Exercises:

1. In this exercise, we will verify the **Chain Rule**. Consider the functions $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ and $g\colon \mathbb{R}^3 \to \mathbb{R}$ defined by
$$f(x_1, x_2) = (x_1x_2, x_1 + x_2, x_1^2) \qquad g(x_1, x_2, x_3) = x_1x_2x_3.$$
   (a) Compute $(Df)_x$ and $(Dg)_y$ for $x \in \mathbb{R}^2$ and $y \in \mathbb{R}^3$.
   (b) Determine the formula for $g \circ f\colon \mathbb{R}^2 \to \mathbb{R}$ and compute $(D(g \circ f))_x$ for $x \in \mathbb{R}^2$.
   (c) Compute $(Dg)_{f(x)} \circ (Df)_x$ for $x \in \mathbb{R}^2$, and compare this to part (b).

2. In this exercise, we will verify the **Product Rule** (for the bilinear form determined by the standard inner product on $\mathbb{R}^2$). Consider the functions $f, g\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by
$$f(x_1, x_2) = (x_2^2, x_1^2) \qquad g(x_1, x_2) = (\cos(x_1), \sin(x_2)).$$
   (a) Compute $(Df)_x$ and $(Dg)_x$ for $x \in \mathbb{R}^2$.
   (b) Define $h\colon \mathbb{R}^2 \to \mathbb{R}$ by
$$h(x_1, x_2) = \langle f(x_1, x_2), g(x_1, x_2)\rangle .$$
   Compute $(Dh)_x$ for $x \in \mathbb{R}^2$.
   (c) Show that for any $x, v \in \mathbb{R}^2$,
$$(Dh)_x(v) = \langle (Df)_x(v), g(x)\rangle + \langle f(x), (Dg)_x(v)\rangle .$$
   [**Hint:** by linearity, it suffices to check the above equality for $v = e_1, e_2$.]

3. Let $U \subset \mathbb{R}^n$ be open, and let $f\colon U \to \mathbb{R}^m$ be differentiable on U. Let $[p, q] \subset U$ be a segment. Assume the set
$$C = \{(Df)_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)\colon x \in [p, q]\}$$
is closed and **convex**: whenever $T, S \in C$ we have $tT + (1 - t)S \in C$ for $t \in [0, 1]$. Show that there exists $\theta \in [p, q]$ such that
$$f(q) - f(p) = (Df)_\theta(q - p);$$
that is, the direct generalization of one-dimensional the Mean Value Theorem holds. You may (and should) use the following special case of the **Hahn–Banach separation theorem** without proof: if A and B are non-empty, disjoint convex sets in $\mathbb{R}^m$ and A is open, then there exists a vector $v \in \mathbb{R}^m$ and a scalar $c \in \mathbb{R}$ such that
$$\langle v, a\rangle < c \leq \langle v, b\rangle$$
for all $a \in A$ and $b \in B$.
[**Hint:** proceed by contradiction and use the separation theorem in conjunction with the one-dimensional Mean Value Theorem.]

4. Let $U \subset \mathbb{R}^n$ be open and connected, and let $f\colon U \to \mathbb{R}^m$ be differentiable on U with $(Df)_p = 0$ for all $p \in U$. Show that f is constant.

5. Let (E, d) be an arbitrary metric space and let $[a, b] \subset \mathbb{R}$. Equip $[a, b] \times E$ with the product metric:
$$d_2((x, y), (x', y')) = \sqrt{|x - x'|^2 + d(y, y')^2} \qquad x, x' \in [a, b],\ y, y' \in E.$$
Let $f\colon [a, b] \times E \to \mathbb{R}$ be a continuous function. Show that
$$F(y) = \int_a^b f(x, y)\ dx$$
is continuous on E.
[**Hint:** recall that $[a, b]$ is compact.]

Solutions:

1. (a) We claim

$$(Df)_x = \begin{bmatrix} x_2 & x_1 \\ 1 & 1 \\ 2x_1 & 0 \end{bmatrix} \qquad \text{and} \qquad (Dg)_y = \begin{bmatrix} y_2y_3 & y_1y_3 & y_1y_2 \end{bmatrix}.$$

One can easily show that the corresponding Taylor remainders are sublinear, but here we simply note that the entries are clearly the partial derivatives, which are continuous. Consequently, a theorem of class implies $(Df)_x$ and $(Dg)_y$ exist and are given by the above formulas. □

(b) We compute:

$$g \circ f(x_1, x_2) = g(x_1x_2, x_1 + x_2, x_1^2) = x_1x_2(x_1 + x_2)x_1^2 = x_1^4x_2 + x_1^3x_2^2.$$

The partial derivatives are easily computed and seen to be continuous. Hence we have

$$(D(g \circ f))_x = \begin{bmatrix} 4x_1^3x_2 + 3x_1^2x_2^2 & x_1^4 + 2x_1^3x_2 \end{bmatrix}.$$

□

(c) Writing $f(x) = (f_1(x), f_2(x), f_3(x))$, we compute

$$\begin{aligned}
(Dg)_{f(x)} \circ (Df)_x &= \begin{bmatrix} f_2(x)f_3(x) & f_1(x)f_3(x) & f_1(x)f_2(x) \end{bmatrix} \cdot \begin{bmatrix} x_2 & x_1 \\ 1 & 1 \\ 2x_1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} x_1^3 + x_1^2x_2 & x_1^3x_2 & x_1^2x_2 + x_1x_2^2 \end{bmatrix} \begin{bmatrix} x_2 & x_1 \\ 1 & 1 \\ 2x_1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} x_1^3x_2 + x_1^2x_2^2 + x_1^3x_2 + 2x_1^3x_2 + 2x_1^2x_2^2 & x_1^4 + x_1^2x_2 + x_1^2x_2 \end{bmatrix} \\
&= \begin{bmatrix} 4x_1^3x_2 + 3x_1^2x_2^2 & x_1^4 + 2x_1^3x_2 \end{bmatrix},
\end{aligned}$$

which agrees with out computation in part (b). □

2. (a) By computing the partial derivatives (and noting their continuity), we obtain:

$$(Df)_x = \begin{bmatrix} 0 & 2x_2 \\ 2x_1 & 0 \end{bmatrix} \qquad \text{and} \qquad (Dg)_x = \begin{bmatrix} -\sin(x_1) & 0 \\ 0 & \cos(x_2) \end{bmatrix}$$

□

(b) Observe that

$$h(x_1, x_2) = x_2^2\cos(x_1) + x_1^2\sin(x_2).$$

So by computing partial derivatives (and noting their continuity), we obtain:

$$(Dh)_x = \begin{bmatrix} -x_2^2\sin(x_1) + 2x_1\sin(x_2) & 2x_2\cos(x_1) + x_1^2\cos(x_2) \end{bmatrix}$$

(c) We will make use of the hint. Observe that $(Dh)_x(e_j)$ is just the jth column of $(Dh)_x$, which we can read off from part (b). Now, we compute

$$\begin{aligned}
\langle (Df)_x(e_1), g(x)\rangle + \langle f(x), (Dg)_x(e_1)\rangle &= \left\langle \begin{pmatrix} 0 \\ 2x_1 \end{pmatrix}, \begin{pmatrix} \cos(x_1) \\ \sin(x_2) \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} x_2^2 \\ x_1^2 \end{pmatrix}, \begin{pmatrix} -\sin(x_1) \\ 0 \end{pmatrix} \right\rangle \\
&= 2x_1\sin(x_2) - x_2^2\sin(x_1),
\end{aligned}$$

which is the first entry of $(Dh)_x$. Similarly,

$$\begin{aligned}
\langle (Df)_x(e_2), g(x)\rangle + \langle f(x), (Dg)_x(e_2)\rangle &= \left\langle \begin{pmatrix} 2x_2 \\ 0 \end{pmatrix}, \begin{pmatrix} \cos(x_1) \\ \sin(x_2) \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} x_2^2 \\ x_1^2 \end{pmatrix}, \begin{pmatrix} 0 \\ \cos(x_2) \end{pmatrix} \right\rangle \\
&= 2x_2\cos(x_1) + x_1^2\cos(x_2),
\end{aligned}$$

which is the second entry of $(Dh)_x$. □

3. Suppose, towards a contradiction, that for every $\theta \in [p,q]$ we have $f(q) - f(p) \neq (Df)_\theta(q-p)$. Define

$$B := \{(Df)_x(p-q) \colon x \in [p,q]\} \subset \mathbb{R}^m.$$

Since C is closed and convex, it follows that B is closed and convex. By assumption, B is disjoint from $\{f(q) - f(p)\}$. Since B is closed, its complement is open and so there exists $r > 0$ such that the ball in $\mathbb{R}^m$ with center $f(q) - f(p)$ and radius r is disjoint from B. Call this ball A, and note that open balls are convex. The Hahn–Banach separation theorem yields a vector $v \in \mathbb{R}^m$ and a scalar $c \in \mathbb{R}$ such that

$$\langle v, a\rangle < c \leq \langle v, b\rangle$$

for all $a \in A$ and $b \in B$.

Now, consider $g \colon [0,1] \to \mathbb{R}$ be defined by

$$g(t) = \langle v, f(p + t(q-p))\rangle = v^{\mathrm{T}} f(p + t(q-p)).$$

Then g is the composition of differentiable functions (see below) and hence is differentiable. So by the one dimensional Mean Value Theorem we have for some $t \in (0,1)$

$$\langle v, f(q) - f(p)\rangle = g(1) - g(0) = g'(t).$$

Now, $g = h_1 \circ f \circ h_2$, where $h_1 \colon \mathbb{R}^m \to \mathbb{R}$ is defined by $h_1(x) = v^{\mathrm{T}} x$ and $h_2 \colon \mathbb{R} \to \mathbb{R}^n$ is defined by $h_2(t) = p + t(p-q)$. These are linear operators (plus a constant for h_2), and so $(Dh_1)_x = v^{\mathrm{T}}$ for all $x \in \mathbb{R}^m$ and $(Dh_2)_t = q - p$ for all $t \in [0,1]$. Thus the Chain Rule implies

$$g'(t) = (Dh_1)_{f\circ h_2(t)} (Df)_{h_2(t)}(q-p) = v^{\mathrm{T}}(Df)_{p+t(q-p)}(q-p) = \left\langle v, (Df)_{p+t(q-p)}(q-p)\right\rangle.$$

Combining this with the previous equality yields for some $t \in (0,1)$

$$\langle v, f(q) - f(p)\rangle = \left\langle v, (Df)_{p+t(q-p)}(q-p)\right\rangle.$$

But this contradicts $\langle v, a\rangle < c \leq \langle v, b\rangle$ for all $a \in A$ and $b \in B$. Thus there must exist some $\theta \in [p,q]$ such that $f(q) - f(p) = (Df)_\theta(q-p)$. □

4. Observe that $U \ni p \mapsto (Df)_p = 0$ is continuous. Hence f is of class C^1 and therefore by the C^1 Mean Value Theorem we have for any segment $[p,q] \subset U$

$$f(q) - f(p) = \left(\int_0^1 (Df)_{p+t(q-p)} dt\right)(q-p) = \left(\int_0^1 0\ dt\right)(q-p) = 0(q-p) = 0.$$

That is $f(q) = f(p)$.

Now, fix $p \in U$ and let

$$A = \{q \in U \colon q \text{ can be connected to } p \text{ via segments contained in } U\}.$$

So if $q \in A$, then there exists $q_1, \ldots, q_n \in U$ such that $[p, q_1], [q_1, q_2], \ldots, [q_n, q] \subset U$. By the above argument, we therefore have

$$f(p) = f(q_1) = f(q_2) = \cdots = f(q_n) = f(q).$$

Thus $f(q) = f(p)$ for all $q \in A$. If we can show that $A = U$, then we will have shown f is constant. First note that A is open. Indeed, if $q \in A \subset U$, then there is a ball centered at q contained in U. Every element in this ball can be connected to q via a segment contained in U (namely a radius of the ball), and hence every element can be connected to p by first using the segments that reach q followed by the radius of the ball. Hence this ball is also contained in A, and A is open. We also claim that $U \setminus A$ is open. Indeed, if $q \in U \setminus A$ then we can once more find a ball centered at q and contained in U. Now, no element of this ball can lie in A since then the series of segments in U connecting it to p followed by a radius of the ball would yield a series of segments in U connecting q to p, contradicting $q \notin A$. Thus the ball is contained in $U \setminus A$, which means $U \setminus A$ is open. Since U is connected, we can only have A and $U \setminus A$ open if one of them is empty. We clearly have $p \in A$, so it must be that $U \setminus A = \emptyset$. Hence $A = U$ as desired. □

5. Fix $y \in E$ and let $\epsilon > 0$. For each $x \in [a,b]$, we use the continuity of f to find $\delta(x) > 0$ such that if $(x', y') \in E$ satisfies $d_2((x,y),(x',y')) < \delta(x)$, then

$$|f(x,y) - f(x',y')| < \frac{\epsilon}{2(b-a)}.$$

For each $x \in [a,b]$, denote $I_x := (x - \frac{1}{\sqrt{2}}\delta(x), x + \frac{1}{\sqrt{2}}\delta(x))$. Observe that $\{I_x\}_{x\in[a,b]}$ is an open cover for the compact set $[a,b]$. Hence there exists a finite subcover:

$$\{I_{x_1}, \ldots, I_{x_n}\};$$

Define $\delta := \min\{\frac{1}{\sqrt{2}}\delta(x_1), \ldots, \frac{1}{\sqrt{2}}\delta(x_n)\} > 0$.

Now, suppose $y' \in E$ satisfies $d(y,y') < \delta$. Then for any $x \in [a,b]$, there is some $j \in \{1,\ldots,n\}$ such that $x \in I_{x_j}$, which means $|x - x_j| < \frac{1}{\sqrt{2}}\delta(x_j)$. Consequently,

$$d_2((x,y'),(x_j,y)) = \sqrt{|x-x_j|^2 + d(y',y)^2} < \sqrt{\frac{1}{2}\delta(x_j)^2 + \delta^2} \leq \sqrt{\delta(x_j)^2} = \delta(x_j)$$

and

$$d_2((x_j,y),(x,y)) = \sqrt{|x_j - x|^2 + d(y,y)} < \sqrt{\frac{1}{2}\delta(x_j)^2} < \delta(x_j).$$

By definition of $\delta(x_j)$, this implies

$$|f(x,y') - f(x,y)| \leq |f(x,y') - f(x_j,y)| + |f(x_j,y) - f(x,y)| < \frac{\epsilon}{2(b-a)} + \frac{\epsilon}{2(b-a)} = \frac{\epsilon}{b-a}.$$

Since this holds for all $x \in [a,b]$, we have

$$|F(y') - F(y)| \leq \int_a^b |f(x,y) - f(x,y')|\ dx < \int_a^b \frac{\epsilon}{b-a}\ dx = \epsilon.$$

Thus F is continuous at y. Since $y \in E$ was arbitrary, we see that F is continuous on E. □