Exercise Problem Sets 15

Dec. 23. 2022

Problem 1. Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}^m$ with $f = (f_1, \cdots, f_m)$.

1. Suppose that f is differentiable on U and the line segment joining x and y lies in U. Then there exist points $c_1, \cdots, c_m$ on that segment such that
$$f_i(y) - f_i(x) = (Df_i)(c_i)(y - x) \qquad \forall\, i = 1, \cdots, m.$$

2. Suppose in addition that U is convex (the convexity of sets is defined in Problem 7 in Exercise 11). Show that for each $x, y \in U$ and vector $v \in \mathbb{R}^m$, there exists c on the line segment joining x and y such that
$$v \cdot \big[f(x) - f(y)\big] = v \cdot (Df)(c)(x - y)\,.$$
In particular, show that if $\sup\limits_{x \in U} \|(Df)(x)\|_{\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)} \leqslant M$, then
$$\|f(x) - f(y)\|_{\mathbb{R}^m} \leqslant M\|x - y\|_{\mathbb{R}^n} \qquad \forall\, x, y \in U\,.$$

Proof. Let $\gamma : [0,1] \to \mathbb{R}^n$ be given by $\gamma(t) = (1-t)x + ty$. Then by the chain rule, for each $i = 1, \cdots, m$, $(f_i \circ \gamma) : [0,1] \to \mathbb{R}$ is differentiable on $(0,1)$; thus the mean value theorem (for functions of one real variable) implies that there exists $t_i \in (0,1)$ such that

$$f_i(y) - f_i(x) = (f_i \circ \gamma)(1) - (f_i \circ \gamma)(0) = (f_i \circ \gamma)'(t_i) = (Df_i)(c_i)\big(\gamma'(t_i)\big)\,,$$

where $c_i = \gamma(t_i)$. Part 1 is concluded since $\gamma'(t_i) = y - x$.

For $v \in \mathbb{R}^n$, let $g(t) = v \cdot f(ty + (1-t)x)$. Then $g : [0,1] \to \mathbb{R}$ is differentiable; thus the mean value theorem (for functions of one real variable) implies that there exists $0 < t_0 < 1$ such that

$$v \cdot \big[f(y) - f(x)\big] = g(1) - g(0) = g'(t_0) = v \cdot (Df)(t_0 y + (1-t_0)x)(x - y)\,.$$

Letting $c = t_0 y + (1-t_0)x$, we conclude that $v \cdot \big[f(x) - f(y)\big] = v \cdot (Df)(c)(x - y)$.

Finally, let $v = f(y) - f(x)$. By the discussion above there exists $c \in \overline{xy}$ such that

$$\|f(y) - f(x)\|_{\mathbb{R}^m}^2 = v \cdot \big[f(y) - f(x)\big] = v \cdot (Df)(c)(x - y)\,.$$

The Cauchy-Schwarz inequality further implies that

$$\begin{aligned}\|f(y) - f(x)\|_{\mathbb{R}^m}^2 &\leqslant \|f(y) - f(x)\|_{\mathbb{R}^m}\|(Df)(c)(x - y)\|_{\mathbb{R}^m} \\ &\leqslant \|f(y) - f(x)\|_{\mathbb{R}^m}\|(Df)(c)\|_{\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)}\|x - y\|_{\mathbb{R}^n}\,.\end{aligned}$$

Therefore, if $\sup\limits_{x \in U} \|(Df)(x)\|_{\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)} \leqslant M$, we conclude that

$$\|f(y) - f(x)\|_{\mathbb{R}^m} \leqslant M\|x - y\|_{\mathbb{R}^n} \qquad \forall\, x, y \in U\,.$$ □

Problem 2. Let $U \subseteq \mathbb{R}^n$ be open and connected, and $f : U \to \mathbb{R}$ be a function such that $\dfrac{\partial f}{\partial x_j}(x) = 0$ for all $x \in U$. Show that f is constant in U.

Proof. First, we show that if $B(a,r)$ is a ball in U, then f is constant on U. In fact, by the fact that balls are convex set, Problem 1 implies that

$$\big|f(y) - f(x)\big| \leqslant \sup_{z \in B(a,r)} \|(Df)(z)\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R})} \|x - y\|_{\mathbb{R}^n} \qquad \forall\, x, y \in B(a,r)\,.$$

Since $\dfrac{\partial f}{\partial x_j}(x) = 0$ for all $x \in B(a,r)$, we find that $\|(Df)(z)\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R})} = 0$ for all $z \in B(a,r)$; thus $f(y) = f(x)$ for all $x, y \in B(a,r)$.

Suppose that $f = c$ in $B(a,r)$. Let $E = f^{-1}(\{c\})$. Note that the fact $\dfrac{\partial f}{\partial x_j}(x) = 0$ for all $x \in U$ implies that Df is continuous on U; thus f is continuously differentiable on U. In particular, f is continuous; thus $f^{-1}(\{c\})$ is closed relative to U. Suppose that $f^{-1}(\{c\}) = F \cap U$ for some closed set F in $\mathbb{R}^n$. Next we show that $U \backslash F = \varnothing$ so that $f = c$ on U.

Suppose the contrary that $U \backslash F \neq \varnothing$. Let $E_1 = U \cap F^{\complement}$ and $E_2 = U \cap F$. Then $U = E_1 \cup E_2$ and Problem 6 in Exercise 7 shows that

$$E_1 \cap \overline{E_2} \subseteq E_1 \cap \bar{F} = U \cap F^{\complement} \cap F = \varnothing\,.$$

Therefore, $\overline{E_1} \cap E_2 \neq \varnothing$ for otherwise U is disconnected by Proposition 3.65 in the lecture note. This implies that there exists $x \in \overline{E_1} \cap E_2$; thus there exists $\{x_k\}_{k=1}^{\infty} \subseteq U \backslash F$ such that $x_k \to x$ as $k \to \infty$. Since $x \in U$, there exists $\epsilon > 0$ such that $B(x,\epsilon) \subseteq U$; thus the convergence of $\{x_k\}_{k=1}^{\infty}$ implies that there exists $N > 0$ such that $x_k \in B(x,\epsilon)$ for all $k \geqslant N$. By the discussion above, f is constant on $B(x,\epsilon)$; thus $f(x_k) = f(x) = c$ for all $k \geqslant N$, a contradiction to that $x_k \notin F$. □

Problem 3. Let $U \subseteq \mathbb{R}^n$ be open, and for each $1 \leqslant i, j \leqslant n$, $a_{ij} : U \to \mathbb{R}$ be differentiable functions. Define $A = [a_{ij}]$ and $J = \det(A)$. Show that

$$\frac{\partial J}{\partial x_k} = \operatorname{tr}\Big(\operatorname{Adj}(A)\frac{\partial A}{\partial x_k}\Big) \qquad \forall\, 1 \leqslant k \leqslant n\,,$$

where for a square matrix $M = [m_{ij}]$, $\operatorname{tr}(M)$ denotes the trace of M, $\operatorname{Adj}(M)$ denotes the adjoint matrix of M, and $\dfrac{\partial M}{\partial x_k}$ denotes the matrix whose (i,j)-th entry is given by $\dfrac{\partial m_{ij}}{\partial x_k}$.

Hint: Show that

$$\frac{\partial J}{\partial x_k} = \begin{vmatrix} \frac{\partial a_{11}}{\partial x_k} & a_{12} & \cdots & a_{1n} \\ \frac{\partial a_{21}}{\partial x_k} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ \frac{\partial a_{n1}}{\partial x_k} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \frac{\partial a_{12}}{\partial x_k} & a_{13} & \cdots & a_{1n} \\ a_{21} & \frac{\partial a_{22}}{\partial x_k} & a_{23} & \cdots & a_{2n} \\ \vdots & & & & \vdots \\ a_{n1} & \frac{\partial a_{n2}}{\partial x_k} & a_{n3} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & \cdots & a_{(n-1)1} & \frac{\partial a_{1n}}{\partial x_k} \\ a_{21} & \cdots & a_{(n-1)2} & \frac{\partial a_{2n}}{\partial x_k} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & a_{(n-1)n} & \frac{\partial a_{n1}}{\partial x_k} \end{vmatrix}$$

and rewrite this identity in the form which is asked to prove. You can also show the differentiation formula by applying the chain rule to the composite function $F \circ g$ of maps $g : U \to \mathbb{R}^{n^2}$ and $F : \mathbb{R}^{n^2} \to \mathbb{R}$ defined by $g(x) = \big(a_{11}(x), a_{12}(x), \cdots, a_{nn}(x)\big)$ and $F(a_{11}, \cdots, a_{nn}) = \det\big([a_{ij}]\big)$. Check first what $\dfrac{\partial F}{\partial a_{ij}}$ is.

Proof. Let $A = [a_{ij}]$ and $\text{Adj}(A) = [c_{ij}]$. Then $\dfrac{\partial F}{\partial a_{ij}} = c_{ji}$ since the cofactor expansion implies that

$$\det(A) = a_{i1}c_{1i} + a_{i2}c_{2i} + \cdots + a_{in}c_{ni} \qquad \text{for each } 1 \leqslant i \leqslant n.$$

Therefore, for $J = \det(A)$, we have

$$\frac{\partial J}{\partial x_k}(x) = \frac{\partial (F \circ g)}{\partial x_k}(x) = \sum_{i,j=1}^{n} \frac{\partial F}{\partial a_{ij}}(g(x))\frac{\partial a_{ij}}{\partial x_k}(x) = \sum_{i,j=1}^{n} c_{ji}(x)\frac{\partial a_{ij}}{\partial x_k}(x)$$

and the result is concluded from the fact that $\text{tr}\Big(\text{Adj}(A)\dfrac{\partial A}{\partial x_k}\Big) = \sum\limits_{i,j=1}^{n} c_{ji}\dfrac{\partial a_{ij}}{\partial x_k}$. □

Problem 4. 1. If $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $g : B \subseteq \mathbb{R}^m \to \mathbb{R}^\ell$ are twice differentiable and $f(A) \subseteq B$, then for $x_0 \in A$, $u, v \in \mathbb{R}^n$, show that

$$\begin{aligned} &D^2(g \circ f)(x_0)(u, v) \\ &\quad = (D^2 g)(f(x_0))\big((Df)(x_0)(u), Df(x_0)(v)\big) + (Dg)(f(x_0))\big((D^2 f)(x_0)(u, v)\big)\,. \end{aligned}$$

2. If $p : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map plus some constant; that is, $p(x) = Lx + c$ for some $L \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$, and $f : A \subseteq \mathbb{R}^m \to \mathbb{R}^s$ is k-times differentiable, prove that

$$D^k(f \circ p)(x_0)(u^{(1)}, \cdots, u^{(k)}) = (D^k f)\big(p(x_0)\big)\big((Dp)(x_0)(u^{(1)}), \cdots, (Dp)(x_0)(u^{(k)})\big)\,.$$

Problem 5. Let $f(x, y)$ be a real-valued function on $\mathbb{R}^2$. Suppose that f is of class $\mathscr{C}^1$ (that is, all first partial derivatives are continuous on $\mathbb{R}^2$) and $\dfrac{\partial^2 f}{\partial x \partial y}$ exists and is continuous. Show that $\dfrac{\partial^2 f}{\partial y \partial x}$ exists and $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x}$.

Hint: Mimic the proof of Clairaut's Theorem.

Proof. Let $(a, b) \in \mathbb{R}^2$. For real numbers $h, k \neq 0$, define $Q : \mathbb{R}^2 \to \mathbb{R}$ and $\varphi : \mathbb{R}^2 \to \mathbb{R}$ by

$$Q(h, k) = \frac{1}{hk}\big[f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b)\big]$$

and

$$\psi(x, y) = f(x + h, y) - f(x, y)\,.$$

Then $Q(h, k) = \dfrac{1}{hk}\big[\psi(a, b + k) - \psi(a, b)\big]$. By the mean value theorem (for functions of one real variable),

$$\begin{aligned} Q(h, k) &= \frac{1}{hk}\frac{\partial \psi}{\partial y}(a, b + \theta_1 k)k = \frac{1}{h}\Big[\frac{\partial f}{\partial y}(a + h, b + \theta_1 k) - \frac{\partial f}{\partial y}(a, b + \theta_1 k)\Big] \\ &= \frac{1}{h}\frac{\partial^2 f}{\partial x \partial y}(a + \theta_2 h, b + \theta_1 k)h = \frac{\partial^2 f}{\partial x \partial y}(a + \theta_2 h, b + \theta_1 k) \end{aligned}$$

for some function $\theta_1 = \theta(h, k)$ and $\theta_2 = \theta_2(h, k)$ satisfying $\theta_1, \theta_2 \in (0, 1)$. Since $\dfrac{\partial^2 f}{\partial x \partial y}$ is continuous, we find that

$$\lim_{(h,k)\to(0,0)} Q(h, k) = \lim_{(h,k)\to(0,0)} \frac{\partial^2 f}{\partial x \partial y}(a + \theta_2 h, b + \theta_1 k) = \frac{\partial^2 f}{\partial x \partial y}(a, b)\,.$$

On the other hand, since the limit $\lim_{(h,k)\to(0,0)} Q(h,k)$ exists,

$$\begin{aligned}\frac{\partial^2 f}{\partial x \partial y}(a,b) &= \lim_{(h,k)\to(0,0)} Q(h,k) = \lim_{k\to 0}\lim_{h\to 0} Q(h,k)\\ &= \lim_{k\to 0}\frac{1}{k}\Big[\lim_{h\to 0}\Big(\frac{f(a+h,b+k)-f(a,b+k)}{h} - \frac{f(a+h,b)-f(a,b)}{h}\Big)\Big]\\ &= \lim_{k\to 0}\frac{1}{k}\Big[\frac{\partial f}{\partial x}(b+k) - \frac{\partial f}{\partial x}(b)\Big];\end{aligned}$$

thus the limit $\lim_{k\to 0}\frac{f_x(a,b+k)-f_x(a,b)}{k}$ exists and equals $\frac{\partial^2 f}{\partial x \partial y}(a,b)$. By the definition of partial derivatives, $\frac{\partial^2 f}{\partial y \partial x}(a,b)$ exists and $\frac{\partial^2 f}{\partial y \partial x}(a,b) = \frac{\partial^2 f}{\partial x \partial y}(a,b)$. □

Problem 6. Let $U \subseteq \mathbb{R}^n$ be open, and $\psi : U \to \mathbb{R}^n$ be a function of class $\mathscr{C}^2$. Suppose that $(D\psi)(x) \in \mathrm{GL}(n)$ for all $x \in \mathbb{R}^n$, and define $J = \det([D\psi])$ and $A = [D\psi]^{-1}$, where $[D\psi]$ is the Jacobian matrix of ψ. Write $[A] = [a_{ij}]$.

1. Show that for each $1 \leqslant i,j,k \leqslant n$, $a_{ij} : \mathbb{R}^n \to \mathbb{R}$ is differentiable, and
$$\frac{\partial a_{ij}}{\partial x_k} = -\sum_{r,s=1}^n a_{ir}\frac{\partial^2 \psi_r}{\partial x_k \partial x_s} a_{sj}\,.$$

2. Show the Piola identity
$$\sum_{i=1}^n \frac{\partial}{\partial x_i}(Ja_{ij})(x) = 0 \qquad \forall\, 1 \leqslant j \leqslant n \text{ and } x \in U\,. \tag{0.1}$$

Proof. Note that since $A = [D\psi]^{-1}$, we have
$$\sum_{r=1}^n a_{ir}\frac{\partial \psi_r}{\partial x_s} = \sum_{r=1}^n \frac{\partial \psi_i}{\partial x_r} a_{rs} = \delta_{is}\,,$$
where δ_{is} is the Kronecker delta.

1. The product rule implies that
$$\sum_{r=1}^n \Big(\frac{\partial a_{ir}}{\partial x_k}\frac{\partial \psi_r}{\partial x_s} + a_{ir}\frac{\partial^2 \psi_r}{\partial x_k \partial x_s}\Big) = 0\,;$$
thus
$$\sum_{r=1}^n \frac{\partial a_{ir}}{\partial x_k}\frac{\partial \psi_r}{\partial x_s} = -\sum_{r=1}^n a_{ir}\frac{\partial^2 \psi_r}{\partial x_k \partial x_s}\,.$$
Therefore,
$$\sum_{s=1}^n a_{sj}\sum_{r=1}^n \frac{\partial a_{ir}}{\partial x_k}\frac{\partial \psi_r}{\partial x_s} = -\sum_{s=1}^n\sum_{r=1}^n a_{ir}\frac{\partial^2 \psi_r}{\partial x_k \partial x_s}a_{sj} = -\sum_{r,s=1}^n a_{ir}\frac{\partial^2 \psi_r}{\partial x_k \partial x_s}a_{sj}\,,$$
and Part 1 follows from the fact that $\sum_{s=1}^n \frac{\partial \psi_r}{\partial x_s}a_{sj} = \delta_{rj}$ and $\sum_{r=1}^n \delta_{rj}\frac{\partial a_{ir}}{\partial x_k} = \frac{\partial a_{ij}}{\partial x_k}$.

2. Note that since $(D\psi) \in \mathrm{GL}(n)$, by the property of the adjoint matrix we obtain that

$$JA = \det([D\psi])[D\psi]^{-1} = \mathrm{Adj}([D\psi])$$

which implies that the (i,j)-entry of $\mathrm{Adj}([D\psi])$ is Ja_{ij}. Therefore, using the result in Problem 3 shows that

$$\frac{\partial J}{\partial x_i} = \mathrm{tr}\Big(\mathrm{Adj}([D\psi])\frac{\partial [D\psi]}{\partial x_i}\Big) = \sum_{r,s=1}^{n} Ja_{rs}\frac{\partial}{\partial x_i}\frac{\partial \psi_s}{\partial x_r} = \sum_{r,s=1}^{n} Ja_{rs}\frac{\partial^2 \psi_s}{\partial x_i \partial x_r}\,;$$

thus the product rule implies that

$$\begin{aligned}\sum_{i=1}^{n}\frac{\partial}{\partial x_i}(Ja_{ij}) &= \sum_{i=1}^{n}\frac{\partial J}{\partial x_i}a_{ij} + \sum_{i=1}^{n} J\frac{\partial a_{ij}}{\partial x_i} = \sum_{i,r,s=1}^{n} Ja_{rs}\frac{\partial^2 \psi_s}{\partial x_i \partial x_r}a_{ij} - \sum_{i,r,s=1}^{n} Ja_{ir}\frac{\partial^2 \psi_r}{\partial x_i \partial x_s}a_{sj}\\ &= \sum_{i,r,s=1}^{n} Ja_{rs}\frac{\partial^2 \psi_s}{\partial x_i \partial x_r}a_{ij} - \sum_{i,r,s=1}^{n} Ja_{rs}\frac{\partial^2 \psi_s}{\partial x_r \partial x_i}a_{ij}\\ &= \sum_{i,r,s=1}^{n} Ja_{rs}\Big(\frac{\partial^2 \psi_s}{\partial x_i \partial x_r} - \frac{\partial^2 \psi_s}{\partial x_r \partial x_i}\Big)a_{ij}\end{aligned}$$

and the conclusion follows from Clairaut's Theorem. □