Elements of Calculus

Prof. Sören Petrat, Constructor University

Lecture notes from Spring 2025

Session 18 (Week 9B)

6. Multivariable Calculus

6.1 Total and Partial Derivatives

Topic for Week 9B: Definitions of Total/Partial and Directional Derivatives

In this chapter we discuss functions with many variables and their derivatives.

In the most general case: $f: \mathbb{R}^n \to \mathbb{R}^m, (x_1, \dots, x_n) \longmapsto \left.\begin{pmatrix} f_1(x_1,\dots,x_n) \\ \vdots \\ f_m(x_1,\dots,x_n) \end{pmatrix}\right\}$ n variables, m components

Usually:

- a fct. $f: \mathbb{R} \to \mathbb{R}^m$ $(m \geq 2)$ is called a curve in $\mathbb{R}^m$,
- a fct. $f: \mathbb{R}^n \to \mathbb{R}$ $(n \geq 2)$ is called a scalar field,
- a fct. $f: \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \geq 2)$ is called a vector field.

In general, we need vectors and matrices to describe fct.s $f: \mathbb{R}^n \to \mathbb{R}^m$ and their derivatives. Therefore, let us recall the following notation and results from Elements of Linear Algebra:

- We write vectors $x \in \mathbb{R}^n$ as $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Their norm is $\|x\| = \left(\sum_{j=1}^n x_j^2 \right)^{\frac{1}{2}}$ and the scalar product is $\langle x, y \rangle = \sum_{j=1}^n x_j y_j$.
- Special vectors are the basis vectors $e_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow j\text{-th component} \\ \\ \\ \\ \end{matrix}$, i.e., $x = \sum_{j=1}^n x_j e_j$
- A map $L: \mathbb{R}^n \to \mathbb{R}^m$ is linear if $L(\lambda x + y) = \lambda L(x) + L(y)$ $\forall x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}$. For linear maps we usually write $L(x) = Lx$.

• linear maps $L: \mathbb{R}^n \to \mathbb{R}^m$ are in one-to-one correspondence to $m \times n$ matrices

$$A_L = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$ by choosing a basis.

Choosing a basis (e_i) of $\mathbb{R}^n$ and $(\tilde{e}_i)$ of $\mathbb{R}^m$, we have $(Lx)_i = \langle \tilde{e}_i, Lx \rangle = \langle \tilde{e}_i, L(\sum_j x_j e_j) \rangle$
$= \sum_j x_j \underbrace{\langle \tilde{e}_i, L e_j \rangle}_{=: a_{ij}}$.

Recall $(Ax)_i = \sum_{j=1}^n a_{ij} x_j$, $(AB)_{ik} = \sum_{j=1}^n a_{ij} b_{jk}$ for A $m \times n$, and B $n \times p$ matrix.
(then AB is an $m \times p$ matrix)

• For linear maps $\mathbb{R}^n \to \mathbb{R}^m$ we define the operator norm $\|L\| := \sup_{\substack{u \in \mathbb{R}^n \\ \|u\| = 1}} \|Lu\| < \infty$.

Since $\|L \frac{u}{\|u\|}\| \leq \|L\|$, we have $\|Lu\| \leq \|L\| \, \|u\|$.

First, we want to define differentiability of $f: \mathbb{R}^n \to \mathbb{R}^m$. Recall that for functions $f: \mathbb{R} \to \mathbb{R}$ we defined differentiability at x_0 as:

$\exists m \in \mathbb{R}$ s.t. for small enough h: $f(x_0 + h) = f(x_0) + mh + r_x(h)$, with $\lim_{h \to 0} \left| \frac{r_x(h)}{h} \right| = 0$.

Clearly, $L_m: \mathbb{R} \to \mathbb{R}$, $h \mapsto mh$ is a linear map.

Key insight: For $f: \mathbb{R}^n \to \mathbb{R}^m$, we still define differentiability by requiring that a linear approximation is possible.

Definition: Let $f: \mathbb{R}^n \to \mathbb{R}^m$. Then f is called differentiable at $x_0 \in \mathbb{R}^n$ if there is a linear map $A: \mathbb{R}^n \to \mathbb{R}^m$ s.t.

$$\underbrace{\boxed{f(x_0+h) = f(x_0) + Ah + r_{x_0}(h) \quad \text{with} \quad \lim_{h\to 0} \frac{\|r_{x_0}(h)\|}{\|h\|} = 0.}}$$

In other words: $\lim_{h \to 0} \frac{\|f(x_0+h) - f(x_0) - Ah\|}{\|h\|} = 0.$

We call $A \equiv Df|_{x_0} \equiv f'(x_0)$ the **total derivative** of f at x_0.

If f is differentiable for all $x_0 \in \mathbb{R}^n$, we say f is differentiable in $\mathbb{R}^n$.

Note:

- Everything is defined analogously when f is only defined on a subset of $\mathbb{R}^n$. Such a subset needs to be the generalization of an open interval, e.g., $B_r(x) = \{y \in \mathbb{R}^n : \|x-y\| < r\}$.
- Clearly differentiability at $x_0 \in \mathbb{R}^n$ implies continuity at x_0 since $\frac{\|r_{x_0}(h)\|}{\|h\|} \to 0$ implies $\|r_{x_0}(h)\| \to 0$.
- For $f: \mathbb{R}^2 \to \mathbb{R}$, differentiability means that we can approximate f near x_0 by a tangent plane. (We provide visualizations later.)

Let us prove that the derivative is unique (if it exists).

Lemma: If $f: \mathbb{R}^n \to \mathbb{R}^n$ is differentiable at $x_0 \in \mathbb{R}^n$, then the derivative $Df|_{x_0}$ is unique.

Proof: Suppose both A_1 and A_2 are derivatives. Then $B := A_1 - A_2$ satisfies

$$\frac{\|Bh\|}{\|h\|} = \frac{1}{\|h\|} \| f(x_0+h) - f(x_0) - r_{1,x_0}(h) - (f(x_0+h) - f(x_0) - r_{2,x_0}(h)) \|$$

$$\leq \frac{\|r_{1,x_0}(h)\|}{\|h\|} + \frac{\|r_{2,x_0}(h)\|}{\|h\|} \xrightarrow{h \to 0} 0.$$

Now fix any $u \in \mathbb{R}^n, u \neq 0$ and choose $h = tu$, $t \in \mathbb{R}$. Then:

$$0 \xleftarrow{t \to 0} \frac{\|(A_1 - A_2)h\|}{\|h\|} = \frac{\|(A_1 - A_2)tu\|}{\|tu\|} = \frac{\|(A_1 - A_2)u\|}{\|u\|}, \text{ i.e., } A_1 u = A_2 u \;\; \forall u \in \mathbb{R}^n$$

$$\Rightarrow A_1 = A_2. \qquad \square$$

Example: $f(x_1, x_2) = \begin{pmatrix} x_1^2 + x_1 x_2 \\ 2x_1 - x_2^2 \end{pmatrix}$

$$f(x_1 + h_1, x_2 + h_2) = \begin{pmatrix} (x_1 + h_1)^2 + (x_1 + h_1)(x_2 + h_2) \\ 2(x_1 + h_1) - (x_2 + h_2)^2 \end{pmatrix} = \underbrace{\begin{pmatrix} x_1^2 + x_1 x_2 \\ 2x_1 - x_2^2 \end{pmatrix}}_{= f(x_1, x_2)} + \underbrace{\begin{pmatrix} 2x_1 h_1 + x_1 h_2 + x_2 h_1 \\ 2h_1 - 2h_2 x_2 \end{pmatrix}}_{= \underbrace{\begin{pmatrix} 2x_1 + x_2 & x_1 \\ 2 & -2x_2 \end{pmatrix}}_{Df|_x} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}} + \underbrace{\begin{pmatrix} h_1^2 + h_1 h_2 \\ -h_2^2 \end{pmatrix}}_{=r(h)}$$

with $\frac{\|r_h\|^2}{\|h\|^2} = \frac{h_1^4 + 2h_1^3 h_2 + h_1^2 h_2^2 + h_2^4}{h_1^2 + h_2^2} \xrightarrow{h_1, h_2 \to 0} 0.$

Next we consider derivatives in different directions:

Definition: $f: \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x_0 \in \mathbb{R}^n$ in the direction $u \in \mathbb{R}^n$, $\|u\| = 1$, if

$$\lim_{t \to 0} \frac{f(x_0 + tu) - f(x_0)}{t}$$ exists. Then this limit is denoted by $D_u f|_{x_0}$ and called **directional derivative**. (or derivative in direction u)

If f is differentiable in the direction e_j, we call $D_{e_j} f|_{x_0} = \frac{\partial f}{\partial x_j}(x_0)$ the j-th **partial derivative** of f at x_0.

In other words: $\frac{\partial f_k}{\partial x_j}(\tilde{x}) = \lim_{t \to 0} \frac{f_k(\tilde{x}_1, \dots, \tilde{x}_{j-1}, \tilde{x}_j + t, \tilde{x}_{j+1}, \dots, \tilde{x}_n) - f_k(\tilde{x})}{t}$.

the 1-dimensional derivative of f_k in the variable x_j only (keeping all other variables fixed)

Ex.: $f(x_1,x_2) = \begin{pmatrix} x_1^2 + x_1 x_2 \\ 2x_1 - x_2^2 \end{pmatrix} \Rightarrow \frac{\partial f}{\partial x_1} = \begin{pmatrix} 2x_1 + x_2 \\ 2 \end{pmatrix}, \frac{\partial f}{\partial x_2} = \begin{pmatrix} x_1 \\ -2x_2 \end{pmatrix}$

Note that in the example we have $Df|_x = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right)$.