

1 Norms and Differentiation

1/21: 1. Let V be a vector space over \mathbb{R} . Recall that a norm on V is a function

$$\|\cdot\| : V \rightarrow \mathbb{R}$$

such that

- $\|\lambda \mathbf{v}\| = |\lambda| \cdot \|\mathbf{v}\|$ for all $\lambda \in \mathbb{R}$, $\mathbf{v} \in V$;
- $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ for all $\mathbf{v}, \mathbf{w} \in V$;
- $\|\mathbf{v}\| = 0$ iff $\mathbf{v} = \mathbf{0}$.

A norm defines a metric on V given by $d(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|$. We will say that two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if there exist constants $C_1, C_2 \in \mathbb{R}$ with $0 < C_1 \leq C_2$ such that for all $\mathbf{v} \in V$,

$$C_1 \|\mathbf{v}\|_2 \leq \|\mathbf{v}\|_1 \leq C_2 \|\mathbf{v}\|_2$$

In this problem, you will show that any two norms on a finite dimensional vector space are equivalent.

- (a) Show that if $\|\cdot\|_1$ and $\|\cdot\|_2$ are two equivalent norms on V , then a subset $U \subset V$ is open with respect to $\|\cdot\|_1$ iff it is open with respect to $\|\cdot\|_2$. (Recall that a subset $U \subset V$ is open with respect to a norm $\|\cdot\|$ if for every $\mathbf{v} \in U$, there exists an $\epsilon > 0$ such that for every $\mathbf{w} \in V$ satisfying $\|\mathbf{v} - \mathbf{w}\| < \epsilon$, $\mathbf{w} \in U$.)

Proof. Suppose first that $U \subset V$ is open with respect to $\|\cdot\|_1$. To prove that U is open with respect to $\|\cdot\|_2$, it will suffice to show that for every $\mathbf{v} \in U$, there exists an $\epsilon > 0$ such that for every $\mathbf{w} \in V$ satisfying $\|\mathbf{v} - \mathbf{w}\|_2 < \epsilon$, $\mathbf{w} \in U$. Let $\mathbf{v} \in U$ be arbitrary. By hypothesis, there exists an $\epsilon' > 0$ such that for every $\mathbf{w} \in V$ satisfying $\|\mathbf{v} - \mathbf{w}\|_1 < \epsilon'$, $\mathbf{w} \in U$. Additionally, since $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, we know that there exists a $C_2 \in \mathbb{R}$ with $0 < C_2$ such that for all $\mathbf{v} \in V$, $\|\mathbf{v}\|_1 \leq C_2 \|\mathbf{v}\|_2$. Thus, if we choose $\epsilon = \epsilon'/C_2$ and let \mathbf{w} be an arbitrary element of V that satisfies $\|\mathbf{v} - \mathbf{w}\|_2 < \epsilon$, we have that

$$\|\mathbf{v} - \mathbf{w}\|_1 \leq C_2 \|\mathbf{v} - \mathbf{w}\|_2 < C_2 \epsilon = \epsilon'$$

This implies that $\mathbf{w} \in U$, as desired.

The proof is symmetric in the reverse direction. □

- (b) Let V be a finite dimensional vector space over \mathbb{R} with basis $\mathbf{e}_1, \dots, \mathbf{e}_n$. Let

$$\|\cdot\|_1 : V \rightarrow \mathbb{R}$$

denote the function given by

$$\|a_1 \mathbf{e}_1 + \dots + a_n \mathbf{e}_n\|_1 = \sum_{i=1}^n |a_i|$$

Show that $\|\cdot\|_1$ defines a norm on V .

Proof. To prove that $\|\cdot\|_1$ defines a norm on V , it will suffice to verify that $\|\cdot\|_1$ satisfies the three norm criteria. Let's begin.

For the first one, let $\lambda \in \mathbb{R}$ and $(a_1 \mathbf{e}_1 + \dots + a_n \mathbf{e}_n) \in V$ be arbitrary. Then

$$\begin{aligned} \|\lambda(a_1 \mathbf{e}_1 + \dots + a_n \mathbf{e}_n)\|_1 &= \|\lambda a_1 \mathbf{e}_1 + \dots + \lambda a_n \mathbf{e}_n\|_1 \\ &= \sum_{i=1}^n |\lambda a_i| \\ &= |\lambda| \cdot \sum_{i=1}^n |a_i| \\ &= |\lambda| \cdot \|a_1 \mathbf{e}_1 + \dots + a_n \mathbf{e}_n\|_1 \end{aligned}$$

as desired.

For the second one, let $(a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n), (b_1\mathbf{e}_1 + \cdots + b_n\mathbf{e}_n) \in V$ be arbitrary. Then

$$\begin{aligned} \|(a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n) + (b_1\mathbf{e}_1 + \cdots + b_n\mathbf{e}_n)\|_1 &= \|(a_1 + b_1)\mathbf{e}_1 + \cdots + (a_n + b_n)\mathbf{e}_n\|_1 \\ &= \sum_{i=1}^n |a_i + b_i| \\ &= \sum_{i=1}^n |a_i| + \sum_{i=1}^n |b_i| \\ &= \|a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n\|_1 + \|b_1\mathbf{e}_1 + \cdots + b_n\mathbf{e}_n\|_1 \end{aligned}$$

as desired.

For the third one, suppose first that $\|a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n\|_1 = 0$. Then $|a_1| + \cdots + |a_n| = 0$, so we must have $a_i = 0$ for all i by the properties of the absolute value. This implies that

$$a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n = 0\mathbf{e}_1 + \cdots + 0\mathbf{e}_n = \mathbf{0}$$

as desired. On the other hand, suppose that $a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n = \mathbf{0}$. Then since $\mathbf{e}_1, \dots, \mathbf{e}_n$ is a basis (hence linearly independent), we must have that $a_1 = \cdots = a_n = 0$. This implies that

$$\|a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n\|_1 = \|0\mathbf{e}_1 + \cdots + 0\mathbf{e}_n\|_1 = \sum_{i=1}^n |0| = 0$$

as desired. □

(c) Let V and $\|\cdot\|_1$ be as in the previous part, and let

$$\|\cdot\| : V \rightarrow \mathbb{R}$$

be another norm on V . Show that the function $\|\cdot\| : V \rightarrow \mathbb{R}$ is continuous with respect to the metric defined by $\|\cdot\|_1$. Deduce that there exist constants $C_1, C_2 \in \mathbb{R}$ such that for all $\mathbf{v} \in V$ with $\|\mathbf{v}\|_1 = 1$,

$$C_1 \leq \|\mathbf{v}\| \leq C_2$$

(Hint: Use the fact that the unit sphere with respect to $\|\cdot\|_1$ is compact.)

Proof. First off, note that for every $(a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n) \in V$,

$$\begin{aligned} \|a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n\| &\leq \sum_{i=1}^n |a_i| \cdot \|\mathbf{e}_i\| \\ &\leq \max \|\mathbf{e}_i\| \cdot \sum_{i=1}^n |a_i| \\ &= \max \|\mathbf{e}_i\| \cdot \|a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n\|_1 \end{aligned}$$

To prove that $\|\cdot\|$ is continuous with respect to $\|\cdot\|_1$, it will suffice to show that it is continuous at every $\mathbf{x} \in V$. Let $\mathbf{x} \in V$ be arbitrary. We want to demonstrate that for every $\epsilon > 0$, there exists a $\delta > 0$ such that if $\mathbf{y} \in V$ and $\|\mathbf{y} - \mathbf{x}\|_1 < \delta$, then $|\|\mathbf{y}\| - \|\mathbf{x}\|| < \epsilon$. Let $\epsilon > 0$ be arbitrary. Choose $\delta = \epsilon / \max \|\mathbf{e}_i\|$. Let $\mathbf{y} \in V$ be such that $\|\mathbf{y} - \mathbf{x}\|_1 < \delta$. Then

$$\|\|\mathbf{y}\| - \|\mathbf{x}\|\| \leq \|\mathbf{y} - \mathbf{x}\| \leq \max \|\mathbf{e}_i\| \cdot \|\mathbf{y} - \mathbf{x}\|_1 < \max \|\mathbf{e}_i\| \cdot \delta = \epsilon$$

as desired.

A particular result of the above argument is that $\|\cdot\|_1$ is continuous. Consequently, we know that the unit sphere with respect to $\|\cdot\|_1$, i.e., the set $(\|\cdot\|_1)^{-1}([0, 1])$ is closed. Additionally, we know that it is bounded by the unit shell $\|\mathbf{x}\|_1 = 1$. Thus, since V is equivalent to \mathbb{R}^n via an

isomorphism, we can apply the Heine-Borel theorem to the image of the unit sphere of V in \mathbb{R}^n and then go back.

Having established that $\|\cdot\|$ is continuous with respect to the metric defined by $\|\cdot\|_1$, we now turn our attention to the second part of the problem. Since $\|\cdot\|$ is a continuous real function and the unit sphere with respect to $\|\cdot\|_1$ is compact, we have that the image of the unit sphere under $\|\cdot\|$ is closed and bounded. It follows since the unit shell $\{\mathbf{v} \in V : \|\mathbf{v}\|_1 = 1\}$ is a subset of the unit sphere that the image of the unit shell under $\|\cdot\|$ is bounded. Consequently, we may choose $C_1, C_2 \in \mathbb{R}$ such that for all $\mathbf{v} \in V$, $C_1 \leq \|\mathbf{v}\| \leq C_2$. \square

- (d) Prove that any two norms on a finite dimensional vector space are equivalent.

Proof. Let $\|\cdot\|_a$ and $\|\cdot\|_b$ be norms on a finite dimensional vector space V . By part (c), there exist constants $C_{1a}, C_{2a}, C_{1b}, C_{2b} \in \mathbb{R}$ such that for all $\mathbf{v} \in V$ with $\|\mathbf{v}\|_1 = 1$,

$$C_{1a} \leq \|\mathbf{v}\|_a \leq C_{2a} \qquad C_{1b} \leq \|\mathbf{v}\|_b \leq C_{2b}$$

Multiplying through the above inequalities by select positive constants gives

$$\begin{aligned} \frac{C_{1b}}{C_{2a}} C_{1a} &\leq \frac{C_{1b}}{C_{2a}} \|\mathbf{v}\|_a \leq \frac{C_{1b}}{C_{2a}} C_{2a} & 1C_{1b} &\leq 1\|\mathbf{v}\|_b \leq 1C_{2b} & \frac{C_{2b}}{C_{1a}} C_{1a} &\leq \frac{C_{2b}}{C_{1a}} \|\mathbf{v}\|_a \leq \frac{C_{2b}}{C_{1a}} C_{2a} \\ \frac{C_{1b}}{C_{2a}} C_{1a} &\leq \frac{C_{1b}}{C_{2a}} \|\mathbf{v}\|_a \leq C_{1b} & C_{1b} &\leq \|\mathbf{v}\|_b \leq C_{2b} & C_{2b} &\leq \frac{C_{2b}}{C_{1a}} \|\mathbf{v}\|_a \leq \frac{C_{2b}}{C_{1a}} C_{2a} \end{aligned}$$

It follows by consecutive applications of the transitive property that

$$\frac{C_{1b}}{C_{2a}} \|\mathbf{v}\|_a \leq \|\mathbf{v}\|_b \leq \frac{C_{2b}}{C_{1a}} \|\mathbf{v}\|_a$$

As suggested by the above equation, choose

$$C_1 = \frac{C_{1b}}{C_{2a}} \qquad C_2 = \max\left(\frac{C_{2b}}{C_{1a}}, \frac{C_{1b}}{C_{2a}}\right)$$

By our choice of C_1 and C_2 , we know that $C_1 \leq C_2$; we now confirm that they are both strictly positive. To do so, it will suffice to show that $C_{1a}, C_{2a}, C_{1b}, C_{2b}$ are strictly positive.

First, note that since every norm is continuous (part (c)) and the unit shell is a closed subset of a compact set (hence compact itself), the image of the unit shell under any norm is compact. Moreover, since the norm is a nonnegative function and $\mathbf{0}$ is not an element of the unit shell, the elements of the image are strictly positive. This combined with the fact that 0 cannot be a limit point of the image (which is closed by the Heine-Borel theorem) proves that there exists a positive lower bound on the image. Therefore, we may take $C_{1a}, C_{2a}, C_{1b}, C_{2b}$ to be strictly positive. \square

2. Let $U \subset \mathbb{R}^n$ be an open subset, and suppose that a function

$$f : U \rightarrow \mathbb{R}^m$$

is differentiable at a point $\mathbf{x}_0 \in U$. For a real number $\lambda > 0$, let g_λ denote the function

$$g_\lambda(\mathbf{x}) = \frac{f(\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0)) - f(\mathbf{x}_0)}{\lambda}$$

Prove that g_λ converges to the linear function $g(\mathbf{x}) = Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$ as $\lambda \rightarrow 0$, where the limit is taken in the topology of uniform convergence on compact sets. In other words, for every compact subset $K \subset \mathbb{R}^n$, prove that the restriction $g_\lambda|_K$ converges uniformly to $g|_K$. This is a precise formulation of the idea that a differentiable function looks linear when “zoomed in” at a point.

Proof. Let K be an arbitrary compact subset of \mathbb{R}^n . To prove that $g_\lambda|_K$ converges uniformly to $g|_K$, it will suffice to show that for every $\epsilon > 0$, there exists a $\delta > 0$ such that if $\lambda < \delta$, then $|g_\lambda|_K(\mathbf{x}) - g|_K(\mathbf{x})| < \epsilon$ for all $\mathbf{x} \in K \cap U$. Let's begin.

Let $\epsilon > 0$ be arbitrary. First off, since K is compact (hence bounded), we know that there exists some $M \in \mathbb{R}$ such that $\|\mathbf{x} - \mathbf{x}_0\| \leq M$ for all $\mathbf{x} \in K$. Additionally, since f is differentiable at \mathbf{x}_0 , we know that

$$\lim_{\lambda(\mathbf{x}-\mathbf{x}_0) \rightarrow \mathbf{0}} \frac{f(\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0)) - f(\mathbf{x}_0) - Df(\mathbf{x}_0)(\lambda(\mathbf{x} - \mathbf{x}_0))}{\|\lambda(\mathbf{x} - \mathbf{x}_0)\|} = \mathbf{0}$$

for all $\mathbf{x} \in K$. In particular, going forward let \mathbf{x} denote the element of $K \cap U$ that maximizes f near \mathbf{x}_0 (we know that such an element exists by the compactness of K and the differentiability [hence continuity] of f at \mathbf{x}_0). Note that we will also take $\mathbf{x} \neq \mathbf{x}_0$ since the other case is trivial. With this particular \mathbf{x} in mind and the help of the above limit, we know that there exists a $\delta' > 0$ such that if $\lambda(\mathbf{x} - \mathbf{x}_0) < \delta'$, then

$$\left\| \frac{f(\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0)) - f(\mathbf{x}_0) - Df(\mathbf{x}_0)(\lambda(\mathbf{x} - \mathbf{x}_0))}{\|\lambda(\mathbf{x} - \mathbf{x}_0)\|} \right\| < \frac{\epsilon}{M}$$

at this \mathbf{x} and every other element of the domain within a sufficiently small neighborhood. Now choose $\delta = \delta'/M$. Suppose $\lambda < \delta$. Then

$$\begin{aligned} \lambda(\mathbf{x} - \mathbf{x}_0) &\leq \lambda \cdot M \\ &< \delta \cdot M \\ &= \delta' \end{aligned}$$

so we may utilize the above inequality. In particular, it follows that under this supposition,

$$\begin{aligned} \|g_\lambda|_K(\mathbf{x}) - g|_K(\mathbf{x})\| &= \left\| \frac{f(\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0)) - f(\mathbf{x}_0)}{\lambda} - Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \right\| \\ &= \left\| \frac{f(\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0)) - f(\mathbf{x}_0) - \lambda Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)}{\lambda} \cdot \frac{\|\mathbf{x} - \mathbf{x}_0\|}{\|\mathbf{x} - \mathbf{x}_0\|} \right\| \\ &= \left\| \frac{f(\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0)) - f(\mathbf{x}_0) - Df(\mathbf{x}_0)(\lambda(\mathbf{x} - \mathbf{x}_0))}{\|\lambda(\mathbf{x} - \mathbf{x}_0)\|} \right\| \cdot \|\mathbf{x} - \mathbf{x}_0\| \\ &< \frac{\epsilon}{M} \cdot \|\mathbf{x} - \mathbf{x}_0\| \\ &\leq \frac{\epsilon}{M} \cdot M \\ &= \epsilon \end{aligned}$$

as desired. □

3. (a) Let

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

Prove that $D_{\mathbf{v}}f(\mathbf{0})$ exists for all vectors $\mathbf{v} \in \mathbb{R}^2$ but that f is not continuous at $(0, 0)$ (and in particular, not differentiable there).

Proof. Let $\mathbf{v} = (h, k) \in \mathbb{R}^2$ be an arbitrary nonzero vector. We have that

$$\begin{aligned} \frac{f(\mathbf{0} + t\mathbf{u}) - f(\mathbf{0})}{t} &= \frac{1}{t} \frac{(th)^2(tk)}{(th)^4 + (tk)^2} \\ &= \frac{h^2k}{t^2h^4 + k^2} \end{aligned}$$

Thus, to prove that $D_{\mathbf{v}}f(\mathbf{0})$ exists, it will suffice to show that the above fraction goes to zero as $t \rightarrow 0$. We divide into two cases ($k = 0$ and $k \neq 0$). If $k = 0$, then by the sum, product, and

quotient rules of limits, we have that

$$\begin{aligned} 0 &= \frac{h^2 \cdot 0}{0 \cdot h^4 + 0^2} \\ &= \frac{h^2 k}{(\lim_{t \rightarrow 0} t^2) \cdot h^4 + k^2} \\ &= \lim_{t \rightarrow 0} \frac{h^2 k}{t^2 h^4 + k^2} \end{aligned}$$

as desired. If $k \neq 0$, then we have by the same rules of limits that

$$\begin{aligned} \frac{h^2}{k} &= \frac{h^2 k}{0 \cdot h^4 + k^2} \\ &= \frac{h^2 k}{(\lim_{t \rightarrow 0} t^2) \cdot h^4 + k^2} \\ &= \lim_{t \rightarrow 0} \frac{h^2 k}{t^2 h^4 + k^2} \end{aligned}$$

as desired.

To prove that f is not continuous at $(0, 0)$, it will suffice to show that there exists an $\epsilon > 0$ such that for all $\delta > 0$, there exists $(x, y) \in \mathbb{R}^2$ such that $\|(x, y) - (0, 0)\| < \delta$ and $|f(x, y) - f(0, 0)| \geq \epsilon$. Choose $\epsilon = 1/2$. Let $\delta > 0$ be arbitrary. Choose $(y, y^2) \in \mathbb{R}^2$ such that $\|(y, y^2) - (0, 0)\| = \|(y, y^2)\| < \delta$. This combined with the fact that

$$\begin{aligned} |f(y, y^2) - f(0, 0)| &= \left| \frac{(y)^2(y^2)}{(y)^4 + (y^2)^2} \right| \\ &= \left| \frac{y^4}{2y^4} \right| \\ &= \left| \frac{1}{2} \right| \\ &= \epsilon \end{aligned}$$

completes the proof. It follows by the contrapositive of the proof that differentiability implies continuity that f is not differentiable at $(0, 0)$. \square

(b) Let

$$f(x, y) = \begin{cases} (x^2 + y^2) \sin\left(\frac{1}{\sqrt{x^2 + y^2}}\right) & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

Prove that f is differentiable at zero but both partial derivatives are not continuous at zero.

Proof. Let A denote the zero transformation. To prove that f is differentiable at $\mathbf{0}$ with total derivative A , it will suffice to show that $[f(\mathbf{h}) - f(\mathbf{0}) - A\mathbf{h}]/\|\mathbf{h}\| \rightarrow 0$ as $\mathbf{h} \rightarrow \mathbf{0}$. First off, we have that

$$\begin{aligned} \frac{f(\mathbf{h}) - f(\mathbf{0}) - A\mathbf{h}}{\|\mathbf{h}\|} &= \frac{(h_1^2 + h_2^2) \sin\left(1/\sqrt{h_1^2 + h_2^2}\right) - 0 - 0}{\|\mathbf{h}\|} \\ &= \frac{\|\mathbf{h}\|^2}{\|\mathbf{h}\|} \sin\left(\frac{1}{\|\mathbf{h}\|}\right) \\ &= \|\mathbf{h}\| \sin\left(\frac{1}{\|\mathbf{h}\|}\right) \end{aligned}$$

Thus, since $\|\mathbf{h}\| \rightarrow 0$ as $\mathbf{h} \rightarrow \mathbf{0}$ and $\sin(1/\|\mathbf{h}\|)$ is bounded on $[-1, 1]$, we have the desired result.

As to the other part of the question, we have by the sum, product, and chain rules of derivatives (as well as the power and sine rules, which we'll take without proof) that the partial derivative WLOG with respect to x is

$$\begin{aligned}
 \frac{\partial f}{\partial x} &= \frac{\partial}{\partial x} \left((x^2 + y^2) \sin \left(\frac{1}{\sqrt{x^2 + y^2}} \right) \right) \\
 &= \frac{\partial}{\partial x} (x^2 + y^2) \cdot \sin \left(\frac{1}{\sqrt{x^2 + y^2}} \right) + (x^2 + y^2) \frac{\partial}{\partial x} \left(\sin \left(\frac{1}{\sqrt{x^2 + y^2}} \right) \right) \\
 &= \left(\frac{\partial}{\partial x} (x^2) + \frac{\partial}{\partial x} (y^2) \right) \cdot \sin \left(\frac{1}{\sqrt{x^2 + y^2}} \right) + (x^2 + y^2) \cdot \frac{\partial}{\partial u} (\sin(u)) \cdot \frac{\partial}{\partial x} ((x^2 + y^2)^{-1/2}) \\
 &= (2x + 0) \cdot \sin \left(\frac{1}{\sqrt{x^2 + y^2}} \right) + (x^2 + y^2) \cdot \cos \left(\frac{1}{\sqrt{x^2 + y^2}} \right) \cdot \frac{\partial}{\partial v} (v^{-1/2}) \cdot \frac{\partial}{\partial x} (x^2 + y^2) \\
 &= 2x \sin \left(\frac{1}{\sqrt{x^2 + y^2}} \right) + (x^2 + y^2) \cdot \cos \left(\frac{1}{\sqrt{x^2 + y^2}} \right) \cdot -\frac{1}{2(x^2 + y^2)^{3/2}} \cdot 2x \\
 &= 2x \sin \left(\frac{1}{\sqrt{x^2 + y^2}} \right) - \frac{x}{\sqrt{x^2 + y^2}} \cos \left(\frac{1}{\sqrt{x^2 + y^2}} \right)
 \end{aligned}$$

when $x \neq 0$ and is equal to 0 when $x = 0$ (from the Jacobian of A). However, $\partial f/\partial x$ is not continuous at zero since if we require $|\partial f/\partial x| < 1$, we can show that for any $\delta > 0$, there exists \mathbf{h} with $\|\mathbf{h}\| < \delta$ such that $(\partial f/\partial x)_{\mathbf{h}} \geq 1$. Indeed, use the Archimedean principle to choose $1/n\pi < \delta$ and let $\mathbf{h} = (1/n\pi, 0)$; it follows that

$$\begin{aligned}
 \left| \left(\frac{\partial f}{\partial x} \right)_{\mathbf{h}} \right| &= \left| 2 \cdot \frac{1}{n\pi} \sin \left(\frac{1}{\sqrt{(1/n\pi)^2 + 0^2}} \right) - \frac{1/n\pi}{\sqrt{(1/n\pi)^2 + 0^2}} \cos \left(\frac{1}{\sqrt{(1/n\pi)^2 + 0^2}} \right) \right| \\
 &= \left| \frac{2}{n\pi} \sin(n\pi) - \frac{1}{1} \cos(n\pi) \right| \\
 &= \left| \frac{2}{n\pi} \cdot 0 - 1 \cdot \pm 1 \right| \\
 &= 1
 \end{aligned}$$

as desired. □

4. Let $U \subset \mathbb{R}^n$ be an open subset, and $f : U \rightarrow \mathbb{R}$ be a function. Suppose that for $\mathbf{a} \in U$, the partial derivatives $\partial f/\partial x_i$ ($i = 1, \dots, n$) exist and are bounded in a neighborhood of \mathbf{a} . Prove that f is continuous at \mathbf{a} .

Proof. To prove that f is continuous at \mathbf{a} , it will suffice to show that $f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) \rightarrow 0$ as $\mathbf{h} \rightarrow \mathbf{0}$. We have

$$\begin{aligned}
 f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) &= f(a_1 + h_1, a_2 + h_2, a_3 + h_3, \dots, a_n + h_n) - f(a_1, a_2 + h_2, a_3 + h_3, \dots, a_n + h_n) \\
 &\quad + f(a_1, a_2 + h_2, a_3 + h_3, \dots, a_n + h_n) - f(a_1, a_2, a_3 + h_3, \dots, a_n + h_n) \\
 &\quad + \dots \\
 &\quad + f(a_1, \dots, a_{n-1}, a_n + h_n) - f(\mathbf{a})
 \end{aligned}$$

Applying the MVT to each term yields

$$f(a_1, \dots, a_i + h_i, \dots, a_n + h_n) - f(a_1, \dots, a_i, \dots, a_n + h_n) = h_i \frac{\partial f}{\partial x_i}(a_1, \dots, c_i(\mathbf{h}), \dots, a_n + h_n)$$

for some $c_i(\mathbf{h}) \in (a_i, a_i + h_i) \cup (a_i + h_i, a_i)$. It follows that

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a_1, \dots, c_i(\mathbf{h}), \dots, a_n + h_n)$$

But since the partial derivatives are bounded and the h_i 's are decreasing, each term in the summation goes to zero as $\mathbf{h} \rightarrow \mathbf{0}$, yielding the desired result. \square

5. Let

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

(a) Show that f is of class C^1 on \mathbb{R}^2 .

Proof. To prove that f is of class C^1 on \mathbb{R}^2 , it will suffice to show that the partial derivatives of f exist and are continuous on \mathbb{R}^2 (note that by Theorem 6.2 of Munkres (1991), this result would also imply differentiability of f across \mathbb{R}^2). As in Problem 3b, we can use the various rules of partial derivatives to determine that

$$\frac{\partial f}{\partial x} = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2} \qquad \frac{\partial f}{\partial y} = \frac{x(x^4 - 4x^2y^2 - y^4)}{(x^2 + y^2)^2}$$

for $(x, y) \neq (0, 0)$. We now confirm that both partial derivatives evaluate to zero at $(0, 0)$. To do so, we can show that the total derivative of f at zero is the zero transformation and pick the values from the Jacobian. To confirm that f is differentiable at $\mathbf{0}$ with total derivative equal to the zero transformation (again denoted by A), it will suffice to show that $[f(\mathbf{h}) - f(\mathbf{0}) - A\mathbf{h}]/\|\mathbf{h}\| \rightarrow 0$ as $\mathbf{h} \rightarrow \mathbf{0}$. First off, we switch to polar coordinates:

$$\begin{aligned} \frac{xy(x^2 - y^2)}{x^2 + y^2} &= \frac{(r \cos \theta)(r \sin \theta)((r \cos \theta)^2 - (r \sin \theta)^2)}{r^2} \\ &= \frac{r^4 \sin \theta \cos \theta (\cos^2 \theta - \sin^2 \theta)}{r^2} \\ &= r^2 \cdot \frac{\sin(2\theta)}{2} \cdot \cos(2\theta) \\ &= \frac{r^2}{2} \cdot \sin(2\theta) \cos(2\theta) \\ &= \frac{r^2}{2} \cdot \frac{\sin(2 \cdot 2\theta)}{2} \\ &= \frac{r^2}{4} \cdot \sin(4\theta) \end{aligned}$$

From here, we have that

$$\begin{aligned} \frac{f(\mathbf{h}) - f(\mathbf{0}) - A\mathbf{h}}{\|\mathbf{h}\|} &= \frac{\frac{r^2}{4} \cdot \sin(4\theta) - 0 - 0}{r} \\ &= \frac{r}{4} \cdot \sin(4\theta) \end{aligned}$$

Thus, since $r \rightarrow 0$ as $\mathbf{h} \rightarrow \mathbf{0}$ and $\sin(4\theta)$ is bounded on $[-1, 1]$, we have the desired result.

We now verify the continuity of the partial derivatives. We divide into two cases ($(x, y) \neq (0, 0)$ and $(x, y) = (0, 0)$). In the first case, since the expression for each partial derivative is composed of the sums, products, and quotients (with nonzero denominators) of polynomials (continuous functions), we know that both partial derivatives are continuous at every $(x, y) \neq (0, 0)$. At $(0, 0)$, however, is a different matter. We consider only $\partial f / \partial x$ in the following discussion since

the argument as it pertains to $\partial f/\partial y$ is symmetric. With respect to $\partial f/\partial x$, we invoke polar coordinates again to find that

$$\frac{\partial f}{\partial x} = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2} = \frac{r}{4}(3\sin(3\theta) - \sin(5\theta))$$

It follows that as $\mathbf{h} \rightarrow \mathbf{0}$, $r \rightarrow 0$, and so $\partial f/\partial x \rightarrow 0$, as desired. \square

(b) Show that both $\partial^2 f/\partial x \partial y$ and $\partial^2 f/\partial y \partial x$ exist at $(0, 0)$, but that

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) \neq \frac{\partial^2 f}{\partial y \partial x}(0, 0)$$

Proof. We have that

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y}(0, 0) &= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) \Big|_{(0,0)} & \frac{\partial^2 f}{\partial y \partial x}(0, 0) &= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) \Big|_{(0,0)} \\ &= \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial y} \Big|_{(h,0)} - \frac{\partial f}{\partial y} \Big|_{(0,0)}}{h} & &= \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial x} \Big|_{(0,h)} - \frac{\partial f}{\partial x} \Big|_{(0,0)}}{h} \\ &= \lim_{h \rightarrow 0} \frac{\frac{h(h^4 - 4h^2 \cdot 0^2 - 0^4)}{(h^2 + 0^2)^2} - 0}{h} & &= \lim_{h \rightarrow 0} \frac{\frac{h(0^4 + 4 \cdot 0^2 h^2 - h^4)}{(0^2 + h^2)^2} - 0}{h} \\ &= \lim_{h \rightarrow 0} 1 & &= \lim_{h \rightarrow 0} -1 \\ &= 1 & &= -1 \end{aligned}$$

so that

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1 \neq -1 = \frac{\partial^2 f}{\partial y \partial x}(0, 0)$$

as desired. \square

6. Let M_n denote the space of n -by- n matrices (which can be identified with \mathbb{R}^{n^2}), and let

$$\text{GL}_n \subset M_n$$

denote the subset of invertible matrices. In this problem, you will show that the operation of matrix inverse $\text{inv} : \text{GL}_n \rightarrow \text{GL}_n$ defined by

$$\text{inv}(A) = A^{-1}$$

is smooth and compute its derivative.

(a) Let $H \in M_n$ be a matrix such that $\|H\| < 1$ (where $\|\cdot\|$ denotes the operator norm). Show that $I + H$ is invertible (where $I \in M_n$ is the identity matrix). Use this to show that $\text{GL}_n \subset M_n$ is open. In particular, for a matrix $A \in \text{GL}_n$, if inv is differentiable at A , then the total derivative can be regarded as a linear function

$$D \text{inv}(A) : M_n \rightarrow M_n$$

Proof. Suppose for the sake of contradiction that $I + H$ is not invertible. Then $I + H$ has a nontrivial null space. Let $\mathbf{x} \in \ker(I + H)$ WLOG have $\|\mathbf{x}\| = 1$. It follows that

$$\begin{aligned} \mathbf{0} &= (I + H)\mathbf{x} \\ -\mathbf{x} &= H\mathbf{x} \\ \|\mathbf{x}\| &= \|H\mathbf{x}\| \\ 1 &= \|H\mathbf{x}\| \\ &\leq \|H\| \|\mathbf{x}\| \\ &= \|H\| \end{aligned}$$

a contradiction.

To prove that GL_n is open, it will suffice to show that for every $A \in \text{GL}_n$, there exists a neighborhood $N_r(A) \subset \text{GL}_n$. Let $A \in \text{GL}_n$ be arbitrary. Choose $r = 1/\|A^{-1}\|$. Let $B \in M_n$ be such that $\|B - A\| < r$. It follows that

$$\begin{aligned}\|A^{-1}B - I\| &= \|A^{-1}(B - A)\| \\ &\leq \|A^{-1}\| \cdot \|B - A\| \\ &< \|A^{-1}\| \cdot \frac{1}{\|A^{-1}\|} \\ &= 1\end{aligned}$$

Thus, by the first part of the proof, $I + (A^{-1}B - I) = A^{-1}B$ is invertible. This combined with the fact that A^{-1} is invertible proves that B is invertible, as desired. \square

(b) Show directly that inv is differentiable at the identity I with derivative

$$D\text{inv}(I)(X) = -X$$

for $X \in M_n$. (Hint: Show that for $\|H\| < 1$, $(I + H)^{-1} - I + H = H^2(I + H)^{-1}$.)

Proof. We first prove the hint. Let $H \in M_n$ have $\|H\| < 1$. Then $I + H$ is invertible, as per part (a). Thus, we may state that

$$\begin{aligned}(I + H)^{-1} - I + H &= [(I + H)^{-1} - I + H](I + H)(I + H)^{-1} \\ &= [I - (I + H) + (H + H^2)](I + H)^{-1} \\ &= H^2(I + H)^{-1}\end{aligned}$$

as desired.

To prove the desired result, it will suffice to show that $[\text{inv}(I + H) - \text{inv}(I) - (-H)]/\|H\| \rightarrow 0$ as $H \rightarrow 0$ (where 0 denotes the zero transformation). However, since $\|H\| \rightarrow 0$ implies $H \rightarrow 0$ and $\|[\text{inv}(I + H) - \text{inv}(I) - (-H)]/\|H\|\| \rightarrow 0$ implies $[\text{inv}(I + H) - \text{inv}(I) - (-H)]/\|H\| \rightarrow 0$, we need only show the normed conditions. Let's begin. We have that

$$\frac{\text{inv}(I + H) - \text{inv}(I) - (-H)}{\|H\|} = \frac{H^2(I + H)^{-1}}{\|H\|}$$

and that

$$\begin{aligned}\|(I + H)^{-1}\| &= \|I - H + H^2 - H^3 + \dots\| \\ &\leq \|I\| + \|H\| + \|H^2\| + \|H^3\| + \dots \\ &\leq \|I\| + \|H\| + \|H\|^2 + \|H\|^3 + \dots \\ &= \frac{1}{1 - \|H\|}\end{aligned}$$

It follows that

$$\begin{aligned}\left\|\frac{H^2(I + H)^{-1}}{\|H\|}\right\| &= \frac{1}{\|H\|} \cdot \|H^2(I + H)^{-1}\| \\ &\leq \frac{1}{\|H\|} \cdot \|H\| \cdot \|H\| \cdot \|(I + H)^{-1}\| \\ &= \|H\| \cdot \|(I + H)^{-1}\| \\ &\leq \|H\| \cdot \frac{1}{1 - \|H\|}\end{aligned}$$

Thus, since the bottom term in the above equation is composed of a term that converges to zero as $\|H\| \rightarrow 0$ and a term that is bounded by 1 as $\|H\| \rightarrow 0$ (for sufficiently small $\|H\|$), we now that it converges to zero as $\|H\| \rightarrow 0$. Therefore, $\|H^2(I + H)^{-1}/\|H\|\| \rightarrow 0$ as $\|H\| \rightarrow 0$, as desired. \square

- (c) Let $\text{mult} : M_n \times M_n \rightarrow M_n$ be the function defined by matrix multiplication, i.e.,

$$\text{mult}(A_1, A_2) = A_1 A_2$$

Show that mult is smooth and the derivative at $(A_1, A_2) \in M_n \times M_n$ is given by

$$D \text{mult}((A_1, A_2))(H_1, H_2) = A_1 H_2 + H_1 A_2$$

for $(H_1, H_2) \in M_n \times M_n$.

Proof. As in part (b), to prove the desired result, it will suffice to show that the norm of

$$\begin{aligned} & \frac{\text{mult}((A_1, A_2) + (H_1, H_2)) - \text{mult}(A_1, A_2) - D \text{mult}((A_1, A_2))(H_1, H_2)}{\|(H_1, H_2)\|} \\ &= \frac{\text{mult}(A_1 + H_1, A_2 + H_2) - \text{mult}(A_1, A_2) - (A_1 H_2 + H_1 A_2)}{\|(H_1, H_2)\|} \\ &= \frac{(A_1 A_2 + A_1 H_2 + H_1 A_2 + H_1 H_2) - A_1 A_2 - (A_1 H_2 + H_1 A_2)}{\|(H_1, H_2)\|} \\ &= \frac{H_1 H_2}{\|(H_1, H_2)\|} \end{aligned}$$

goes to zero as $\|(H_1, H_2)\| \rightarrow 0$ (note that we define a norm on $M_n \times M_n$ using the product topology, i.e., $\|(A, B)\| = \|A\| + \|B\|$). Indeed, since

$$\|H_1\| \leq \|H_1\| + \|H_2\| = \|(H_1, H_2)\| \quad \|H_2\| \leq \|H_1\| + \|H_2\| = \|(H_1, H_2)\|$$

we have that

$$\begin{aligned} \left\| \frac{H_1 H_2}{\|(H_1, H_2)\|} \right\| &\leq \frac{1}{\|(H_1, H_2)\|} \cdot \|H_1\| \cdot \|H_2\| \\ &\leq \frac{1}{\|(H_1, H_2)\|} \cdot \|(H_1, H_2)\| \cdot \|(H_1, H_2)\| \\ &= \|(H_1, H_2)\| \end{aligned}$$

Thus, since the desired norm is directly bounded by $\|(H_1, H_2)\|$, clearly it goes to zero as $\|(H_1, H_2)\| \rightarrow 0$, as desired.

As to smoothness, it follows from the fact that the derivative is as differentiable as the original function that we can continue differentiating forever. \square

- (d) Use the chain rule to show that inv is differentiable at every $A \in \text{GL}_n$ and that

$$D \text{inv}(A)(X) = -A^{-1} X A^{-1}$$

Deduce that inv is smooth.

Proof. The derivative is as differentiable as the original function (with the help of mult); thus, inv is smooth. \square