Differentiability: a summary

October 24[th], 2019

Real-valued case – $U \subset \mathbb{R}^n$ open and $f : U \to \mathbb{R}$.

Name	Nature	Notation and definition
Directional derivative at $\mathbf{x} \in U$ along $\mathbf{v} \in \mathbb{R}^n$	Real number	$\partial_{\mathbf{v}} f(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{v}) - f(\mathbf{x})}{t}$
i-th partial derivative at $\mathbf{x} \in U$	Real number	$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \partial_{\mathbf{e}_i} f(\mathbf{x})$
Gradient at $\mathbf{x} \in U$	Vector in $\mathbb{R}^n$	$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right)$
Differential at $\mathbf{x} \in U$ *“f is differentiable at $\mathbf{x}$”*	Linear function $d_{\mathbf{x}} f : \mathbb{R}^n \to \mathbb{R}$	$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + d_{\mathbf{x}} f(\mathbf{h}) + E(\mathbf{h})$ with $\lim_{\mathbf{h} \to \mathbf{0}} \frac{E(\mathbf{h})}{\Vert\mathbf{h}\Vert} = 0$

See the slides from Oct 10 for the geometric intuitions about these objects.

Real-valued case – $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ open, $\mathbf{x} \in U$

Partial derivatives exist on U and are continuous at $\mathbf{x}$ $\Longrightarrow$ Differentiable at $\mathbf{x}$ (and Differentiable at $\mathbf{x}$ $\not\Longrightarrow$ Partial derivatives exist on U and are continuous at $\mathbf{x}$)

Differentiable at $\mathbf{x}$ $\Longrightarrow$ Continuous at $\mathbf{x}$ (and Continuous at $\mathbf{x}$ $\not\Longrightarrow$ Differentiable at $\mathbf{x}$)

Differentiable at $\mathbf{x}$ $\Longrightarrow$

Interpretation of $\nabla f(\mathbf{x})$:
direction: dir. of fastest increase through $\mathbf{x}$
magnitude: steepness, instantaneous rate of change in that direction
Moreover: it is orthogonal to the level sets

Differentiable at $\mathbf{x}$ $\Longrightarrow$

Directional derivatives at $\mathbf{x}$ exist
and moreover

- $\partial_{\mathbf{v}} f(\mathbf{x}) = d_{\mathbf{x}} f(\mathbf{v})$
- $d_{\mathbf{x}} f(\mathbf{h}) = \nabla f(\mathbf{x}) \cdot \mathbf{h}$
- $\partial_{\mathbf{v}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{v}$

(and Directional derivatives at $\mathbf{x}$ exist $\not\Longrightarrow$ Differentiable at $\mathbf{x}$)

See the slides from Oct 10 for more details and for the counter-examples.

Partial derivatives at $\mathbf{x}$ exist $\not\Longrightarrow$ Directional derivatives at $\mathbf{x}$ exist

All the directional derivatives at $\mathbf{x}$ exist $\not\Longrightarrow$ Continuous at $\mathbf{x}$

Vector-valued case – $U \subset \mathbb{R}^n$ open and $\mathbf{f} : U \to \mathbb{R}^k$.

We denote by f_i the components of $\mathbf{f}$, i.e. $\mathbf{f} = (f_1, \dots, f_k) : U \to \mathbb{R}^k$

Name	Nature	Notation and definition
Differential (or total derivative) at $\mathbf{x} \in U$ "$\mathbf{f}$ *is differentiable at* $\mathbf{x}$"	Linear function $d_{\mathbf{x}}\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^k$	$\mathbf{f}(\mathbf{x}+\mathbf{h}) = \mathbf{f}(\mathbf{x}) + d_{\mathbf{x}}\mathbf{f}(\mathbf{h}) + \mathbf{E}(\mathbf{h})$ with $\lim_{\mathbf{h}\to\mathbf{0}} \frac{\mathbf{E}(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0}$
Jacobian matrix of $\mathbf{f}$ at $\mathbf{x} \in U$	$(k \times n)$-matrix	$D\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_k}{\partial x_n}(\mathbf{x}) \end{pmatrix}$

Vector-valued case – $\mathbf{f} : U \to \mathbb{R}^k$, $U \subset \mathbb{R}^n$ open, $\mathbf{x} \in U$

We denote by $f_i : U \to \mathbb{R}$ the components of $\mathbf{f}$, i.e. $\mathbf{f} = (f_1, \ldots, f_k) : U \to \mathbb{R}^k$

We proved that $\mathbf{f}$ is differentiable at $\mathbf{x}$ if and only if its components f_i are too.
It allowed us to use the results from the real-valued case to prove the following theorems:

The partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist on U and are continuous at $\mathbf{x}$

$\Longrightarrow$ $\mathbf{f}$ is differentiable at $\mathbf{x}$

$\mathbf{f}$ is differentiable at $\mathbf{x}$ $\Longrightarrow$ $\mathbf{f}$ is continuous at $\mathbf{x}$

$\mathbf{f}$ is differentiable at $\mathbf{x}$ $\Longrightarrow$ All the directional derivatives $\partial_{\mathbf{v}} f_i(\mathbf{x})$ of the components f_i at $\mathbf{x}$ exist and moreover

$$\operatorname{Mat}\left(d_{\mathbf{x}}\mathbf{f}\right) = D\mathbf{f}(\mathbf{x})$$

The Chain Rule

Let $U \subset \mathbb{R}^n$ open, $\mathbf{f} : \begin{array}{ccc} U & \to & \mathbb{R}^l \\ \mathbf{x} & \mapsto & \mathbf{f}(\mathbf{x}) \end{array}$, $V \subset \mathbb{R}^l$ open, $\mathbf{g} : \begin{array}{ccc} V & \to & \mathbb{R}^k \\ \mathbf{y} & \mapsto & \mathbf{g}(\mathbf{y}) \end{array}$.
Assume that $\mathbf{f}(U) \subset V$ so that $\mathbf{g} \circ \mathbf{f} : U \to \mathbb{R}^k$ is well-defined.

Let $\mathbf{x} \in U$.
If $\mathbf{f}$ is differentiable at $\mathbf{x}$ and $\mathbf{g}$ is differentiable at $\mathbf{f}(\mathbf{x})$ then $\mathbf{g} \circ \mathbf{f}$ is differentiable at $\mathbf{x}$.

- Chain rule formula for the differentials:

$$d_{\mathbf{x}}(\mathbf{g} \circ \mathbf{f}) = \left(d_{\mathbf{f}(\mathbf{x})}\mathbf{g}\right) \circ \left(d_{\mathbf{x}}\mathbf{f}\right)$$

- Chain rule formula for the Jacobian matrices:

$$D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = D(\mathbf{g})(\mathbf{f}(\mathbf{x})) \cdot D(\mathbf{f})(\mathbf{x})$$

- Chain rule for the partial derivatives:

$$\frac{\partial(g_i \circ \mathbf{f})}{\partial x_j}(\mathbf{x}) = \sum_{\alpha=1}^{l} \frac{\partial g_i}{\partial y_\alpha}(\mathbf{f}(\mathbf{x})) \cdot \frac{\partial f_\alpha}{\partial x_j}(\mathbf{x})$$

The Chain Rule

Let $U \subset \mathbb{R}^n$ open, $\mathbf{f} : \begin{array}{ccc} U & \to & \mathbb{R}^l \\ \mathbf{x} & \mapsto & \mathbf{f}(\mathbf{x}) \end{array}$, $V \subset \mathbb{R}^l$ open, $\mathbf{g} : \begin{array}{ccc} V & \to & \mathbb{R}^k \\ \mathbf{y} & \mapsto & \mathbf{g}(\mathbf{y}) \end{array}$.
Assume that $\mathbf{f}(U) \subset V$ so that $\mathbf{g} \circ \mathbf{f} : U \to \mathbb{R}^k$ is well-defined.

Let $\mathbf{x} \in U$.
If $\mathbf{f}$ is differentiable at $\mathbf{x}$ and $\mathbf{g}$ is differentiable at $\mathbf{f}(\mathbf{x})$ then $\mathbf{g} \circ \mathbf{f}$ is differentiable at $\mathbf{x}$.

- Chain rule formula for the differentials:

$$d_{\mathbf{x}}(\mathbf{g} \circ \mathbf{f}) = \left(d_{\mathbf{f}(\mathbf{x})}\mathbf{g}\right) \circ \left(d_{\mathbf{x}}\mathbf{f}\right)$$

- Chain rule formula for the Jacobian matrices:

$$D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = D(\mathbf{g})(\mathbf{f}(\mathbf{x})) \cdot D(\mathbf{f})(\mathbf{x})$$

- Chain rule for the partial derivatives:

$$\frac{\partial (g_i \circ \mathbf{f})}{\partial x_j}(\mathbf{x}) = \sum_{\alpha=1}^{l} \frac{\partial g_i}{\partial y_\alpha}(\mathbf{f}(\mathbf{x})) \cdot \frac{\partial f_\alpha}{\partial x_j}(\mathbf{x})$$

We derive the second formula from the first one by noticing that $D(\mathbf{f})(\mathbf{x}) = \mathrm{Mat}(d_{\mathbf{x}}\mathbf{f})$. And we derive the third formula from the second one by looking at the (i, j)-component of the matrices (the RHS is just the matrix multiplication formula).

The Chain Rule

Let $U \subset \mathbb{R}^n$ open, $\mathbf{f} : \begin{array}{ccc} U & \to & \mathbb{R}^l \\ \mathbf{x} & \mapsto & \mathbf{f}(\mathbf{x}) \end{array}$, $V \subset \mathbb{R}^l$ open, $\mathbf{g} : \begin{array}{ccc} V & \to & \mathbb{R}^k \\ \mathbf{y} & \mapsto & \mathbf{g}(\mathbf{y}) \end{array}$.
Assume that $\mathbf{f}(U) \subset V$ so that $\mathbf{g} \circ \mathbf{f} : U \to \mathbb{R}^k$ is well-defined.

Let $\mathbf{x} \in U$.
If $\mathbf{f}$ is differentiable at $\mathbf{x}$ and $\mathbf{g}$ is differentiable at $\mathbf{f}(\mathbf{x})$ then $\mathbf{g} \circ \mathbf{f}$ is differentiable at $\mathbf{x}$.

- Chain rule formula for the differentials:

$$d_{\mathbf{x}}(\mathbf{g} \circ \mathbf{f}) = \left(d_{\mathbf{f}(\mathbf{x})}\mathbf{g}\right) \circ \left(d_{\mathbf{x}}\mathbf{f}\right)$$

- Chain rule formula for the Jacobian matrices:

$$D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = D(\mathbf{g})(\mathbf{f}(\mathbf{x})) \cdot D(\mathbf{f})(\mathbf{x})$$

- Chain rule for the partial derivatives:

$$\frac{\partial(g_i \circ \mathbf{f})}{\partial x_j}(\mathbf{x}) = \sum_{\alpha=1}^{l} \frac{\partial g_i}{\partial y_\alpha}(\mathbf{f}(\mathbf{x})) \cdot \frac{\partial f_\alpha}{\partial x_j}(\mathbf{x})$$

The last formula may seem *difficult* but after using it several times you'll notice that it is easy to use in practice, it generalizes the chain rule from MAT135/137/157 in a natural way.

Beware!

Your worst enemy in calculus is going to be the notation!

- There are as many notations as people: if you pick two different textbooks/mathematicians randomly, they probably don't use the same notations for the directional derivatives, the partial derivatives, the differentials, the Jacobian matrices…
  For instance, below are some notations more or less commonly used for the partial derivative of $f : \mathbb{R}^2 \to \mathbb{R}$ with respect to the first variable (i.e. the directional derivative along $\mathbf{e}_1$):

$$\frac{\partial f}{\partial x},\ \partial_x f,\ \partial_{\mathbf{e}_1} f,\ \partial_1 f,\ f_x,\ f'_x,\ D_x f,\ D_{\mathbf{e}_1} f,\ D_1 f,\ D^1 f,\ D^{\mathbf{e}_1} f, \dots$$

- The notations might be confusing at first: be sure that you understand what you are reading and/or writing! Rely on the context to avoid any confusion!
  For instance, given a function $f : \mathbb{R}^2 \to \mathbb{R}$, $\frac{\partial f}{\partial x}$ simply denotes the derivative with respect to the first variable (i.e. the directional derivative along $\mathbf{e}_1$), do not try to interpret the x in the denominator ∂x, that's just a notation.
  Therefore, if you see $\frac{\partial f}{\partial x}(x^2, xyz)$, it means that you **first** compute the partial derivative and **then** that you evaluate it at (x^2, xyz).
  You should **not** compute $f(x^2, xyz)$ and then take the derivative with respect to x.