MAT257 Assignment 4

(Author's name here)

October 15, 2021

1. For all parts of this question, we are given a continuous function $g : \mathbb{R} \to \mathbb{R}$.
(a) We will find the partial derivatives of the following function:

$$f(x, y) := \int_a^{x+y} g.$$

First, keeping y constant, we obtain:

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial x} &= \frac{\partial}{\partial x} \int_a^{x+y} g \\
&= \frac{d}{du} \int_a^u g \bigg|_{u=x+y} \cdot \frac{\partial}{\partial x}(x+y) \qquad \text{(Applying Chain Rule)} \\
&= g(x+y) \cdot \frac{\partial}{\partial x}(x+y) \qquad \text{(Applying Fundamental Theorem of Calculus)} \\
&= g(x+y) \cdot 1 \qquad \text{(Keeping } y \text{ constant)} \\
&= g(x+y).
\end{aligned}$$

Next, keeping x constant, we obtain:

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial y} &= \frac{\partial}{\partial y} \int_a^{x+y} g \\
&= \frac{d}{du} \int_a^u g \bigg|_{u=x+y} \cdot \frac{\partial}{\partial y}(x+y) \qquad \text{(Applying Chain Rule)} \\
&= g(x+y) \cdot \frac{\partial}{\partial y}(x+y) \qquad \text{(Applying Fundamental Theorem of Calculus)} \\
&= g(x+y) \cdot 1 \qquad \text{(Keeping } x \text{ constant)} \\
&= g(x+y).
\end{aligned}$$

Overall, we obtain:

$$\boxed{\frac{\partial f(x,y)}{\partial x} = \frac{\partial f(x,y)}{\partial y} = g(x+y).}$$

□

(b) We will find the partial derivatives of the following function:

$$f(x, y) := \int_y^{x} g.$$

First, keeping y constant, we obtain:

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial x} &= \frac{\partial}{\partial x} \int_y^{x} g \\
&= g(x). \qquad \text{(Applying Fundamental Theorem of Calculus while keeping } y \text{ constant)}
\end{aligned}$$

Next, keeping x constant, we obtain:

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial y} &= \frac{\partial}{\partial y} \int_y^{x} g \\
&= \frac{\partial}{\partial y}\left(- \int_x^{y} g \right) \\
&= -g(y). \qquad \text{(Applying Fundamental Theorem of Calculus while keeping } x \text{ constant)}
\end{aligned}$$

Overall, we obtain:

$$\boxed{\frac{\partial f(x,y)}{\partial x} = g(x), \qquad \frac{\partial f(x,y)}{\partial y} = -g(y)}.$$

$\square$

(c) We will find the partial derivatives of the following functions:

$$f(x,y) := \int_a^{xy} g.$$

First, keeping y constant, we obtain:

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial x} &= \frac{\partial}{\partial x} \int_a^{xy} g \\
&= \frac{d}{du} \int_a^u g \bigg|_{u=xy} \cdot \frac{\partial}{\partial x} xy \qquad \text{(Applying Chain Rule)} \\
&= g(xy) \cdot \frac{\partial}{\partial x} xy \qquad \text{(Applying Fundamental Theorem of Calculus)} \\
&= g(xy)y. \qquad \text{(Keeping } y \text{ constant)}
\end{aligned}$$

Next, keeping x constant, we obtain:

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial y} &= \frac{\partial}{\partial y} \int_a^{xy} g \\
&= \frac{d}{du} \int_a^u g \bigg|_{u=xy} \cdot \frac{\partial}{\partial y} xy \qquad \text{(Applying Chain Rule)} \\
&= g(xy) \cdot \frac{\partial}{\partial y} xy \qquad \text{(Applying Fundamental Theorem of Calculus)} \\
&= g(xy)x. \qquad \text{(Keeping } x \text{ constant)}
\end{aligned}$$

Overall, we obtain:

$$\boxed{\frac{\partial f(x,y)}{\partial x} = g(xy)y, \qquad \frac{\partial f(x,y)}{\partial y} = g(xy)x.}$$

$\square$

(d) We will find the partial derivatives of the following function:

$$f(x,y) := \int_a^{\int_b^y g} g.$$

First, keeping y constant, we notice that x does not contribute to the formula of $f(x,y)$, so $f(x,y)$ becomes constant. Thus, applying MAT157 material, we obtain $\frac{\partial}{\partial x} f(x,y) = 0$.
Next, keeping x constant, we obtain:

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial y} &= \frac{\partial}{\partial y} \int_a^{\int_b^y g} g \\
&= \frac{d}{du} \int_a^u g \bigg|_{u=\int_b^y g} \cdot \frac{\partial}{\partial y} \int_b^y g \qquad \text{(Applying Chain Rule)} \\
&= g\left(\int_b^y g\right) g(y). \qquad \text{(Applying Fundamental Theorem of Calculus twice)}
\end{aligned}$$

Overall, we obtain:

$$\boxed{\frac{\partial f(x,y)}{\partial x} = 0, \qquad \frac{\partial f(x,y)}{\partial y} = g\left(\int_b^y g\right) g(y)}.$$

$\square$

2. Given the function:

$$f(x,y) := x^{x^{x^{x^y}}} + (\log x)(\arctan(\arctan(\arctan(\sin(\cos xy) - \log(x+y))))),$$

we will find the partial derivative $D_2 f(1,y)$.
First, for all $y_0 \in \mathbb{R}$, we have:

$$f(1,y_0) = 1^{1^{1^{1^{y_0}}}} + (\log 1) \cdot (\text{some number}) = 1 + 0 = 1.$$

Then, by definition of partial derivatives, we obtain the following for all $y \in \mathbb{R}$:

$$D_2 f(1,y) = \lim_{h\to 0} \frac{f(1,y+h) - f(1,y)}{h} = \lim_{h\to 0} \frac{1-1}{h} = \lim_{h\to 0} 0 = \boxed{0}.$$

$\square$

3. Given continuous functions $g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$, we define $f : \mathbb{R}^2 \to \mathbb{R}$ by:

$$f(x,y) = \int_0^x g_1(t,0)dt + \int_0^y g_2(x,t)dt.$$

(a) First, we will show that $\frac{\partial}{\partial y} f(x,y) = g_2(x,y)$.
Keeping x constant, we obtain:

$$\begin{aligned}
\frac{\partial}{\partial y} f(x,y) &= \frac{\partial}{\partial y}\left(\int_0^x g_1(t,0)dt + \int_0^y g_2(x,t)dt\right) \\
&= \frac{\partial}{\partial y}\int_0^y g_2(x,t)dt \qquad \text{(Since } \int_0^x g_1(t,0)dt \text{ is constant if } x \text{ is constant)} \\
&= g_2(x,y), \qquad \text{(Applying Fundamental Theorem of Calculus)}
\end{aligned}$$

as required. □

(b) Next, we will find a function $f_b : \mathbb{R}^2 \to \mathbb{R}$ such that $\frac{\partial}{\partial x} f(x,y) = g_1(x,y)$.
Let us define $f_b : \mathbb{R}^2 \to \mathbb{R}$ by "swapping" the first and second coordinates in the definition of f above. Formally, we define:

$$f_b(x,y) := \int_0^y g_2(0,t)dt + \int_0^x g_1(t,y)dt.$$

(These integrals exist because g_1, g_2 are continuous.) Then, keeping y constant, we obtain:

$$\begin{aligned}
\frac{\partial}{\partial x} f_b(x,y) &= \frac{\partial}{\partial x}\left(\int_0^y g_2(0,t)dt + \int_0^x g_1(t,y)dt\right) \\
&= \frac{\partial}{\partial x}\int_0^x g_1(t,y)dt \qquad \text{(Since } \int_0^y g_2(0,t)dt \text{ is constant if } y \text{ is constant)} \\
&= g_1(x,y), \qquad \text{(Applying Fundamental Theorem of Calculus)}
\end{aligned}$$

as required. □

(c) Next, we will find a function $f_c : \mathbb{R}^2 \to \mathbb{R}$ such that $\frac{\partial}{\partial x} f_c(x,y) = x$ and $\frac{\partial}{\partial y} f_c(x,y) = y$.
Let us define $f_c : \mathbb{R}^2 \to \mathbb{R}$ by $f_c(x,y) := \frac{1}{2}(x^2 + y^2)$. Then, keeping y constant, we obtain:

$$\begin{aligned}
\frac{\partial}{\partial x} f_c(x,y) &= \frac{\partial}{\partial x}\frac{1}{2}(x^2 + y^2) \\
&= \frac{1}{2} \cdot 2x \qquad \text{(Keeping } y \text{ constant)} \\
&= x.
\end{aligned}$$

Keeping x constant, we also obtain:

$$\begin{aligned}
\frac{\partial}{\partial y} f_c(x,y) &= \frac{\partial}{\partial y}\frac{1}{2}(x^2 + y^2) \\
&= \frac{1}{2} \cdot 2y \qquad \text{(Keeping } x \text{ constant)} \\
&= y.
\end{aligned}$$

Therefore, our choice of $f_c : \mathbb{R}^2 \to \mathbb{R}$ satisfies $\frac{\partial}{\partial x} f_c(x,y) = x$ and $\frac{\partial}{\partial y} f_c(x,y) = y$, as required. □

(d) Finally, we will find a function $f_d : \mathbb{R}^2 \to \mathbb{R}$ such that $\frac{\partial}{\partial x} f_d(x,y) = y$ and $\frac{\partial}{\partial y} f_d(x,y) = x$.

Let us define $f_d : \mathbb{R}^2 \to \mathbb{R}$ by $f_d(x,y) := xy$. Then, keeping y constant, we obtain:

$$\begin{aligned} \frac{\partial}{\partial x} f_d(x,y) &= \frac{\partial}{\partial x} xy \\ &= y. \qquad \text{(Keeping } y \text{ constant)} \end{aligned}$$

Keeping x constant, we also obtain:

$$\begin{aligned} \frac{\partial}{\partial y} f_d(x,y) &= \frac{\partial}{\partial y} xy \\ &= x. \qquad \text{(Keeping } x \text{ constant)} \end{aligned}$$

Therefore, our choice of $f_d : \mathbb{R}^2 \to \mathbb{R}$ satisfies $\frac{\partial}{\partial x} f_d(x,y) = y$ and $\frac{\partial}{\partial y} f_d(x,y) = x$, as required. $\square$

4. Given $f : \mathbb{R}^n \to \mathbb{R}$ and $x \in \mathbb{R}^n$, we define the directional derivative in the direction x to be:

$$D_x f(a) := \lim_{h\to 0} \frac{f(a+hx)-f(a)}{h}$$

whenever this limit exists.
(a) First, we will show that $D_{e_i} f(a) = D_i f(a)$, where e_i is the standard basis vector in $\mathbb{R}^n$ whose i^{th} coordinate is 1 and whose other coordinates are 0.
By definition of directional derivatives, we have:

$$\begin{aligned} D_{e_i} f(a) &= \lim_{h\to 0} \frac{f(a+he_i)-f(a)}{h} \\ &= \lim_{h\to 0} \frac{f((a_1,\ldots,a_{i-1},a_i,a_{i+1},\ldots,a_n)+h(0,\ldots,0,1,0,\ldots,0))-f(a_1,\ldots,a_n)}{h} \\ &= \frac{f(a_1,\ldots,a_{i-1},a_i+h,a_{i+1},\ldots,a_n)-f(a_1,\ldots,a_n)}{h}, \end{aligned}$$

and by definition of partial derivatives, this equals $D_i f(a)$. Therefore, $D_{e_i} f(a) = D_i f(a)$, as required. $\square$
(b) Next, we will show that $D_{tx} f(a) = tD_x f(a)$ for all $t \in \mathbb{R}$.
(Note: Following the grader's advice, the $t = 0$ case was removed from the solution to clean it up after it was graded.)
We will assume $t \neq 0$ so that $tx \neq 0$ is a valid direction and the directional derivative $D_{tx} f(a)$ makes sense. By definition, we know that:

$$D_{tx} f(a) = \lim_{k\to 0} \frac{f(a+k\cdot tx)-f(a)}{k}.$$

We wish to show that this limit equals $tD_x f(a)$. Let any $\epsilon > 0$ be given. By definition, we also know that:

$$D_x f(a) = \lim_{h\to 0} \frac{f(a+hx)-f(a)}{h}.$$

Then, since $t \neq 0$, we have $\frac{\epsilon}{|t|} > 0$, so there exists some $\delta' > 0$ such that all $h \in \mathbb{R}$ which satisfy $0 < |h| < \delta'$ also satisfy:

$$\left|\frac{f(a+hx)-f(a)}{h} - D_x f(a)\right| < \frac{\epsilon}{|t|}. \qquad (*)$$

Now, let us pick $\delta := \frac{\delta'}{|t|} > 0$. Then, for all $k \in \mathbb{R}$ which satisfy $0 < |k| < \delta$, we obtain:

$$\begin{aligned} \left|\frac{f(a+ktx)-f(a)}{k} - tD_x f(a)\right| &= \left|t\cdot\frac{f(a+ktx)-f(a)}{kt} - tD_x f(a)\right| \\ &= |t|\left|\frac{f(a+ktx)-f(a)}{kt} - D_x f(a)\right|. \end{aligned}$$

Next, let us choose the substitution $h := kt$. Since t and k are nonzero, we obtain $|h| = |k||t| > 0$, and we also obtain:

$$|h| = |k||t| < \delta|t| = \frac{\delta'}{|t|}|t| = \delta'.$$

Overall, $0 < |h| < \delta'$, so $(*)$ applies to h. As a result:

$$\begin{aligned}
\left|\frac{f(a+ktx)-f(a)}{k} - tD_x f(a)\right| &= |t|\left|\frac{f(a+ktx)-f(a)}{kt} - D_x f(a)\right| \\
&= |t|\left|\frac{f(a+hx)-f(a)}{h} - D_x f(a)\right| \\
&< |t| \cdot \frac{\epsilon}{|t|} \qquad \text{(Applying } (*)) \\
&= \epsilon.
\end{aligned}$$

Thus, for all $\epsilon > 0$, we found $\delta > 0$ such that all $k \in \mathbb{R}$ which satisfy $0 < |k| < \delta$ also satisfy:

$$\left|\frac{f(a+ktx)-f(a)}{k} - tD_x f(a)\right| < \epsilon,$$

so we conclude that:

$$D_{tx} f(a) = \lim_{k \to 0} \frac{f(a+ktx)-f(a)}{k} = tD_x f(a)$$

for all nonzero t. Overall, we proved that $D_{tx} f(a) = tD_x f(a)$ for all $t \in \mathbb{R}$ no matter whether t is zero or nonzero, as required. □

(c) Finally, if f is differentiable at a, we will show that $D_x f(a) = Df(a)(x)$ and conclude that $D_{x+y} f(a) = D_x f(a) + D_y f(a)$.

To prove that $D_x f(a) = Df(a)(x)$, let us consider the following two cases:

Case 1: $x = 0$. Then, since $Df(a)$ is a linear map, we obtain $Df(a)(x) = Df(a)(0) = 0$. We also obtain:

$$\begin{aligned}
D_x f(a) &= D_0 f(a) \\
&= \lim_{t \to 0} \frac{f(a + t \cdot 0) - f(a)}{t} \\
&= \lim_{t \to 0} \frac{f(a) - f(a)}{t} \\
&= 0 \\
&= Df(a)(x).
\end{aligned}$$

Thus, $D_x f(a) = Df(a)(x)$ for this case, as desired.

Case 2: $x \neq 0$. Then, since $D_x f(a)$ is defined to be $\lim_{t \to 0} \frac{f(a+tx)-f(a)}{t}$, it is enough to show that this limit exists and equals $Df(a)(x)$. To do this, let any $\epsilon > 0$ be given. Then, by definition of $Df(a)$, we know that the error function $e(h) := f(a+h) - f(a) - Df(a)(h)$ is in $o(h)$. As a result,

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - Df(a)(h)}{|h|} = 0.$$

Then, since $x \neq 0$, we have $\frac{\epsilon}{|x|} > 0$, so there exists $\delta' > 0$ such that all $h \in \mathbb{R}^n$ which satisfy $0 < |h| < \delta'$ also satisfy:

$$\left|\frac{f(a+h) - f(a) - Df(a)(h)}{|h|}\right| < \frac{\epsilon}{|x|}. \qquad (1)$$

Now, since $x \neq 0$, we can define $\delta := \frac{\delta'}{|x|} > 0$. Then, for all $t \in \mathbb{R}$ which satisfy $0 < |t| < \delta$, since t and x are nonzero, we have $|tx| > 0$. We also have:

$$|tx| = |t||x| < \delta|x| = \frac{\delta'}{|x|}|x| = \delta'.$$

Overall, we obtain $0 < |tx| < \delta'$, so we can plug $h = tx$ into (1) to obtain:

$$\begin{aligned}
\frac{\epsilon}{|x|} &> \left|\frac{f(a+tx) - f(a) - Df(a)(tx)}{|tx|}\right| \\
&= \left|\frac{f(a+tx) - f(a) - Df(a)(tx)}{|t||x|}\right| \\
&= \frac{1}{|x|}\left|\frac{f(a+tx) - f(a) - Df(a)(tx)}{t}\right| \\
&= \frac{1}{|x|}\left|\frac{f(a+tx) - f(a) - tDf(a)(x)}{t}\right| && \text{(Since } Df(a) \text{ is linear)} \\
&= \frac{1}{|x|}\left|\frac{f(a+tx) - f(a)}{t} - Df(a)(x)\right|.
\end{aligned}$$

Multiplying both sides by $|x|$, we obtain:

$$\left|\frac{f(a+tx) - f(a)}{t} - Df(a)(x)\right| < \epsilon. \tag{2}$$

Therefore, for all $\epsilon > 0$, we found $\delta > 0$ such that all $t \in \mathbb{R}$ which satisfy $0 < |t| < \delta$ also satisfy (2), so we conclude that:

$$D_x f(a) = \lim_{t \to 0} \frac{f(a+tx) - f(a)}{t} = Df(a)(x).$$

At this point, we have proven that $D_x f(a) = Df(a)(x)$ for all $x \in \mathbb{R}^n$ no matter whether x is zero or nonzero, as required. Finally, for all $x, y \in \mathbb{R}^n$, it follows that:

$$D_{x+y} f(a) = Df(a)(x+y) = Df(a)(x) + Df(a)(y) = D_x f(a) + D_y f(a),$$

where the middle step holds because $Df(a)$ is a linear map, as required. $\square$

5. If a function $f : \mathbb{R}^n \to \mathbb{R}$ is homogeneous of degree m and differentiable, we will prove that:

$$\sum_{i=1}^{n} x_i D_i f(x) = m f(x)$$

for all $x \in \mathbb{R}^n$.

Let any $x \in \mathbb{R}^n$ be given. First, let us define the function $g : \mathbb{R} \to \mathbb{R}$ by $g(t) := f(tx)$. Then, we will compute $g'(1)$ in two different ways. On the one hand, since f is homogeneous of degree m, we have $g(t) = f(tx) = t^m f(x)$. Then, since $f(x)$ is a constant coefficient, we obtain $g'(t) = mt^{m-1} f(x)$, and plugging in $t = 1$ yields $g'(1) = m \cdot 1^{m-1} f(x) = mf(x)$. On the other hand, if we define $h : \mathbb{R} \to \mathbb{R}^n$ by $h(t) := tx$, then we can write g as the composition:

$$g(t) = f(tx) = f(h(t)) = (f \circ h)(t).$$

Now, since x is constant, $h(t) = xt$ is a linear map in terms of t with matrix $x = (x_1, \ldots, x_n)^T$. Then, Spivak's Theorem 2-3(2) tells us that $h'(t) = (x_1, \ldots, x_n)^T$. Next, since f is differentiable, we can apply Spivak's Theorem 2-7 to express the differential of f at any point $y \in \mathbb{R}^n$ as $Df(y) = (D_1 f(y), \ldots, D_n f(y))$. Then, applying the Chain Rule, we obtain:

$$g'(t) = f'(h(t)) \cdot h'(t) = f'(tx) \cdot h'(t) = (D_1 f(tx), \ldots, D_n f(tx)) \cdot (x_1, \ldots, x_n)^T = \sum_{i=1}^{n} x_i D_i f(tx).$$

Plugging in $t = 1$, we obtain $g'(1) = \sum_{i=1}^{n} x_i D_i f(x)$. Therefore, since $g'(1)$ equals both $\sum_{i=1}^{n} x_i D_i f(x)$ and $mf(x)$, we conclude that $\sum_{i=1}^{n} x_i D_i f(x) = mf(x)$, as required. $\square$

6. Given any differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ such that $f(0) = 0$, we will prove that there exist functions $g_i : \mathbb{R}^n \to \mathbb{R}$ such that:

$$f(x) = \sum_{i=1}^{n} x_i g_i(x).$$

(Note: Following the grader's advice, the solution below was edited to clean up the presentation after it was graded.)

We will define the functions g_i as follows. For $x = 0$, we define $g_i(0) = 0$ for all $1 \leq i \leq n$. Next, for any point $x \neq 0$, x has at least one nonzero coordinate. Then, we can define the values $g_j(x)$ at this point x as follows:

$$g_j(x) = \begin{cases} \frac{f(x)}{x_j}, & \text{if } j \text{ is the smallest index such that } x_j \text{ is nonzero;} \\ 0, & \text{otherwise.} \end{cases}$$

It remains to show that $f(x) = \sum_{i=1}^{n} x_i g_i(x)$ at all points $x \in \mathbb{R}^n$. If $x = 0$, we have:

$$\begin{aligned} \text{LHS} &= f(0) \\ &= 0 \end{aligned} \qquad\qquad \begin{aligned} \text{RHS} &= \sum_{i=1}^{n} 0 \cdot g_i(0) \\ &= 0, \end{aligned}$$

so LHS = RHS. If $x \neq 0$, x has at least one nonzero coordinate, so let j be the smallest index such that x_j is nonzero. Then, we obtain:

$$\begin{aligned} \text{RHS} &= \sum_{i=1}^{n} x_i g_i(x) \\ &= \sum_{\substack{1 \leq i \leq n \\ i \neq j}} x_i g_i(x) + x_j g_j(x) \\ &= \sum_{\substack{1 \leq i \leq n \\ i \neq j}} x_i \cdot 0 + x_j \cdot \frac{f(x)}{x_j} && \text{(Applying definitions of } g_i(x) \text{ and } g_j(x)) \\ &= 0 + f(x) \\ &= \text{LHS}, \end{aligned}$$

so LHS = RHS. Overall, we proved that our choices for $g_i : \mathbb{R}^n \to \mathbb{R}$ satisfy $f(x) = \sum_{i=1}^{n} x_i g_i(x)$ for all $x \in \mathbb{R}^n$, as required. $\square$

Notes on intuition

Now, let us develop some intuition on how to approach these problems and find these solutions. (Note: This section was not submitted on Crowdmark.)

1. Overall, all sub-problems of this problem had $f(x,y)$ defined as some integral, which motivates us to solve all of them using the Fundamental Theorem of Calculus. In fact, this approach works for the entire problem, with some small tricks for a few sub-problems. For instance, part (b) needed the standard trick of reversing the integral's direction from $\int_y^x g$ to $-\int_x^y g$. Moreover, part (d) may seem intimidating at first, but after we use the Chain Rule to solve parts (a) and (c), this helps us to realize that the Chain Rule also helps with part (d).

2. Since we are finding $D_2 f(1,y)$, we only care about how f changes along the y-direction, so we might as well plug in $x=1$ directly and keep x constant. It turns out that plugging in $x=1$ simplifies the problem dramatically, and it is easy to finish. (This question may remind you of the social media posts of the form "Find (*complicated and scary expression*) $\cdot\, 0$".)

3. First, for part (a), we observe that the first integral is independent of y, so it will not change when we move along the y-direction to find $D_2 f(x,y)$. Then, we focus on the second integral, which is relatively simple to differentiate along y, similarly to Problem 1.
   Next, the problem statement for part (b) is very similar to the statement for part (a). This helps us to infer that the solution for part (b) will mirror that of part (a). Indeed, we construct $f_b : \mathbb{R}^2 \to \mathbb{R}$ such that:

   - The first integral in f_b is independent of x, similarly to how the first integral of f is independent of y.
   - The second integral in f_b is easy to differentiate with the Fundamental Theorem of Calculus, with x only appearing once in the expression, similarly to the second integral of f.

   Now, it is trickier to construct f_c and f_d for parts (c) and (d) because we care about partial derivatives along both directions. For part (c), we notice that $D_1 f_c(x,y)$ only depends on x and that $D_2 f_c(x,y)$ only depends on y. In other words, the change in the x-direction only depends on x, and the change in the y-direction only depends on y. This suggests that f_c can be decomposed into a component which only depends on x and a component which only depends on y. Indeed, the integral of x with respect to x is $\frac{1}{2}x^2$, so we deduce that the x-component should be $\frac{1}{2}x^2$, and we deduce similarly that the y-component should be $\frac{1}{2}y^2$. Thus, $f_c(x,y) = \frac{1}{2}x^2 + \frac{1}{2}y^2$.
   We cannot use the same trick for part (d). Instead, we could try to integrate y along x to obtain xy. In fact, once we try $f_d(x,y) := xy$, we discover that this f_d already works, so we are done.

4. (a) Intuitively, $D_x f(a)$ represents the rate of change of f along the x-direction, so $D_{e_i} f(a)$ represents the rate of change along the e_i-direction. On the other hand, we also defined the partial derivative $D_i f(a)$ so that it represents the rate of change along the e_i-direction. This tells us that the equality $D_{e_i} f(a) = D_i f(a)$ should follow from the definitions of D_{e_i} and D_i – indeed, our solution confirms this.
   (b) Intuitively, when we travel along the tx-direction, then f should change t times as fast as when we travel along the x-direction. Then, in some sense, the situation for $D_{tx} f(a)$ is essentially the situation for $D_x f(a)$, scaled by a factor of t. This motivates us to consider $\frac{\epsilon}{|t|}$ and $\frac{\delta'}{|t|}$ in our solution. Then, the rest of the solution consists of epsilon-delta formalities.

(c) Our proof that $D_x f(a) = Df(a)(x)$ is motivated by the solution for part (a). Since we are differentiating along the x-direction instead of the e_i-direction, we also have to consider scalings of ϵ and δ by a factor of $|x|$, similarly to part (b). From here on, the proof for $D_x f(a) = Df(a)(x)$ mostly consists of epsilon-delta formalities. Finally, the proof for $D_{x+y} f(a) = D_x f(a) + D_y f(a)$ follows almost directly from plugging in $D_x f(a) = Df(a)(x)$.

5. Our solution for this problem is strongly motivated by the textbook's hint. Let us discuss how one might be able to approach this problem without the hint:

   - First, consider the left-hand side, $\sum_{i=1}^n x_i D_i f(x)$. Each term $D_i f(x)$ means the partial derivative in the e_i-direction (where $e_1, \ldots, e_n$ is the standard basis of $\mathbb{R}^n$), and each coefficient x_i means that we move in the e_i direction at a rate of x_i. Combining these movements, we find that we are observing the change in f as we move in the x-direction. This procedure reminds us of the directional derivative introduced in Problem 4.
   - Additionally, we hope to use homogeneity somewhere in our solution, and we can only use it for multiples of x. In other words, starting from x, we must move in the x-direction to continue obtaining multiples of x. This is a second hint that we should consider the directional derivative in the x-direction.

   Combining these hints, we formalize this directional derivative idea using $g(t) := f(tx)$, which inputs multiples of x into f. Finally, it is relatively straightforward to apply this definition and compute $g'(1)$ to solve this problem.

6. We presented an unconventional solution, in the sense that it does not use calculus. This solution was motivated by the fact that writing $f(x)$ as a sum:

   $$f(x) = \sum_{i=1}^n g_i(x) x_i$$

   amounts to writing $f(x)$ as a linear combination of the x_i-coordinates. Note that $f(x), x_1, \ldots, x_n$ are in the one-dimensional vector space $\mathbb{R}^1$. Then, if one of the vectors $x_1, \ldots, x_n$ is nonzero, then $(x_1, \ldots, x_n)$ will span $\mathbb{R}^1$, which allows the linear combination above to exist. This approach will work for all nonzero $x \in \mathbb{R}^n$, and the special case $x = 0$ is easy once we are given $f(0) = 0$.