}_{n\times m}$, the collection of all $n \times m$ real matrices, equipped with the Frobenius norm $\|\cdot\|_F$ introduced in Problem 5 of Exercise 6, and $f : X \to \mathbb{R}$ be defined by $f(A) = \|A\|_F^2$. Show that f is differentiable on X and find $(Df)(A)$ for $A \in X$.

Proof. First we note that $f(A) = \mathrm{tr}(AA^{\mathrm{T}})$, where $\mathrm{tr}(M)$ denotes the trace of M is M is a square matrix. Let $A = [a_{ij}] \in X$. Then for $\delta A \in X$, we have

$$\begin{aligned} f(A+\delta A) - f(A) &= \mathrm{tr}\big[(A+\delta A)(A+\delta A)^{\mathrm{T}}\big] - \mathrm{tr}(AA^{\mathrm{T}}) \\ &= \mathrm{tr}\big(AA^{\mathrm{T}} + A\delta A^{\mathrm{T}} + \delta A A^{\mathrm{T}} + \delta A \delta A^{\mathrm{T}}\big) - \mathrm{tr}(AA^{\mathrm{T}}) \\ &= \mathrm{tr}(A\delta A^{\mathrm{T}}) + \mathrm{tr}(\delta A A^{\mathrm{T}}) + \mathrm{tr}(\delta A\delta A^{\mathrm{T}})\,. \end{aligned}$$

Define $L_A : X \to \mathbb{R}$ by $L(B) = \mathrm{tr}(AB^{\mathrm{T}}) + \mathrm{tr}(BA^{\mathrm{T}})$. Then Problem 2 of Exercise 13 shows that $L \in \mathscr{B}(X, \mathbb{R})$. Therefore, by the fact that

$$\lim_{\delta A\to 0} \frac{\big|f(A+\delta A) - f(A) - L_A(\delta A)\big|}{\|\delta A\|_F} = \lim_{\delta A\to 0} \frac{\big|\mathrm{tr}(\delta A\delta A^{\mathrm{T}})\big|}{\|\delta A\|_F} = \lim_{\delta A\to 0} \frac{\|\delta A\|_F^2}{\|\delta A\|_F} = \lim_{\delta A\to 0} \|\delta A\|_F = 0\,,$$

we conclude that f is differentiable at A and $(Df)(A) = L_A$. □

Problem 5. Let $\|\cdot\|_F$ denote the Frobenius norm of matrices given in Problem 5 of Exercise 6. For an $m \times n$ matrix $A = [a_{ij}]$, we look for an $m \times k$ matrix $C = [c_{ij}]$ and an $k \times n$ matrix $R = [r_{ij}]$, where $1 \leqslant k \leqslant \min\{m, n\}$, such that $\|A - CR\|_F^2$ is minimized. This is to minimize the function

$$f(C,R) = \|A - CR\|_F^2 = \mathrm{tr}((A-CR)(A-CR)^{\mathrm{T}}) = \sum_{i=1}^{n}\sum_{j=1}^{m}\Big(a_{ij} - \sum_{\ell=1}^{k} c_{i\ell} r_{\ell j}\Big)^2\,.$$

Show that if $C \in \mathbb{R}^{m\times k}$ and $R \in \mathbb{R}^{k\times n}$ minimize f, then C, R satisfy

$$(A - CR)R^{\mathrm{T}} = 0 \qquad \text{and} \qquad C^{\mathrm{T}}(A - CR) = 0\,.$$

Proof. Since

$$\begin{aligned} f(C,R) &= \mathrm{tr}((A-CR)(A^{\mathrm{T}} - R^{\mathrm{T}}C^{\mathrm{T}})) \\ &= \mathrm{tr}(AA^{\mathrm{T}}) - \mathrm{tr}(CRA^{\mathrm{T}}) - \mathrm{tr}(AR^{\mathrm{T}}C^{\mathrm{T}}) + \mathrm{tr}(CRR^{\mathrm{T}}C^{\mathrm{T}}) \\ &= \mathrm{tr}(AA^{\mathrm{T}}) - 2\mathrm{tr}(AR^{\mathrm{T}}C^{\mathrm{T}}) + \mathrm{tr}(CRR^{\mathrm{T}}C^{\mathrm{T}})\,, \end{aligned}$$

we find that

$$\begin{aligned} (Df)(C,R)(\delta C, \delta R) &= -2\mathrm{tr}(AR^{\mathrm{T}}(\delta C)^{\mathrm{T}}) + \mathrm{tr}((\delta C)RR^{\mathrm{T}}C^{\mathrm{T}}) + \mathrm{tr}(CRR^{\mathrm{T}}(\delta C)^{\mathrm{T}}) \\ &\quad - 2\mathrm{tr}(A(\delta R)^{\mathrm{T}}C^{\mathrm{T}}) + \mathrm{tr}(C(\delta R)R^{\mathrm{T}}C^{\mathrm{T}}) + \mathrm{tr}(CR(\delta R)^{\mathrm{T}}C^{\mathrm{T}}) \\ &= -2\mathrm{tr}(AR^{\mathrm{T}}(\delta C)^{\mathrm{T}}) + 2\mathrm{tr}(CRR^{\mathrm{T}}(\delta C)^{\mathrm{T}}) \\ &\quad - 2\mathrm{tr}(A(\delta R)^{\mathrm{T}}C^{\mathrm{T}}) + 2\mathrm{tr}(CR(\delta R)^{\mathrm{T}}C^{\mathrm{T}}) \\ &= -2\mathrm{tr}((A-CR)R^{\mathrm{T}}(\delta C)^{\mathrm{T}}) - 2\mathrm{tr}((A-CR)(\delta R)^{\mathrm{T}}C^{\mathrm{T}}) \\ &= -2\mathrm{tr}((A-CR)R^{\mathrm{T}}(\delta C)^{\mathrm{T}}) - 2\mathrm{tr}(C^{\mathrm{T}}(A-CR)(\delta R)^{\mathrm{T}})\,, \end{aligned}$$

where we have used that $\mathrm{tr}(PQ) = \mathrm{tr}(QP)$ to obtain the last equality. By the fact that $\mathrm{tr}(PQ) = 0$ for all Q if and only if $P = 0$, we conclude that if C, R minimize f, then it holds the desired identity.
□

Alternative proof. If f attains its minimum at $C = [c_{ij}]$ and $R = [r_{ij}]$, then

$$\frac{\partial f}{\partial c_{pq}}(C,R) = 2\sum_{i=1}^{n}\sum_{j=1}^{m}\Big[\Big(a_{ij} - \sum_{\ell=1}^{k} c_{i\ell}r_{\ell j}\Big)\sum_{s=1}^{k}\delta_{ip}\delta_{sq}r_{sj}\Big] = 0\,,$$
$$\frac{\partial f}{\partial r_{pq}}(C,R) = 2\sum_{i=1}^{n}\sum_{j=1}^{m}\Big[\Big(a_{ij} - \sum_{\ell=1}^{k} c_{i\ell}r_{\ell j}\Big)\sum_{s=1}^{k}\delta_{sp}\delta_{jq}c_{is}\Big] = 0\,,$$

where $\delta_{\cdot\cdot}$ is the Kronecker delta. Therefore, for all p, q,

$$\sum_{j=1}^{m}\Big(a_{pj} - \sum_{\ell=1}^{k} c_{p\ell}r_{\ell j}\Big)r_{qj} = \sum_{i=1}^{n}\Big(a_{iq} - \sum_{\ell=1}^{k} c_{i\ell}r_{\ell q}\Big)c_{ip} = 0$$

which implies desired identity. □

Problem 6. Let $X = \mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$ equipped with norm $\|\cdot\|$, and $f : \mathrm{GL}(n) \to \mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$ be defined by $f(L) = L^{-2} \equiv L^{-1} \circ L^{-1}$. Show that f is differentiable on $\mathrm{GL}(n)$ and find $(Df)(L)$ for $L \in \mathrm{GL}(n)$.

Proof. Let $L \in \mathrm{GL}(n)$. By the fact that

$$K^{-1} - L^{-1} = -K^{-1}(K-L)L^{-1} \text{ and } K^{-2} - L^{-2} = -K^{-2}(K-L)L^{-1} - K^{-1}(K-L)L^{-2}\,,$$

we have

$$\begin{aligned} K^{-2} - L^{-2} &= -\big[L^{-2} - K^{-2}(K-L)L^{-1} - K^{-1}(K-L)L^{-2}\big](K-L)L^{-1} \\ &\quad - \big[L^{-1} - K^{-1}(K-L)L^{-1}\big](K-L)L^{-2} \\ &= -L^{-2}(K-L)L^{-1} - L^{-1}(K-L)L^{-2} + K^{-2}(K-L)L^{-1}(K-L)L^{-1} \\ &\quad + K^{-1}(K-L)L^{-2}(K-L)L^{-1} + K^{-1}(K-L)L^{-1}(K-L)L^{-2}\,; \end{aligned}$$

thus

$$\begin{aligned} &\|K^{-2} - L^{-2} + L^{-2}(K-L)L^{-1} + L^{-1}(K-L)L^{-2}\| \\ &\quad \leqslant \Big[\|K^{-2}\|\|L^{-1}\|^2 + 2\|K^{-1}\|\|L^{-1}\|\|L^{-2}\|\Big]\|K-L\|^2\,. \end{aligned} \tag{$\star$}$$

This motivates us to define $(Df)(L) \in \mathscr{B}(X, X)$ by

$$(Df)(L)(H) = -L^{-2}HL^{-1} - L^{-1}HL^{-2} \qquad \forall\, H \in X\,, \tag{$\diamond$}$$

and ($\star$) implies that

$$\lim_{K\to L}\frac{\|f(K) - f(L) - (Df)(L)(K-L)\|}{\|K-L\|} = 0\,.$$

Therefore, f is differentiable on $\mathrm{GL}(n)$, and $(Df)(L)$ is given by ($\diamond$). □

Problem 7. Let $X = \mathscr{C}([-,1,1];\mathbb{R})$ and $\|\cdot\|_X$ be defined by $\|f\|_X = \max\limits_{x\in[-1,1]}\big|f(x)\big|$, and $(Y, \|\cdot\|_Y) = (\mathbb{R}, |\cdot|)$. Consider the map $\delta : X \to \mathbb{R}$ be defined by $\delta(f) = f(0)$. Show that δ is differentiable on X. Find $(D\delta)(f)$ (for $f \in \mathscr{C}([-1,1];\mathbb{R})$).

Proof. Let $f \in X$ be given. For $h \in X$, we have

$$\delta(f+h) - \delta f = \big(f(0) + h(0)\big) - f(0) = h(0) = \delta h\,;$$

thus we expect that $(D\delta)(f)(h) = \delta h$. We first show that $\delta \in \mathscr{B}(X, \mathbb{R})$.

1. For linearity, for $h_1, h_2 \in X$ and $c \in \mathbb{R}$, we have

$$\delta(ch_1 + h_2) = \big(ch_1 + h_2\big)(0) = ch_1(0) + h_2(0) = c\delta h_1 + \delta h_2\,.$$

2. For boundedness, if $\|h\|_X = 1$, then $\max\limits_{x\in[-1,1]} \big|h(x)\big| = 1$ so that

$$|\delta h| = \big|h(0)\big| \leqslant \max_{x\in[-1,1]} \big|h(x)\big| = 1 < \infty\,.$$

Having established that $\delta \in \mathscr{B}(X, \mathbb{R})$, we note that

$$\lim_{h\to 0} \frac{\big|\delta(f+h) - \delta f - \delta h\big|}{\|h\|_X} = \lim_{h\to 0} \frac{0}{\|h\|_X} = 0\,;$$

thus δ is differentiable at f (for all $f \in X$), and $(D\delta)(f) = \delta$ for all $f \in X$. □

Problem 8. Let $X = \mathscr{C}([a, b]; \mathbb{R})$ and $\|\cdot\|_2$ be the norm induced by the inner product $\langle f, g\rangle = \displaystyle\int_a^b f(x)g(x)\,dx$. Define $I : X \to X$ by

$$I(f)(x) = \int_a^x f(t)^2\,dt \qquad \forall\, x \in [a, b]\,.$$

Show that I is differentiable on X, and find $(DI)(f)$.

Proof. Let $f \in X$ be given. For $h \in X$,

$$I(f+h)(x) - I(f)(x) = \int_a^x \big(f(t) + h(t)\big)^2\,dt - \int_a^x f(t)^2\,dt = \int_a^x \big[2f(t)h(t) + h(t)^2\big]\,dt\,; \tag{$\star\star$}$$

thus we expect that

$$(DI)(f)(h)(x) = 2\int_a^x f(t)h(t)\,dt\,. \tag{$\diamond\diamond$}$$

Define L by $(Lh)(x) = 2\displaystyle\int_a^x f(t)h(t)\,dt$.
Claim: $L \in \mathscr{B}(X, X)$.

1. For linearity, let $h_1, h_2 \in X$ and $c \in \mathbb{R}$. Then

$$L(ch_1 + h_2)(x) = 2\int_a^x f(t)\big(ch_1(t) + h_2(t)\big)\,dt = 2c\int_a^x f(t)h_1(t)\,dt + 2\int_a^x f(t)h_2(t)\,dt$$

which shows that $L(ch_1 + h_2) = cL(h_1) + L(h_2)$.

2. Note that by the Cauchy-Schwarz inequality,

$$\Big|\int_a^x f(t)h(t)\,dt\Big| \leqslant \int_a^b |f(t)||h(t)|\,dt \leqslant \|f\|_2\|h\|_2\,;$$

thus for $\|h\|_2 = 1$,

$$\|L(h)\|_2 = \Big[\int_a^b \Big(\int_a^x f(t)h(t)\,dt\Big)^2 dx\Big]^{\frac{1}{2}} \leqslant \Big(\int_a^b \|f\|_2^2\|h\|_2^2\,dx\Big)^{\frac{1}{2}} \leqslant \sqrt{b-a}\|f\|_2\,.$$

Therefore,

$$\|L\| = \sup_{\|h\|_2=1} \|L(h)\|_2 \leqslant \sqrt{b-a}\|f\|_2 < \infty$$

which shows that L is bounded.

Finally, using $(\star\star)$ we obtain that

$$\begin{aligned}\big\|I(f+h) - I(f) - L(h)\big\|_2 &= \Big[\int_a^b \Big(\int_a^x h(t)^2\,dt\Big)^2 dx\Big]^{\frac{1}{2}} \leqslant \Big[\int_a^b \Big(\int_a^b h(t)^2\,dt\Big)^2 dx\Big]^{\frac{1}{2}}\\ &= \Big[\int_a^b \|h\|_2^4\,dx\Big]^{\frac{1}{2}} = \sqrt{b-a}\|h\|_2^2\,;\end{aligned}$$

thus

$$\lim_{h\to 0} \frac{\big\|I(f+h) - I(f) - (DI)(f)(h)\big\|_2}{\|h\|_2} = 0\,.$$

Therefore, I is differentiable at f for all $f \in X$ and $(DI)(f)$ is given by $(\diamond\diamond)$. □

Problem 9. Let $r > 0$ and $\alpha > 1$. Suppose that $f : B(0,r) \to \mathbb{R}$ satisfies $|f(x)| \leqslant \|x\|^\alpha$ for all $x \in B(0,r)$. Show that f is differentiable at 0. What happens if $\alpha = 1$?

Problem 10. Suppose that $f, g : \mathbb{R} \to \mathbb{R}^m$ are differentiable at a and there is a $\delta > 0$ such that $g(x) \neq 0$ for all $0 < |x-a| < \delta$. If $f(a) = g(a) = 0$ and $(Dg)(a) \neq 0$, show that

$$\lim_{x\to a} \frac{\|f(x)\|}{\|g(x)\|} = \frac{\|(Df)(a)\|}{\|(Dg)(a)\|}\,.$$

Problem 11. Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$. Suppose that the partial derivatives $\dfrac{\partial f}{\partial x_1}, \cdots, \dfrac{\partial f}{\partial x_n}$ are bounded on U; that is, there exists a real number $M > 0$ such that

$$\Big|\frac{\partial f}{\partial x_j}(x)\Big| \leqslant M \qquad \forall\, x \in U \text{ and } j = 1, \cdots, n\,.$$

Show that f is continuous on U.

Hint: Mimic the proof of Theorem 5.40 in the lecture note.

Proof. Assume that $\Big|\dfrac{\partial f}{\partial x_i}(x)\Big| \leqslant M$ for all $x \in U$ and $1 \leqslant i \leqslant n$. Let $a \in U$ be given. Then there exists $r > 0$ such that $B(a,r) \subseteq U$. For $x \in B(a,r)$, let $k = x - a$. Then

$$\begin{aligned}\big|f(x) - f(a)\big| &= \big|f(a_1+k_1, a_2+k_2, \cdots, a_n+k_n) - f(a_1, a_2, \cdots, a_n)\big|\\ &= \Big|\sum_{j=1}^n \big[f(a_1, \cdots, a_{j-1}, a_j+k_j, \cdots, a_n+k_n) - f(a_1, \cdots, a_j, a_{j+1}+k_{j+1}, \cdots, a_n+k_n)\big]\Big|\\ &\leqslant \sum_{j=1}^n \Big|f(a_1, \cdots, a_{j-1}, a_j+k_j, \cdots, a_n+k_n) - f(a_1, \cdots, a_j, a_{j+1}+k_{j+1}, \cdots, a_n+k_n)\Big|\,.\end{aligned}$$

By the Mean Value Theorem, for each $1 \leqslant j \leqslant n$ there exists $\theta_j \in (0,1)$ such that

$$\begin{aligned}&\big|f(a_1,\cdots,a_{j-1},a_j+k_j,\cdots,a_n+k_n) - f(a_1,\cdots,a_j,a_{j+1}+k_{j+1},\cdots,a_n+k_n)\\ &\quad = \frac{\partial f}{\partial x_j}(a_1,\cdots,a_{j-1},a_j+\theta_j k_j,a_{j+1}+k_{j+1},\cdots,a_n+k_n)k_j\,;\end{aligned}$$

thus

$$\big|f(a_1,\cdots,a_{j-1},a_j+k_j,\cdots,a_n+k_n) - f(a_1,\cdots,a_j,a_{j+1}+k_{j+1},\cdots,a_n+k_n)\big| \leqslant M|k_j|\,.$$

Therefore, if $x \in B(a,r)$,

$$\big|f(x)-f(a)\big| = \sum_{j=1}^n M|k_j| \leqslant M\sqrt{n}\Big(\sum_{j=1}^n |k_j|^2\Big)^{\frac{1}{2}} = \sqrt{n}M\|x-a\|_{\mathbb{R}^n}\,.$$

This shows that f is continuous at a. □

Problem 12. Let $U \subseteq \mathbb{R}^n$ be open, and $f: U \to \mathbb{R}$. Show that f is differentiable at $a \in U$ if and only if there exists a vector-valued function $\varepsilon: U \to \mathbb{R}^n$ such that

$$f(x) - f(a) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a)(x_j - a_j) = \varepsilon(x)\cdot(x-a)$$

and $\varepsilon(x) \to 0$ as $x \to a$.

Proof. "⇒" Suppose that f is differentiable at a. Define $\varepsilon: U \to \mathbb{R}^n$ by

$$\varepsilon(x) = \begin{cases} \Big[f(x)-f(a)-\sum\limits_{j=1}^n \dfrac{\partial f}{\partial x_j}(a)(x_j-a_j)\Big]\dfrac{x-a}{\|x-a\|^2} & \text{if } x \neq a\,,\\ 0 & \text{if } x = a\,.\end{cases}$$

Then for $x \neq a$,

$$\big|\varepsilon(x)\big| \leqslant \frac{\Big|f(x)-f(a)-\sum\limits_{j=1}^n \dfrac{\partial f}{\partial x_j}(a)(x_j-a_j)\Big|}{\|x-a\|}$$

which, by the differentiability of f at a, implies that

$$\lim_{x\to a}\big|\varepsilon(x)\big| = 0\,.$$

Moreover,

$$\varepsilon(x)\cdot(x-a) = f(x)-f(a)-\sum_{j=1}^n \frac{\partial f}{\partial x_j}(a)(x_j-a_j)\,.$$

"⇐" Suppose that there exists a vector-valued function $\varepsilon: U \to \mathbb{R}^n$ such that

$$f(x) - f(a) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a)(x_j - a_j) = \varepsilon(x)\cdot(x-a)$$

and $\varepsilon(x) \to 0$ as $x \to a$. Then for $x \neq a$, the Cauchy-Schwarz inequality implies that

$$\frac{\Big|f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j)\Big|}{\|x-a\|} = \frac{|\varepsilon(x)\cdot(x-a)|}{\|x-a\|} \leqslant \|\varepsilon(x)\|\,;$$

thus

$$\lim_{x\to a} \frac{\Big|f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j)\Big|}{\|x-a\|} = 0\,.$$

Therefore, f is differentiable at a with $[(Df)(a)] = \Big[\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_n}(a)\Big]$. □

Problem 13. Let

$$f(x,y) = \begin{cases} \dfrac{x^3 y}{x^4 + y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

and $u \in \mathbb{R}^2$ be a unit vector. Show that the directional derivative of f at the origin exists in all direction, and

$$(D_u f)(0,0) = \Big(\frac{\partial f}{\partial x}(0,0), \frac{\partial f}{\partial y}(0,0)\Big)\cdot u\,.$$

Is f differentiable at $(0,0)$?

Solution. Let $u = (\cos\theta, \sin\theta)$ be a unit vector. Then the directional derivative of f at $(0,0)$ in direction u is

$$\begin{aligned}(D_u f)(0,0) &= \lim_{t\to 0^+} \frac{f(t\cos\theta, t\sin\theta) - f(0,0)}{t} = \lim_{t\to 0^+} \frac{t^4 \cos^3\theta \sin\theta}{t(t^4\cos^4\theta + t^2\sin^2\theta)} \\ &= \lim_{t\to 0^+} \frac{t\cos^3\theta\sin\theta}{t^2\cos^4\theta + \sin^2\theta} = 0\,.\end{aligned}$$

On the other hand,

$$f_x(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = 0 \quad \text{and} \quad f_y(0,0) = \lim_{k\to 0} \frac{f(0,k) - f(0,0)}{k} = 0\,;$$

thus we conclude that $(D_u f)(0,0) = \big(f_x(0,0), f_y(0,0)\big)\cdot u$.

Since $f_x(0,0) = f_y(0,0) = 0$, if f is differentiable at $(0,0)$, we must have

$$\lim_{(x,y)\to(0,0)} \frac{\big|f(x,y) - f(0,0) - 0\cdot(x-0) - 0\cdot(y-0)\big|}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{|x^3 y|}{\sqrt{x^2+y^2}(x^4+y^2)} = 0\,;$$

however, by passing to the limit as $(x,y) \to (0,0)$ along the curve $y = x^2$, we find that

$$0 = \lim_{x\to 0} \frac{|x^3\cdot x^2|}{\sqrt{x^2+x^4}(x^4+x^4)} = \lim_{x\to 0} \frac{1}{2\sqrt{1+x^2}} = \frac{1}{2}\,,$$

a contradiction. Therefore, f is not differentiable at $(0,0)$. □