MATH20132 Calculus of Several Variables. 2020-21

Solutions to Problems 4 Fréchet derivatives

Fréchet differentiable ***scalar-valued*** *functions.*

1. a. Prove, by verifying the definition, that the scalar-valued functions

i. $f : \mathbb{R}^2 \to \mathbb{R}, \mathbf{x} \longmapsto x(x+y)$ and

ii. $g : \mathbb{R}^2 \to \mathbb{R}, \mathbf{x} \longmapsto y(x-y)$

are Fréchet differentiable at a general point $\mathbf{a} = (\alpha, \beta)^T$ and find the Fréchet derivatives $df_{\mathbf{a}}$ and $dg_{\mathbf{a}}$.

b. Use your results to check your answers to Question 1 on Sheet 3.

Solution a. i. With $\mathbf{t} = (s, t)^T$ we have

$$\mathbf{a} + \mathbf{t} = \begin{pmatrix} \alpha + s \\ \beta + t \end{pmatrix},$$

and so

$$\begin{aligned} f(\mathbf{a}+\mathbf{t}) - f(\mathbf{a}) &= (\alpha + s)(\alpha + s + \beta + t) - \alpha(\alpha + \beta) \\ &= \alpha s + \alpha t + s\alpha + s^2 + s\beta + st \\ &= (2\alpha + \beta)s + \alpha t + s^2 + st. \end{aligned}$$

We might guess that the required linear function of s and t is $L(\mathbf{t}) = (2\alpha + \beta)s + \alpha t$ (we have simply omitted any higher powers or products of s and t). To verify the definition first consider

$$f(\mathbf{a}+\mathbf{t}) - f(\mathbf{a}) - L(\mathbf{t}) = s^2 + st.$$

Recall that $|s|, |t| \le |\mathbf{t}|$ so, by the triangle inequality,

$$\left|s^2 + st\right| \le |s|^2 + |s|\,|t| \le 2\,|\mathbf{t}|^2.$$

Then

$$\left|\frac{f(\mathbf{a}+\mathbf{t}) - f(\mathbf{a}) - L(\mathbf{t})}{|\mathbf{t}|}\right| = \left|\frac{s^2 + st}{|\mathbf{t}|}\right| \le \frac{2\,|\mathbf{t}|^2}{|\mathbf{t}|} = 2\,|\mathbf{t}| \to 0$$

as $\mathbf{t} \to 0$. Hence, by the Sandwich Rule,

$$\lim_{\mathbf{t}\to 0} \frac{f(\mathbf{a}+\mathbf{t}) - f(\mathbf{a}) - L(\mathbf{t})}{|\mathbf{t}|} = 0. \quad (1)$$

Thus f is Fréchet differentiable at $\mathbf{a}$ with

$$df_{\mathbf{a}}(\mathbf{t}) = (2\alpha + \beta)\, s + \alpha t.$$

Yet $\mathbf{a}$ is arbitrary so f is Fréchet differentiable on $\mathbb{R}^2$.

For the second function,

$$\begin{aligned} g(\mathbf{a}+\mathbf{t}) - g(\mathbf{a}) &= (\beta + t)\,(\alpha - \beta + s - t) - \beta\,(\alpha - \beta) \\ &= \beta s - 2\beta t + \alpha t + st - t^2 \end{aligned}$$

We guess the linear form is $L(\mathbf{t}) = \beta s + (\alpha - 2\beta)\, t$. There are no new ideas involved in showing that the corresponding (1) holds. Thus g is Fréchet differentiable at $\mathbf{a}$ with

$$dg_{\mathbf{a}}(\mathbf{t}) = \beta s + (\alpha - 2\beta)\, t.$$

Again $\mathbf{a}$ is arbitrary so g is Fréchet differentiable on $\mathbb{R}^2$.

b. Since the functions are Fréchet differentiable we have

$$\begin{aligned} d_{\mathbf{v}} f(\mathbf{a}) &= df_{\mathbf{a}}(\mathbf{v}) = (2\alpha + \beta)\, u + \alpha v \\ d_{\mathbf{v}} g(\mathbf{a}) &= dg_{\mathbf{a}}(\mathbf{v}) = \beta u + (\alpha - 2\beta)\, v. \end{aligned}$$

if $\mathbf{v} = (u, v)^T$. With $\mathbf{a} = (1,2)^T$ and $\mathbf{v} = (2,-1)^T / \sqrt{5}$ we find

$$\begin{aligned} d_{\mathbf{v}} f(\mathbf{a}) &= 4 \times \frac{2}{\sqrt{5}} - \frac{1}{\sqrt{5}} = \frac{7}{\sqrt{5}}, \\ d_{\mathbf{v}} g(\mathbf{a}) &= 2 \times \frac{2}{\sqrt{5}} - 3 \times \left(-\frac{1}{\sqrt{5}}\right) = \frac{7}{\sqrt{5}}. \end{aligned}$$

Hopefully these agree with your answers to Question 1 on Sheet 3.

2. i. Define the function $h : \mathbb{R}^3 \to \mathbb{R}$ by $h(\mathbf{x}) = xy + yz + xz$ where $\mathbf{x} = (x, y, z)^T$. Prove by verifying the definition that f is Fréchet differentiable at a general point $\mathbf{a} = (\alpha, \beta, \gamma)^T \in \mathbb{R}^3$ and find the Fréchet derivative $dh_{\mathbf{a}}$ of h at $\mathbf{a}$.

ii. Use your result to check your answer to Question 3 on Sheet 3.

Solution With $\mathbf{a} = (\alpha, \beta, \gamma)^T \in \mathbb{R}^3$ given and $\mathbf{t} = (s, t, u)^T$ consider

$$\begin{aligned} h(\mathbf{a}+\mathbf{t}) - h(\mathbf{a}) &= (\alpha+s)(\beta+t) + (\beta+t)(\gamma+u) + (\alpha+s)(\gamma+u) \\ &\quad -\alpha\beta - \beta\gamma - \alpha\gamma \\ &= \alpha t + s\beta + st + \beta u + t\gamma + tu + \alpha u + s\gamma + su \\ &= (\beta+\gamma)\, s + (\alpha+\gamma)\, t + (\alpha+\beta)\, u + st + tu + su \end{aligned}$$

We might guess that the required linear function of s, t and u is

$$L(\mathbf{t}) = (\beta+\gamma)\, s + (\alpha+\gamma)\, t + (\alpha+\beta)\, u.$$

To check the definition first consider

$$h(\mathbf{a}+\mathbf{t}) - h(\mathbf{a}) - L(\mathbf{t}) = st + tu + su.$$

Recall that $|s|, |t|, |u| \le |\mathbf{t}|$ so, by the triangle inequality,

$$|st + tu + su| \le |s|\,|t| + |t|\,|u| + |s|\,|u| \le 3\,|\mathbf{t}|^2.$$

Thus

$$\left| \frac{h(\mathbf{a}+\mathbf{t}) - h(\mathbf{a}) - L(\mathbf{t})}{|\mathbf{t}|} \right| = \left| \frac{st+tu+su}{|\mathbf{t}|} \right| \le \frac{3\,|\mathbf{t}|^2}{|\mathbf{t}|} = 3\,|\mathbf{t}| \to 0$$

as $\mathbf{t} \to 0$. Hence, by the sandwich rule,

$$\lim_{\mathbf{t}\to 0} \frac{h(\mathbf{a}+\mathbf{t}) - h(\mathbf{a}) - L(\mathbf{t})}{|\mathbf{t}|} = 0.$$

Thus h is Fréchet differentiable at $\mathbf{a}$ with

$$dh_{\mathbf{a}}(\mathbf{t}) = (\beta+\gamma)\, s + (\alpha+\gamma)\, t + (\alpha+\beta)\, u. \tag{2}$$

Yet $\mathbf{a}$ is arbitrary so h is Fréchet differentiable on $\mathbb{R}^2$.

The lesson to be learnt is that there must be an easier way of verifying a function is Fréchet differentiable then finding the derivative.

ii. Question 3 on Sheet 3 asked for the directional derivative of $h(\mathbf{x}) = xy + yz + xz$ at $\mathbf{a} = (1, 2, 3)^T$ in the direction of the unit vector $\mathbf{v} = (3, 2, 1)^T / \sqrt{14}$. Now we know that h is Fréchet differentiable we know that

$$\begin{aligned} d_{\mathbf{v}} h(\mathbf{a}) &= dh_{\mathbf{a}}(\mathbf{v}) = \frac{(2+3)\times 3 + (1+3)\times 2 + (1+2)\times 1}{\sqrt{14}} \quad \text{by (2)} \\ &= \frac{26}{\sqrt{14}}. \end{aligned}$$

Hopefully the same result as found previously.

*Fréchet differentiable **vector-valued** functions.*

3. Let $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix},$$

for $\mathbf{x} = (x, y)^T \in \mathbb{R}^2$. Prove, by *verifying the definition* that $\mathbf{f}$ is everywhere Fréchet differentiable and find the Fréchet derivative of $\mathbf{f}$ at a general point $\mathbf{a} = (\alpha, \beta)^T$.

Solution With $\mathbf{t} = (s, t)^T$ consider

$$\begin{aligned} \mathbf{f}(\mathbf{a}+\mathbf{t}) - \mathbf{f}(\mathbf{a}) &= \begin{pmatrix} (\alpha+s)^2 - (\beta+t)^2 \\ 2(\alpha+s)(\beta+t) \end{pmatrix} - \begin{pmatrix} \alpha^2 - \beta^2 \\ 2\alpha\beta \end{pmatrix} \\ &= \begin{pmatrix} 2\alpha s + s^2 - 2\beta t - t^2 \\ 2\alpha t + 2\beta s + 2st \end{pmatrix}. \end{aligned}$$

Guess that the linear part in s and t is

$$\mathbf{L}(\mathbf{t}) = \begin{pmatrix} 2\alpha s - 2\beta t \\ 2\alpha t + 2\beta s \end{pmatrix}.$$

Consider

$$\frac{\mathbf{f}(\mathbf{x}+\mathbf{t}) - \mathbf{f}(\mathbf{x}) - \mathbf{L}(\mathbf{t})}{|\mathbf{t}|} = \frac{1}{|\mathbf{t}|} \begin{pmatrix} s^2 - t^2 \\ 2st \end{pmatrix}.$$

Recall that in general, $\lim_{\mathbf{x}\to\mathbf{0}} \mathbf{f}(\mathbf{x}) = \mathbf{b}$ if, and only if $\lim_{\mathbf{x}\to\mathbf{0}} f^i(\mathbf{x}) = b^i$ for all components. So it suffices to prove

$$\lim_{\mathbf{t}\to\mathbf{0}} \frac{s^2 - t^2}{|\mathbf{t}|} = 0 \quad \text{and} \quad \lim_{\mathbf{t}\to\mathbf{0}} \frac{2st}{|\mathbf{t}|} = 0. \tag{3}$$

Yet both will follow from the Sandwich Rule with the observation that $|s|, |t| \leq |\mathbf{t}|$. So, start with the triangle inequality,

$$\left|s^2 - t^2\right| \leq |s|^2 + |t|^2 = |\mathbf{t}|^2 \quad \text{along with} \quad |2st| \leq 2|\mathbf{t}|^2.$$

Then

$$\frac{|s^2-t^2|}{|\mathbf{t}|} \le \frac{|s^2|+|t^2|}{|\mathbf{t}|} = |\mathbf{t}| \to 0 \quad \text{and} \quad \frac{|2st|}{|\mathbf{t}|} \le 2\,|\mathbf{t}| \to 0$$

as $\mathbf{t} \to \mathbf{0}$. Hence the limits in (3) follow, f is Fréchet differentiable at $\mathbf{a}$ and thus everywhere with

$$\mathbf{df_a}(\mathbf{t}) = \begin{pmatrix} 2\alpha s - 2\beta t \\ 2\alpha t + 2\beta s \end{pmatrix}.$$

4. i. Prove that the scalar-valued function $\mathbf{x} \mapsto x^2y$ is everywhere Fréchet differentiable on $\mathbb{R}^2$. In lectures and Problems class this was done for the scalar-valued function $\mathbf{x} \mapsto xy^2$, simply copy that method for the second.

ii. Prove that the vector-valued function $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} xy^2 \\ x^2y \end{pmatrix},$$

is Fréchet differentiable at a general point $\mathbf{a} = (\alpha, \beta)^T \in \mathbb{R}^2$ and find the Fréchet derivative $d\mathbf{f_a}$ of $\mathbf{f}$ at $\mathbf{a}$.

ii. Use the result from part i to check your answer to Question 7 on Sheet 3.

Note the difference in wording between this question and the previous one.

Solution Let $\mathbf{a} = (\alpha, \beta)^T$, $\mathbf{t} = (s, t)^T \in \mathbb{R}^2$. Let $f(\mathbf{x}) = x^2y$. Then

$$\begin{aligned} f(\mathbf{a}+\mathbf{t}) - f(\mathbf{a}) &= f\left(\begin{pmatrix} \alpha+s \\ \beta+t \end{pmatrix}\right) - f\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}\right) \\ &= (\alpha+s)^2(\beta+t) - \alpha^2\beta \\ &= \left(\alpha^2 + 2\alpha s + s^2\right)(\beta+t) - \alpha^2\beta \\ &= 2\alpha\beta s + \alpha^2 t + \beta s^2 + 2\alpha st + s^2 t. \end{aligned}$$

The 'linear part in s and t' of this is $2\alpha\beta s + \alpha^2 t$, so we *guess*

$$df_{\mathbf{a}}(\mathbf{t}) = 2\alpha\beta s + \alpha^2 t.$$

To *check*, consider

$$\frac{f(\mathbf{a}+\mathbf{t}) - f(\mathbf{a}) - (2\alpha\beta s + \alpha^2 t)}{|\mathbf{t}|} = \frac{\beta s^2 + 2\alpha st + s^2 t}{|\mathbf{t}|}.$$

Again, $|s|$, $|t| \leq |\mathbf{t}|$ and $|\alpha|$, $|\beta| \leq |\mathbf{a}|$, so by the triangle inequality

$$\begin{aligned} \left|\beta s^2 + 2\alpha st + s^2 t\right| &\leq |\beta| |s|^2 + 2|\alpha| |s| |t| + |s|^2 |t| \\ &\leq 3|\mathbf{a}| |\mathbf{t}|^2 + |\mathbf{t}|^3 . \end{aligned}$$

Therefore

$$\left|\frac{\beta s^2 + 2\alpha st + s^2 t}{|\mathbf{t}|}\right| \leq 3|\mathbf{a}| |\mathbf{t}| + |\mathbf{t}|^2 \to 0$$

as $\mathbf{t} \to \mathbf{0}$. Hence f is Fréchet differentiable on $\mathbb{R}^2$ with

$$df_{\mathbf{a}}(\mathbf{t}) = 2\alpha\beta s + \alpha^2 t.$$

ii. By a proposition in the notes, $\mathbf{f}$ is Fréchet differentiable at $\mathbf{a}$ iff each component function is Fréchet differentiable at $\mathbf{a}$, and further $(d\mathbf{f}_{\mathbf{a}})^i = df^i_{\mathbf{a}}$ for all i. In this example $f^1(\mathbf{x}) = xy^2$ and from the lectures we know that $df^1_{\mathbf{a}}(\mathbf{t}) = \beta^2 s + 2\alpha\beta t$. Hence

$$d\mathbf{f}_{\mathbf{a}}(\mathbf{t}) = \begin{pmatrix} df^1_{\mathbf{a}}(\mathbf{t}) \\ df^2_{\mathbf{a}}(\mathbf{t}) \end{pmatrix} = \begin{pmatrix} \beta^2 s + 2\alpha\beta t \\ 2\alpha\beta s + \alpha^2 t \end{pmatrix}. \tag{4}$$

iii. Question 7 on Sheet 3 asks for the directional derivative of $\mathbf{f}$ at $\mathbf{a} = (2,1)^T$ in the direction of the unit vector $\mathbf{v} = (1,-1)^T/\sqrt{2}$. Now we know that $\mathbf{f}$ is Fréchet differentiable we know that

$$\begin{aligned} d_{\mathbf{v}}\mathbf{f}(\mathbf{a}) &= d\mathbf{f}_{\mathbf{a}}(\mathbf{v}) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1^2 \times 1 - 2 \times 2 \times 1 \times 1 \\ 2 \times 2 \times 1 \times 1 - 2^2 \times 1 \end{pmatrix} \quad \text{by (4)} \\ &= -\frac{3}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

Jacobian Matrices and Gradient vectors

5. Write down the general Jacobian matrix in each of the following cases and then evaluate them at the given point.

i. $\mathbf{p}(r,\theta) = (r\cos\theta, r\sin\theta)^T$ at the point $(1,\pi)^T$,

ii. $g(u,v,w) = uv + 5u^2 w$ at the point $(2,-3,1)^T$,

iii. $\mathbf{r}(t) = (\cos t,\, \sin t,\, t)^T$, a helix in $\mathbb{R}^3$, at the point $t = 3\pi$.

Solution i.

$$J\mathbf{p}(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \quad \text{and} \quad J\mathbf{p}(1,\pi) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

ii.

$$Jg(u,v,w) = \begin{pmatrix} v + 10uw & u & 5u^2 \end{pmatrix} \quad \text{and} \quad Jg(2,-3,1) = \begin{pmatrix} 17 & 2 & 20 \end{pmatrix}.$$

iii.

$$J\mathbf{r}(t) = \begin{pmatrix} -\sin t \\ \cos t \\ 1 \end{pmatrix} \quad \text{and} \quad J\mathbf{r}(3\pi) = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

6. By returning to Questions 1 and 2 find the gradient vectors of

i. $f : \mathbb{R}^2 \to \mathbb{R}, \mathbf{x} \longmapsto x(x+y)$ and

ii. $g : \mathbb{R}^2 \to \mathbb{R}, \mathbf{x} \longmapsto y(x-y)$

iii. $h : \mathbb{R}^3 \to \mathbb{R}$ by $\mathbf{x} \longmapsto xy + yz + xz$ where $\mathbf{x} = (x,y,z)^T$,

without using partial differentiation. Justify your argument.

Solution i. From Question 1 we have that f is everywhere Fréchet differentiable which is necessary for us to say that, for all $\mathbf{a} = (\alpha,\beta)^T \in \mathbb{R}^2$, we have

$$\begin{aligned} \nabla f(\mathbf{a}) \bullet \mathbf{t} &= df_{\mathbf{a}}(\mathbf{t}) \\ &= (2\alpha+\beta)\, s + \alpha t \quad \text{by Question 1} \\ &= (2\alpha+\beta, \alpha)^T \bullet \mathbf{t}. \end{aligned}$$

True for all $\mathbf{t} \in \mathbb{R}^2$ means $\nabla f(\mathbf{a}) = (2\alpha+\beta,\alpha)^T$, i.e. $\nabla f(\mathbf{x}) = (2x+y, x)^T$ for $\mathbf{x} = (x,y)^T$.

ii. From Question 1 we have that g is everywhere Fréchet differentiable which is necessary for us to say that, for all $\mathbf{a} = (\alpha,\beta)^T \in \mathbb{R}^2$, we have

$$\begin{aligned} \nabla g(\mathbf{a}) \bullet \mathbf{t} &= dg_{\mathbf{a}}(\mathbf{t}) \\ &= \beta s + (\alpha - 2\beta)\, t \quad \text{by Question 1} \\ &= (\beta, \alpha - 2\beta)^T \bullet \mathbf{t}. \end{aligned}$$

True for all $\mathbf{t} \in \mathbb{R}^2$ means $\nabla g(\mathbf{a}) = (\beta, \alpha - 2\beta)^T$, i.e. $\nabla g(\mathbf{x}) = (y, x - 2y)^T$ for $\mathbf{x} = (x, y)^T$.

iii. From Question 2 we have that h is everywhere Fréchet differentiable which is necessary for us to say that, for all $\mathbf{a} = (\alpha, \beta, \gamma)^T \in \mathbb{R}^3$, we have

$$\begin{aligned} \nabla h(\mathbf{a}) \bullet \mathbf{t} &= dh_{\mathbf{a}}(\mathbf{t}) \\ &= (\beta + \gamma)\, s + (\alpha + \gamma)\, t + (\alpha + \beta)\, u \quad \text{by Question 2} \\ &= (\beta + \gamma, \alpha + \gamma, \alpha + \beta)^T \bullet \mathbf{t}. \end{aligned}$$

True for all $\mathbf{t} \in \mathbb{R}^3$ means $\nabla h(\mathbf{a}) = (\beta + \gamma, \alpha + \gamma, \alpha + \beta)^T$, i.e. $\nabla h(\mathbf{x}) = (y + z, x + z, x + y)^T$ for $\mathbf{x} = (x, y, z)^T$.

7. By returning to Question 3 and 4 find the Jacobian matrices of $\mathbf{f}, \mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}, \text{ and } \mathbf{g}(\mathbf{x}) = \begin{pmatrix} xy^2 \\ x^2 y \end{pmatrix}$$

without using partial differentiation. Justify your argument.

Solution From Question 3 we have that $\mathbf{f}$ is everywhere Fréchet differentiable which is the required justification for us to say that, for all $\mathbf{a} = (\alpha, \beta)^T \in \mathbb{R}^2$, we have

$$\begin{aligned} J\mathbf{f}(\mathbf{a})\, \mathbf{t} &= d\mathbf{f}_{\mathbf{a}}(\mathbf{t}) = \begin{pmatrix} 2\alpha s - 2\beta t \\ 2\alpha t + 2\beta s \end{pmatrix} \quad \text{by Question 3} \\ &= \begin{pmatrix} 2\alpha & -2\beta \\ 2\beta & 2\alpha \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 2\alpha & -2\beta \\ 2\beta & 2\alpha \end{pmatrix} \mathbf{t}. \end{aligned}$$

True for all $\mathbf{t} \in \mathbb{R}^2$ means

$$J\mathbf{f}(\mathbf{a}) = \begin{pmatrix} 2\alpha & -2\beta \\ 2\beta & 2\alpha \end{pmatrix}, \text{ i.e. } J\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$$

for $\mathbf{x} = (x, y)^T$.

From Question 4 we have that $\mathbf{g}$ is everywhere Fréchet differentiable and so, for all $\mathbf{a} = (\alpha, \beta)^T \in \mathbb{R}^2$, we have

$$J\mathbf{g}(\mathbf{a})\, \mathbf{t} = d\mathbf{g}_{\mathbf{a}}(\mathbf{t}) = \begin{pmatrix} \beta^2 s + 2\alpha\beta t \\ 2\alpha\beta s + \alpha^2 t \end{pmatrix} = \begin{pmatrix} \beta^2 & 2\alpha\beta \\ 2\alpha\beta & \alpha^2 \end{pmatrix} \mathbf{t}.$$

True for all $\mathbf{t} \in \mathbb{R}^2$ means

$$J\mathbf{g}(\mathbf{a}) = \begin{pmatrix} \beta^2 & 2\alpha\beta \\ 2\alpha\beta & \alpha^2 \end{pmatrix}, \text{ i.e. } J\mathbf{g}(\mathbf{x}) = \begin{pmatrix} 2y^2 & 2xy \\ 2xy & 2x^2 \end{pmatrix}$$

for $\mathbf{x} = (x, y)^T$.

Not *Fréchet differentiable*

8. *Recall the important result for* ***scalar-valued*** *functions*

$$\boxed{f \text{ Fréchet differentiable at } \mathbf{a} \implies \forall \text{ unit } \mathbf{v}, d_{\mathbf{v}}f(\mathbf{a}) \text{ exists and } d_{\mathbf{v}}f(\mathbf{a}) = \nabla f(\mathbf{a}) \bullet \mathbf{v}.}$$

The contrapositive of this is the useful

$$\boxed{\begin{array}{c} \exists \text{ unit } \mathbf{v} : \text{either } d_{\mathbf{v}}f(\mathbf{a}) \text{ does not exist or } d_{\mathbf{v}}f(\mathbf{a}) \neq \nabla f(\mathbf{a}) \bullet \mathbf{v} \\ \implies f \text{ is } \textbf{not} \text{ Fréchet differentiable at } \mathbf{a} \end{array}} \qquad (5)$$

Define the function $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(\mathbf{x}) = \frac{x^2 y}{x^2 + y^2} \quad \text{if } \mathbf{x} \neq 0, \qquad f(\mathbf{0}) = 0.$$

Prove that f is **not** Fréchet differentiable at $\mathbf{0}$.

This function was seen in Question 11iii Sheet 1 where it was shown that $\lim_{\mathbf{x}\to\mathbf{0}} f(\mathbf{x}) = 0$. *Since* $f(\mathbf{0}) = 0$ *this means* f *is continuous at* $\mathbf{0}$. *So we have an example illustrating the important*

$$\boxed{f \text{ continuous at } \mathbf{a} \not\Longrightarrow f \text{ is Fréchet differentiable at } \mathbf{a}.}$$

But f *was also seen in Question 14 on Sheet 3 where it was shown that the directional derivatives* $d_{\mathbf{v}}f(\mathbf{0})$ *exist for all unit* $\mathbf{v}$. *So we have an example of*

$$\boxed{\forall \text{ unit } \mathbf{v}, d_{\mathbf{v}}f(\mathbf{a}) \text{ exists } \not\Longrightarrow f \text{ is Fréchet differentiable at } \mathbf{a}.}$$

Solution By (5) it suffices to find a $\mathbf{v} : d_{\mathbf{v}}f(\mathbf{0}) \neq \nabla f(\mathbf{0}) \bullet \mathbf{v}$.

From Question 14ii on Sheet 3 we have

$$\frac{\partial f}{\partial x}(\mathbf{0}) = \frac{\partial f}{\partial y}(\mathbf{0}) = 0$$

and so $\nabla f(\mathbf{0}) = \mathbf{0}$ and thus $\nabla f(\mathbf{0}) \bullet \mathbf{v} = 0$ for all $\mathbf{v}$.

Yet from Question 14iii on Sheet 3 we have $d_{\mathbf{v}} f(\mathbf{0}) = f(\mathbf{v})$ so we need only choose a $\mathbf{v} : f(\mathbf{v}) \neq 0$. For example, $\mathbf{v} = (1,1)^T / \sqrt{2}$.

In Question 8 we used that fact that there exists a unit $\mathbf{v}$ for which $d_{\mathbf{v}} f(\mathbf{0}) \neq \nabla f(\mathbf{0}) \bullet \mathbf{v}$ to deduce that f is not Fréchet differentiable at $\mathbf{0}$. In the following question we find an example of a function f for which $d_{\mathbf{v}} f(\mathbf{0}) = \nabla f(\mathbf{0}) \bullet \mathbf{v}$ for all unit $\mathbf{v}$ and yet f is still not Fréchet differentiable at $\mathbf{0}$.

9. (Tricky) Define the function $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(\mathbf{x}) = \frac{xy^2\sqrt{x^2+y^2}}{x^2+y^4} \quad \text{if } \mathbf{x} \neq \mathbf{0}; \quad f(\mathbf{0}) = 0.$$

i. Prove that f is continuous at $\mathbf{0}$.

   Hint Note that $x \leq \sqrt{x^2+y^4}$, and similarly for y^2.

ii. Prove from first principles that the directional derivative exists in all directions, and further, satisfies $d_{\mathbf{v}} f(\mathbf{0}) = \nabla f(\mathbf{0}) \bullet \mathbf{v}$ for all unit vectors $\mathbf{v} \in \mathbb{R}^2$.

iii. Prove that, nevertheless, f is **not** Fréchet differentiable at $\mathbf{0}$.

This example illustrates two important points

$$\boxed{\text{continuous} \;\not\Longrightarrow\; \text{Fréchet differentiable.}}$$

and

$$\boxed{\forall \text{ unit } \mathbf{v}, d_{\mathbf{v}} f(\mathbf{a}) \text{ exists and } d_{\mathbf{v}} f(\mathbf{a}) = \nabla f(\mathbf{a}) \bullet \mathbf{v} \;\not\Longrightarrow\; \text{differentiable.}}$$

Solution i. To show f is continuous at $\mathbf{0}$ we start by bounding $|f(\mathbf{x}) - f(\mathbf{0})| = |f(\mathbf{x})|$ from above. By the hint given

$$|x| \leq \sqrt{x^2+y^4} \quad \text{and} \quad \left|y^2\right| \leq \sqrt{x^2+y^4}.$$

(This is probably the hardest step.) Thus

$$|f(\mathbf{x})| \leq \sqrt{x^2+y^2} = |\mathbf{x}| \to 0$$

as $\mathbf{x} \to 0$. Hence, by the Sandwich Rule, $\lim_{\mathbf{x}\to\mathbf{0}} f(\mathbf{x}) = 0 = f(\mathbf{0})$, therefore f is continuous at $\mathbf{0}$.

Alternative proof Use the so-called AM-GM inequality; that the arithmetic mean is greater than the geometric mean. So if $a, b > 0$ then $\sqrt{ab} \leq (a+b)/2$ (follows from $\left(\sqrt{a} - \sqrt{b}\right)^2 \geq 0$). Apply this with $a = x^2$ and $b = y^4$ to get $|x| y^2 < (x^2 + y^4)/2$. This gives the stronger result $|f(\mathbf{x})| < |\mathbf{x}|/2$.

ii. To find the directional derivatives at $\mathbf{0}$ assume that the unit vector $\mathbf{v}$ is given. Let $\mathbf{v} = (h, k)^T$, so $\sqrt{h^2 + k^2} = 1$. Then we have two cases.

The first case is $h \neq 0$, when

$$f(\mathbf{0} + t\mathbf{v}) = \frac{th\,(tk)^2}{(th)^2 + (tk)^4}\sqrt{(th)^2 + (tk)^2} = \frac{t^2 hk^2}{h^2 + t^2 k^4}.$$

Thus

$$\frac{f(\mathbf{0} + t\mathbf{v}) - f(\mathbf{0})}{t} = t\frac{hk^2}{h^2 + t^2 k^4} \to 0 \tag{6}$$

as $t \to 0$.

In the second case $h = 0$ when $f(\mathbf{0} + t\mathbf{v}) = 0$ for all t and we get the same limit in (6).

Therefore in all cases the limit in (6) exists, so f has directional derivatives at $\mathbf{0}$ with $d_{\mathbf{v}} f(\mathbf{0}) = 0$ for all unit vectors $\mathbf{v}$.

Note for the next part that (6) means, in particular, that $d_1 f(\mathbf{0}) = d_{\mathbf{e}_1} f(\mathbf{0}) = 0$ and $d_2 f(\mathbf{0}) = d_{\mathbf{e}_2} f(\mathbf{0}) = 0$. Thus the gradient vector at $\mathbf{0}$ is

$$\nabla f(\mathbf{0}) = \begin{pmatrix} d_1 f(\mathbf{0}) \\ d_2 f(\mathbf{0}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{0}.$$

iii. To prove that f is **not** Fréchet differentiable at $\mathbf{0}$ we assume, for a contradiction, that f **is** Fréchet differentiable at $\mathbf{0}$. This means that the linear function $df_{\mathbf{0}}$ exists. But further, by a result in the notes $df_{\mathbf{0}}(\mathbf{v}) = \nabla f(\mathbf{0}) \bullet \mathbf{v}$ for all vectors $\mathbf{v}$. Yet we have $\nabla f(\mathbf{0}) = \mathbf{0}$ which thus means that $df_{\mathbf{0}}(\mathbf{v}) = 0$ for all $\mathbf{v} \in \mathbb{R}^2$. Hence $df_{\mathbf{0}} = 0 : \mathbb{R}^2 \to \mathbb{R}$ (i.e. it is the map that takes all vectors of $\mathbb{R}^2$ to 0.)

Now look at the definition of $df_{\mathbf{0}}$ as the linear map which satisfies

$$0 = \lim_{\mathbf{t}\to\mathbf{0}} \frac{f(\mathbf{0} + \mathbf{t}) - f(\mathbf{0}) - df_{\mathbf{0}}(\mathbf{t})}{|\mathbf{t}|} = \lim_{\mathbf{t}\to\mathbf{0}} \frac{f(\mathbf{t})}{|\mathbf{t}|} \tag{7}$$

since $f(\mathbf{0}) = 0$ and $df_{\mathbf{0}}(\mathbf{t}) = 0$. The definition of f may have looked complicated but it could have been written as

$$f(\mathbf{x}) = \frac{xy^2\,|\mathbf{x}|}{x^2+y^4}, \quad \text{i.e.} \quad \frac{f(\mathbf{x})}{|\mathbf{x}|} = \frac{xy^2}{x^2+y^4}.$$

Then (7) is saying that

$$0 = \lim_{\mathbf{t}\to\mathbf{0}} \frac{st^2}{s^2+t^4}, \tag{8}$$

where $\mathbf{t} = (s,t)^T$. Yet Question 11iv, Sheet 1, showed this is false. As a recap of that earlier question, if (8) were true we would get the same value, 0, along whatever path we approached the origin. If we choose the path

$$\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} \ell^2 \\ \ell \end{pmatrix}$$

then

$$0 = \lim_{\mathbf{t}\to\mathbf{0}} \frac{st^2}{s^2+t^4} = \lim_{\ell\to 0} \frac{\ell^4}{\ell^4+\ell^4} = \frac{1}{2}.$$

Contradiction.

Aside You can see how this example was constructed. I took a bounded function not continuous at $\mathbf{0}$, called $g(\mathbf{x})$ say. I then defined $f(\mathbf{x}) = g(\mathbf{x})\,|\mathbf{x}|$. Note that $f(\mathbf{0}) = 0$ and $f(\mathbf{0}+t\mathbf{v}) = g(t\mathbf{v})\,|t|$ so $(f(\mathbf{0}+t\mathbf{v}) - f(\mathbf{0}))\,/t = g(t\mathbf{v})\,|t|\,/t$. Thus I also demand that $\lim_{t\to 0} g(t\mathbf{v}) = 0$ for all $\mathbf{v}$. So I had to look through our collection of functions that had the same directional limit at $\mathbf{0}$ in all directions but with different limits along two different paths to $\mathbf{0}$.

Solutions to Additional Questions 4

10. *Product Rule for Gradient vectors* Assume for $f, g : U \to \mathbb{R}$, $U \subseteq \mathbb{R}^n$ that the gradient vectors $\nabla f(\mathbf{a})$ and $\nabla g(\mathbf{a})$ exist at $\mathbf{a} \in U$. Prove that $\nabla (fg)(\mathbf{a})$ exists and satisfies

$$\nabla (fg)(\mathbf{a}) = f(\mathbf{a}) \nabla g(\mathbf{a}) + g(\mathbf{a}) \nabla f(\mathbf{a}).$$

Solution $\nabla f(\mathbf{a})$ and $\nabla g(\mathbf{a})$ exist iff $d_i f(\mathbf{a})$ and $d_i g(\mathbf{a})$ exist for all $1 \leq i \leq n$, iff $d_{\mathbf{e}_i} f(\mathbf{a})$ and $d_{\mathbf{e}_i} g(\mathbf{a})$ exist for all $1 \leq i \leq n$. This is sufficient, by Question 16 on Sheet 3, to deduce that $d_{\mathbf{e}_i}(fg)(\mathbf{a})$ exists and

$$d_i (fg)(\mathbf{a}) = f(\mathbf{a}) d_i g(\mathbf{a}) + g(\mathbf{a}) d_i f(\mathbf{a}),$$

for all $1 \leq i \leq n$. Yet this is simply the equality of coordinates in $\nabla (fg)(\mathbf{a}) = f(\mathbf{a}) \nabla g(\mathbf{a}) + g(\mathbf{a}) \nabla f(\mathbf{a})$.

11. Prove that

$$f(\mathbf{x}) = \frac{xy}{\sqrt{x^2+y^2}}, \ \mathbf{x} = (x,y)^T \neq \mathbf{0}, \quad f(\mathbf{0}) = 0,$$

is continuous but not Fréchet differentiable at $\mathbf{0}$.

Solution Start by noting that

$$f(\mathbf{x}) = \frac{xy}{|\mathbf{x}|} \quad \text{so} \quad |f(\mathbf{x})| = \frac{|x|\,|y|}{|\mathbf{x}|} \leq \frac{|\mathbf{x}|\,|\mathbf{x}|}{|\mathbf{x}|} = |\mathbf{x}| \to 0$$

as $\mathbf{x} \to \mathbf{0}$. Thus, by the Sandwich Rule, $\lim_{\mathbf{x}\to\mathbf{0}} f(\mathbf{x}) = 0 = f(\mathbf{0})$ and so f is continuous at $\mathbf{0}$.

Assume for contradiction that f is Fréchet differentiable at $\mathbf{0}$. Then $df_{\mathbf{0}}$ exists and satisfies $df_{\mathbf{0}}(\mathbf{v}) = d_{\mathbf{v}} f(\mathbf{0})$ for all unit $\mathbf{v}$. It is not hard to check from the definition that $d_{\mathbf{v}} f(\mathbf{0}) = 0$ for all unit $\mathbf{v}$. Thus $df_{\mathbf{0}} = 0$. Within the definition of Fréchet differentiable this gives

$$0 = \lim_{\mathbf{t}\to\mathbf{0}} \frac{f(\mathbf{t}) - f(\mathbf{0}) - df_{\mathbf{0}}(\mathbf{t})}{|\mathbf{t}|} = \lim_{\mathbf{t}\to\mathbf{0}} \frac{f(\mathbf{t})}{|\mathbf{t}|} = \lim_{\mathbf{t}\to\mathbf{0}} \frac{st}{s^2+t^2},$$

where $t = (s,t)^T$. This is the required contradiction since the limit on the right does not exist (Question 11ii, Sheet 1.)