Solution suggestions

Exercise 1. (i) In '(x,y)-notation', f has the following appearance:

$$f(x,y) = \sqrt{x^2+y^2}.$$

With this representation, it is easy to calculate the partial derivatives using standard differentiation rules (more specifically, the chain rule). We obtain

$$\frac{\partial f}{\partial x}(x,y) = \frac{1}{2\sqrt{x^2+y^2}} \cdot 2x = \frac{x}{\sqrt{x^2+y^2}}, \quad \frac{\partial f}{\partial y}(x,y) = \frac{1}{2\sqrt{x^2+y^2}} \cdot 2y = \frac{y}{\sqrt{x^2+y^2}}.$$

We have here implicitly used that we calculate the derivative in a point different from the origin in that the square root only is differentiable in $]0,\infty[$. Now, again since we are speaking about points different from the origin, these two functions are continuous (the denominator is not equal to zero in such points). Therefore, f is differentiable, and the derivative is given by

$$f'(x,y) = \begin{bmatrix} \frac{\partial f}{\partial x}(x,y) & \frac{\partial f}{\partial y}(x,y) \end{bmatrix} = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \end{bmatrix}.$$

(ii) Here, the best idea is probably to actually resort to the limit defining $D_{\mathbf{v}}f(\mathbf{p})$. We have for $\mathbf{v}$ arbitrary

$$D_{\mathbf{v}}f(\mathbf{0}) = \lim_{t\to 0} \frac{f(\mathbf{0}+t\mathbf{v}) - f(\mathbf{0})}{t} = \lim_{t\to 0} \frac{|t\mathbf{v}| - |\mathbf{0}|}{t} = \lim_{t\to 0} \frac{|t|}{t} |\mathbf{v}|$$

This limit does not exist, since the fraction approaches different values depending on if we approach the zero from the right or from the left. Hence, the directional derivatives do not exist, and the function is surely not differentiable.

Note: Some authors define the directional limit through a one-sided limit:

$$\Delta_{\mathbf{v}}f(\mathbf{p}) = \lim_{t\to 0^+} \frac{f(\mathbf{p}+t\mathbf{v}) - f(\mathbf{p})}{t}$$

(the different notation is to differentiate from our definition). This has its motivation - it might in some situations be beneficial to be able to capture the difference in response in f-value when we move in $\mathbf{v}$-direction rather than $-\mathbf{v}$-direction! With this definition, the function at hand would have directional derivatives

$$\Delta_{\mathbf{v}}(f(\mathbf{0})) = |\mathbf{v}|.$$

However, the function is still not differentiable, since $\mathbf{v} \mapsto \mathbf{v}$ surely isn't linear.

□

Exercise 2. (i)

(a)

$$f'(x,y) = \begin{bmatrix} e^y & xe^y \end{bmatrix}, \quad \Rightarrow \quad f'(g(s,t)) = f'(st, s+t) = \begin{bmatrix} e^{s+t} & ste^{s+t} \end{bmatrix} \quad g'(s,t) = \begin{bmatrix} t & s \\ 1 & 1 \end{bmatrix}$$

$$f'(g(s,t))g'(s,t) = \begin{bmatrix} e^{s+t} & ste^{s+t} \end{bmatrix} \begin{bmatrix} t & s \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} te^{s+t} + ste^{s+t} & se^{s+t} + ste^{s+t} \end{bmatrix}.$$

(b)

$$f'(x,y,z) = \begin{bmatrix} y & x & 0 \\ 0 & z & y \end{bmatrix} \Rightarrow f'(g(x,y)) = f'(x,y,x^2+y^2) = \begin{bmatrix} y & x & 0 \\ 0 & x^2+y^2 & y \end{bmatrix} \quad g'(x,y,z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2x & 2y \end{bmatrix}$$

$$f'(g(x,y))g'(x,y) = \begin{bmatrix} y & x & 0 \\ 0 & x^2+y^2 & y \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2x & 2y \end{bmatrix} = \begin{bmatrix} y & x \\ 2xy & x^2+3y^2 \end{bmatrix}.$$

(c) $f'(t) = 2t$, so that $f(g(t)) = 2g(t)$. The chain rule thus yields $(f \circ g)'(t) = 2g(t)g'(t)$.

(d) Let us begin by noticing that $f \circ g = f \circ f \circ f$. This implies that

$$(f \circ g)' = f'\left((f \circ f)(x,y)\right)(f \circ f)'(0,0) = f'\left(f(f(0,0))\right) f'(f(0,0))f'(0,0).$$

We now calculate

$$f'(x,y) = \begin{bmatrix} 1 & 2y \\ 2x & 1 \end{bmatrix}.$$

Since $f(0,0) =< 0,0 >$, we obtain

$$f'\left(f(f(0,0))\right) f'(f(0,0))f'(0,0) = f'(0,0)f'(0,0)f'(0,0) = \begin{bmatrix} 1 & 2 \cdot 0 \\ 2 \cdot 0 & 1 \end{bmatrix}^3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(ii) The functions m and Φ have the derivatives

$$m'(x,y) = \begin{bmatrix} y & x \end{bmatrix}, \quad \Phi'(t) = \begin{bmatrix} f'(t) \\ g'(t) \end{bmatrix}.$$

Since $fg = m \circ \Phi$, we obtain

$$(fg)'(t) = m'(\Phi(t))\Phi'(t) = \begin{bmatrix} g(t) & f(t) \end{bmatrix} \begin{bmatrix} f'(t) \\ g'(t) \end{bmatrix} = g(t)f'(t) + f(t)g'(t).$$

The generalization is similar. Here, we need use the map

$$\mu : \mathbb{R}^{2n} \to \mathbb{R}, < \mathbf{v}, \mathbf{u} > \mapsto \mathbf{v} \cdot \mathbf{u}$$

instead of m, where we interpret a vector in $\mathbb{R}^{2n}$ as two vectors in $\mathbb{R}^n$ next to each other. We can then calculate

$$\frac{\partial \mu}{\partial u_i} = \frac{\partial}{\partial u_i} \sum_{k=1}^{n} u_k v_k = v_i, \quad \frac{\partial \mu}{\partial v_i} = \frac{\partial}{\partial u_i} \sum_{k=1}^{n} u_k v_k = u_i$$

so that

$$\mu'(< \mathbf{u}, \mathbf{v} >) = \begin{bmatrix} \frac{\partial \mu}{\partial u_1} & \dots \frac{\partial \mu}{\partial u_n} & \frac{\partial \mu}{\partial v_1} & \dots \frac{\partial \mu}{\partial v_n} \end{bmatrix} = \begin{bmatrix} \mathbf{v}^T & \mathbf{u}^T \end{bmatrix}$$

, so that (with Φ as above)

$$\begin{aligned}(f \cdot g)'(t) = (\mu \circ \Phi)'(t) = \mu'(\Phi(t))\Phi'(t) &= \begin{bmatrix} g'(t)^T & f'(t)^T \end{bmatrix} \begin{bmatrix} f(t) \\ g(t) \end{bmatrix} \\ &= g'(t)^T f(t) + f'(t)^T g(t) = g'(t) \cdot f(t) + f'(t) \cdot g(t).\end{aligned}$$

□

Exercise 3. As for the formula for R, we have

$$(c \circ \sigma \circ L)(\mathbf{a}, b) = c(\sigma(\mathbf{a} \cdot \mathbf{x} + b)) = (\sigma(\mathbf{a} \cdot \mathbf{x} + b) - y)^2.$$

c is not hard to differentiate, since it is a univariate function: $c'(t) = 2(t - y)$.

We move on to the formula for $L'(\mathbf{a}, b)$. The more mundane to write everything in coordinates and calculate the partial derivatives, as follows:

$$L(\mathbf{a}, b) = \mathbf{a} \cdot \mathbf{x} + b = \sum_{i=1}^{n} a_i x_i + b, \frac{\partial L}{\partial a_j} = x_j, \quad \frac{\partial L}{\partial b} = 1,$$

so that

$$L'(\mathbf{a}, b) = \begin{bmatrix} x_1 & \dots & x_n & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{x}^T & 1 \end{bmatrix}.$$

We can also work directly with the definition of the derivative:

$$L(\mathbf{a} + \mathbf{u}, b + v) = (\mathbf{a} + \mathbf{u}) \cdot \mathbf{x} + b + v = \mathbf{a} \cdot \mathbf{x} + b + \mathbf{x}^T \mathbf{u} + v = L(\mathbf{a}, b) + \begin{bmatrix} \mathbf{x}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ v \end{bmatrix} + 0.$$

This is exactly the equation used for defining the derivative, with $L'(\mathbf{a}, b) = \begin{bmatrix} \mathbf{x}^T & 1 \end{bmatrix}$ and $\epsilon(\mathbf{u}, v) = 0$.

□

Exercise 4. If we let the weight of a single peanut be w and the total weight of a bag be T, the number n of peanuts in the bag is given by

$$n(w, T) = \frac{T}{w}.$$

Let $(w_0, T_0) = (0.45, 500)$ be the ideal nut weight - bag weight. The numbers for an actual bag (w, T) can then be anywhere in $[0.4, 0.5] \times [490, 510]$. The Schrankensatz tells us

$$\begin{aligned}|n(w, T) - n(w_0, T_0)| \leq &\left(\max_{(\tilde{w}, \widetilde{T}) \in [0.4,0.5] \times [490,510]} \left| \frac{\partial n}{\partial w}(\tilde{w}, \widetilde{T}) \right| \right) |w - w_0| \\ &+ \left(\max_{(\tilde{w}, \widetilde{T}) \in [0.4,0.5] \times [490,510]} \left| \frac{\partial n}{\partial T}(\tilde{w}, \widetilde{T}) \right| \right) |T - T_0|\end{aligned}$$

We calculate the maxima of the partial derivatives:

$$\max_{(\tilde{w},\widetilde{T})\in[0.4,0.5]\times[490,510]} \left|\frac{\partial n}{\partial w}(\tilde{w},\widetilde{T})\right| = \max_{(\tilde{w},\widetilde{T})\in[0.4,0.5]\times[490,510]} \left|\frac{-T}{w^2}\right| = \frac{510}{0.4^2}$$
$$\max_{(\tilde{w},\widetilde{T})\in[0.4,0.5]\times[490,510]} \left|\frac{\partial n}{\partial T}(\tilde{w},\widetilde{T})\right| = \max_{(\tilde{w},\widetilde{T})\in[0.4,0.5]\times[490,510]} \left|\frac{1}{w}\right| = \frac{1}{0.4}$$

Thus,

$$|n(w,T) - n(w_0,T_0)| \le \frac{510}{0.4^2}\cdot 0.05 + \frac{1}{0.4}\cdot 10 \le 184.4$$

Considering that $n(w_0,T_0) = 1111\frac{1}{9}$, we get

$$n(w,T) = 1111\tfrac{1}{9} \pm 184.4$$

□