Analysis III for engineering study programs

Ingenuin Gasser

Departments of Mathematics
Universität Hamburg

Technische Universität Hamburg–Harburg
Wintersemester 2021/22

based on slides of Prof. Jens Struckmeier from Wintersemster 2020/21

Content of the course Analysis III.

1. Partial derivatives, differential operators.
2. Vector fields, total differential, directional derivative.
3. Mean value theorems, Taylor's theorem.
4. Extrem values, implicit function theorem.
5. Implicit rapresentaion of curves and surfces.
6. Extrem values under equality constraints.
7. Newton–method, non-linear equations and the least squares method.
8. Multiple integrals, Fubini's theorem, transformation theorem.
9. Potentials, Green's theorem, Gauß's theorem.
10. Green's formulas, Stokes's theorem.

1.1 Partial derivatives

Let

$f(x_1, \dots, x_n)$ a scalar function depending n variables

Example: The constitutive law of an ideal gas $pV = RT$.

Each of the 3 quantities p (pressure), V (volume) and T (emperature) can be expressed as a function of the others (R is the gas constant)

$$
\begin{aligned}
p &= p(V,t) = \frac{RT}{V} \\
V &= V(p,T) = \frac{RT}{p} \\
T &= T(p,V) = \frac{pV}{R}
\end{aligned}
$$

Definition: Let $D \subset \mathbb{R}^n$ be open, $f : D \to \mathbb{R}$, $\mathrm{x}^0 \in D$.

- f is called partially differentiable in x^0 with respect to x_i if the limit

$$\frac{\partial f}{\partial x_i}(\mathrm{x}^0) := \lim_{t \to 0} \frac{f(\mathrm{x}^0 + t\mathrm{e}_i) - f(\mathrm{x}^0)}{t}$$

$$= \lim_{t \to 0} \frac{f(x_1^0, \ldots, x_i^0 + t, \ldots, x_n^0) - f(x_1^0, \ldots, x_i^0, \ldots, x_n^0)}{t}$$

  exists. e_i denotes the i–th unit vector. The limit is called partial derivative of f with respect to x_i at x^0.

- If at every point x^0 the partial derivatives with respect to every variable $x_i, i = 1, \ldots, n$ exist and if the partial derivatives are **continuous functions** then we call f continuous partial differentiable or a $\mathcal{C}^1$–function.

Examples.

- Consider the function

$$f(x_1, x_2) = x_1^2 + x_2^2$$

At any point $\mathrm{x}^0 \in \mathbb{R}^2$ there exist both partial derivatives and both partial derivatives are continuous:

$$\frac{\partial f}{\partial x_1}(\mathrm{x}^0) = 2x_1, \qquad \frac{\partial f}{\partial x_2}(\mathrm{x}^0) = 2x_2$$

Thus f is a $\mathcal{C}^1$–function.

- The function

$$f(x_1, x_2) = x_1 + |x_2|$$

at $\mathrm{x}^0 = (0, 0)^T$ is partial differentiable with respect to x_1, but the partial derivative with respect to x_2 does **not** exist!

An engineering example.

The acoustic pressure of a one dimensional acoustic wave is given by

$$p(x,t) = A\sin(\alpha x - \omega t)$$

The partial derivative

$$\frac{\partial p}{\partial x} = \alpha A\cos(\alpha x - \omega t)$$

describes at a given time t the spacial rate of change of the pressure.

The partial derivative

$$\frac{\partial p}{\partial t} = -\omega A\cos(\alpha x - \omega t)$$

describes for a fixed position x the temporal rate of change of the acoustic pressure.

Rules for differentiation

- Let f, g be differentiable with respect to x_i and $\alpha, \beta \in \mathbb{R}$, then we have the rules

$$\frac{\partial}{\partial x_i}\Big(\alpha f(\mathrm{x}) + \beta g(\mathrm{x})\Big) = \alpha \frac{\partial f}{\partial x_i}(\mathrm{x}) + \beta \frac{\partial g}{\partial x_i}(\mathrm{x})$$

$$\frac{\partial}{\partial x_i}\Big(f(\mathrm{x}) \cdot g(\mathrm{x})\Big) = \frac{\partial f}{\partial x_i}(\mathrm{x}) \cdot g(\mathrm{x}) + f(\mathrm{x}) \cdot \frac{\partial g}{\partial x_i}(\mathrm{x})$$

$$\frac{\partial}{\partial x_i}\left(\frac{f(\mathrm{x})}{g(\mathrm{x})}\right) = \frac{\frac{\partial f}{\partial x_i}(\mathrm{x}) \cdot g(\mathrm{x}) - f(\mathrm{x}) \cdot \frac{\partial g}{\partial x_i}(\mathrm{x})}{g(\mathrm{x})^2} \quad \text{for } g(\mathrm{x}) \neq 0$$

- An alternative notation for the partial derivatives of f with respect to x_i at x^0 is given by

$$D_i f(\mathrm{x}^0) \qquad \text{oder} \qquad f_{x_i}(\mathrm{x}^0)$$

Gradient and nabla–operator.

Definition: Let $D \subset \mathbb{R}^n$ be an open set and $f : D \to \mathbb{R}$ partial differentiable.

- We denote the **row vector**

$$\operatorname{grad} f(\mathrm{x}^0) := \left(\frac{\partial f}{\partial x_1}(\mathrm{x}^0), \ldots, \frac{\partial f}{\partial x_n}(\mathrm{x}^0) \right)$$

  as gradient of f at x^0.

- We denote the symbolic vector

$$\nabla := \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n} \right)^T$$

  as nabla–operator.

- Thus we obtain the **column vector**

$$\nabla f(\mathrm{x}^0) := \left(\frac{\partial f}{\partial x_1}(\mathrm{x}^0), \ldots, \frac{\partial f}{\partial x_n}(\mathrm{x}^0) \right)^T$$

More rules on differentiation.

Let f and g be partial differentiable. Then the following rules on differentiation hold true:

$$\begin{aligned} \operatorname{grad}(\alpha f + \beta g) &= \alpha \cdot \operatorname{grad} f + \beta \cdot \operatorname{grad} g \\ \operatorname{grad}(f \cdot g) &= g \cdot \operatorname{grad} f + f \cdot \operatorname{grad} g \\ \operatorname{grad}\left(\frac{f}{g}\right) &= \frac{1}{g^2}(g \cdot \operatorname{grad} f - f \cdot \operatorname{grad} g), \quad g \neq 0 \end{aligned}$$

Examples:

- Let $f(x, y) = e^x \cdot \sin y$. Then:

$$\operatorname{grad} f(x, y) = (e^x \cdot \sin y, e^x \cdot \cos y) = e^x(\sin y, \cos y)$$

- For $r(\mathrm{x}) := \|\mathrm{x}\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$ we have

$$\operatorname{grad} r(\mathrm{x}) = \frac{\mathrm{x}}{r(\mathrm{x})} = \frac{\mathrm{x}}{\|\mathrm{x}\|_2} \qquad \text{für } \mathrm{x} \neq 0,$$

  where $\mathrm{x} = (x_1, \ldots, x_n)$ denotes a row vector.

Partial differentiability does not imply continuity.

Observation: A partial differentiable function (with respect to all coordinates) is not necessarily a continuous function.

Example: Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined as

$$f(x,y) := \begin{cases} \dfrac{x \cdot y}{(x^2+y^2)^2} & : \quad \text{for } (x,y) \neq 0 \\ 0 & : \quad \text{for } (x,y) = 0 \end{cases}$$

The function is partial differntiable on the **entire** $\mathbb{R}^2$ and we have

$$\begin{aligned} f_x(0,0) &= f_y(0,0) = 0 \\ \frac{\partial f}{\partial x}(x,y) &= \frac{y}{(x^2+y^2)^2} - 4\frac{x^2 y}{(x^2+y^2)^3}, \quad (x,y) \neq (0,0) \\ \frac{\partial f}{\partial y}(x,y) &= \frac{x}{(x^2+y^2)^2} - 4\frac{x y^2}{(x^2+y^2)^3}, \quad (x,y) \neq (0,0) \end{aligned}$$

Example (continuation).

We calculate the partial derivatives at the origin $(0,0)$:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t\to 0}\frac{f(t,0)-f(0,0)}{t} = \frac{\frac{t\cdot 0}{(t^2+0^2)^2}-0}{t} = 0$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{t\to 0}\frac{f(0,t)-f(0,0)}{t} = \frac{\frac{0\cdot t}{(0^2+t^2)^2}-0}{t} = 0$$

But: At $(0,0)$ the function is **not** continuous since

$$\lim_{n\to\infty} f\left(\frac{1}{n},\frac{1}{n}\right) = \frac{\frac{1}{n}\cdot\frac{1}{n}}{\left(\frac{1}{n}\cdot\frac{1}{n}+\frac{1}{n}\cdot\frac{1}{n}\right)^2} = \frac{\frac{1}{n^2}}{\frac{4}{n^4}} = \frac{n^2}{4} \to \infty$$

and thus we have

$$\lim_{(x,y)\to(0,0)} f(x,y) \neq f(0,0) = 0$$

Boundedness of the derivatives implies continuity.

To guarantee the continuity of a partial differentiable function we need additional conditions on f.

Theorem: Let $D \subset \mathbb{R}^n$ be an open set. Let $f : D \to \mathbb{R}$ be partial differentiable in a neighborhood of $\mathrm{x}^0 \in D$ and let the partial derivatives $\dfrac{\partial f}{\partial x_i}$, $i = 1, \ldots, n$, be bounded. Then f is continuous in x^0.

Attention: In the previous example the partial derivatives are not bounded in a neighborhood of $(0,0)$ since

$$\frac{\partial f}{\partial x}(x,y) = \frac{y}{(x^2+y^2)^2} - 4\frac{x^2 y}{(x^2+y^2)^3} \quad \text{für } (x,y) \neq (0,0)$$

Proof of the theorem.

For $\|x - x^0\|_\infty < \varepsilon$, $\varepsilon > 0$ sufficiently small we write:

$$\begin{aligned} f(x) - f(x^0) &= (f(x_1, \ldots, x_{n-1}, x_n) - f(x_1, \ldots, x_{n-1}, x_n^0)) \\ &+ (f(x_1, \ldots, x_{n-1}, x_n^0) - f(x_1, \ldots, x_{n-2}, x_{n-1}^0, x_n^0)) \\ &\vdots \\ &+ (f(x_1, x_2^0, \ldots, x_n^0) - f(x_1^0, \ldots, x_n^0)) \end{aligned}$$

For any difference on the right hand side we consider f as a function in one single variable:

$$g(x_n) - g(x_n^0) := f(x_1, \ldots, x_{n-1}, x_n) - f(x_1, \ldots, x_{n-1}, x_n^0)$$

Since f is partial differentiable g is differentiable and we can apply the mean value theorem on g:

$$g(x_n) - g(x_n^0) = g'(\xi_n)(x_n - x_n^0)$$

for an appropriate ξ_n between x_n and x_n^0.

Proof of the theorem (continuation).

Applying the mean value theorem to every term in the right hand side we obtain

$$\begin{aligned} f(x) - f(x^0) &= \frac{\partial f}{\partial x_n}(x_1, \ldots, x_{n-1}, \xi_n) \cdot (x_n - x_n^0) \\ &+ \frac{\partial f}{\partial x_{n-1}}(x_1, \ldots, x_{n-2}, \xi_{n-1}, x_n^0) \cdot (x_{n-1} - x_{n-1}^0) \\ &\vdots \\ &+ \frac{\partial f}{\partial x_1}(\xi_1, x_2^0, \ldots, x_n^0) \cdot (x_1 - x_1^0) \end{aligned}$$

Using the boundedness of the partial derivatives

$$|f(x) - f(x^0)| \leq C_1|x_1 - x_1^0| + \cdots + C_n|x_n - x_n^0|$$

for $\|x - x^0\|_\infty < \varepsilon$, we obtain the continuity of f at x^0 since

$$f(x) \to f(x^0) \qquad \text{für } \|x - x^0\|_\infty \to 0$$

Higher order derivatives.

Definition: Let f be a scalar function and partial differentiable on an open set $D \subset \mathbb{R}^n$. If the partial derivatives are differentiable we obtain (by differentiating) the partial derivatives of second order of f with

$$\frac{\partial^2 f}{\partial x_j \partial x_i} := \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)$$

Example: Second order partial derivatives of a function $f(x, y)$:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y^2}$$

Let $i_1, \ldots, i_k \in \{1, \ldots, n\}$. Then we define recursively

$$\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \ldots \partial x_{i_1}} := \frac{\partial}{\partial x_{i_k}}\left(\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \partial x_{i_{k-2}} \ldots \partial x_{i_1}}\right)$$

Higher order derivatives.

Definition: The function f is called k–times partial differentiable, if all derivatives of order k,

$$\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \ldots \partial x_{i_1}} \qquad \text{for all } i_1, \ldots, i_k \in \{1, \ldots, n\},$$

exist on D.

Alternative notation:

$$\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \ldots \partial x_{i_1}} = D_{i_k} D_{i_{k-1}} \ldots D_{i_1} f = f_{x_{i_1} \ldots x_{i_k}}$$

If all the derivatives of k–th order are continuous the function f is called k–times continuous partial differentiable or called a $\mathcal{C}^k$–function on D. Continuous functions f are called $\mathcal{C}^0$–functions.

Example: For the function $f(x_1, \ldots, x_n) = \prod_{i=1}^{n} x_i^i$ we have $\frac{\partial^n f}{\partial x_n \ldots \partial x_1} =$?

Partial derivaratives are not arbitrarely exchangeable.

ATTENTION: The order how to execute partial derivatives is in general not arbitrarely exchangeable!

Example: For the function

$$f(x,y) := \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & : \quad \text{for } (x,y) \neq (0,0) \\ 0 & : \quad \text{for } (x,y) = (0,0) \end{cases}$$

we calculate

$$f_{xy}(0,0) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(0,0)\right) = -1$$

$$f_{yx}(0,0) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(0,0)\right) = +1$$

i.e. $f_{xy}(0,0) \neq f_{yx}(0,0)$.

Theorem of Schwarz on exchangeablity.

Satz: Let $D \subset \mathbb{R}^n$ be open and let $f : D \to \mathbb{R}$ be a $\mathcal{C}^2$–function. Then it holds

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x_1, \ldots, x_n) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_1, \ldots, x_n)$$

for all $i, j \in \{1, \ldots, n\}$.

Idea of the proof:

Apply the men value theorem twice.

Conclusion:

If f is a C^k–function, then we can exchange the differentiation in order to calculate partial derivatives up to order k arbitrarely!

Example for the exchangeability of partial derivatives.

Calculate the partial derivative of third order f_{xyz} for the function

$$f(x,y,z) = y^2 z \sin(x^3) + (\cosh y + 17e^{x^2})z^2$$

The order of execution is exchangealbe since $f \in \mathcal{C}^3$.

- Differentiate first with respect to z:

$$\frac{\partial f}{\partial z} = y^2 \sin(x^3) + 2z(\cosh y + 17e^{x^2})$$

- Differentiate then f_z with respect to x (then $\cosh y$ disappears):

$$\begin{aligned} f_{zx} &= \frac{\partial}{\partial x}\left(y^2 \sin(x^3) + 2z(\cosh y + 17e^{x^2})\right) \\ &= 3x^2y^2\cos(x^3) + 68xze^{x^2} \end{aligned}$$

- For the partial derivative of f_{zx} with respect to y we obtain

$$f_{xyz} = 6x^2 y \cos(x^3)$$

The Laplace operator.

The Laplace–operator or Laplacian is defined as

$$\Delta := \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$$

For a scalar function $u(x) = u(x_1, \dots, x_n)$ we have

$$\Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} = u_{x_1 x_1} + \cdots + u_{x_n x_n}$$

Examples of important partial differential equations of second order (i.e. equations containing partial derivatives up to order two):

$$\Delta u - \frac{1}{c^2} u_{tt} = 0 \qquad \text{(wave equation)}$$

$$\Delta u - \frac{1}{k} u_t = 0 \qquad \text{(heat equation)}$$

$$\Delta u = 0 \qquad \text{(Laplace–equation or equation for the potential)}$$

Vector valued functions.

Definition: Let $D \subset \mathbb{R}^n$ be open and let $\mathrm{f} : D \to \mathbb{R}^m$ be a vector valued function.

The function f is called partial differentiable on $\mathrm{x}^0 \in D$, if for all $i = 1, \ldots, n$ the limits

$$\frac{\partial \mathrm{f}}{\partial x_i}(\mathrm{x}^0) = \lim_{t \to 0} \frac{\mathrm{f}(\mathrm{x}^0 + t\mathrm{e}_i) - \mathrm{f}(\mathrm{x}^0)}{t}$$

exist. The calculation is done componentwise

$$\frac{\partial \mathrm{f}}{\partial x_i}(\mathrm{x}^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_i} \\ \vdots \\ \frac{\partial f_m}{\partial x_i} \end{pmatrix} \qquad \text{for } i = 1, \ldots, n$$

Vectorfields.

Definition: If $m = n$ the function $\mathrm{f} : D \to \mathbb{R}^n$ is called a vectorfield on D. If every (coordinate-) function $f_i(\mathrm{x})$ of $\mathrm{f} = (f_1, \ldots, f_n)^T$ is a $\mathcal{C}^k$–function, then f is called $\mathcal{C}^k$–vectorfield.

Examples of vectorfields:

- velocity fields of liquids or gases;
- elektromagnetic fields;
- temperature gradients in solid states.

Definition: Let $\mathrm{f} : D \to \mathbb{R}^n$ be a partial differentiable vector field. The divergence on $\mathrm{x} \in D$ is defined as

$$\operatorname{div} \mathrm{f}(\mathrm{x}^0) := \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}(\mathrm{x}^0)$$

or

$$\operatorname{div} \mathrm{f}(\mathrm{x}) = \nabla^T \mathrm{f}(\mathrm{x}) = (\nabla, \mathrm{f}(\mathrm{x}))$$

Rules of computation and the rotation.

The following rules hold true:

$$\begin{aligned}
\operatorname{div}(\alpha\,\mathrm{f} + \beta\,\mathrm{g}) &= \alpha \operatorname{div} \mathrm{f} + \beta \operatorname{div} \mathrm{g} \quad \text{for } \mathrm{f}, \mathrm{g} : D \to \mathbb{R}^n \\
\operatorname{div}(\varphi \cdot \mathrm{f}) &= (\nabla \varphi, \mathrm{f}) + \varphi \operatorname{div} \mathrm{f} \quad \text{for } \varphi : D \to \mathbb{R}, \mathrm{f} : D \to \mathbb{R}^n
\end{aligned}$$

Remark: Let $f : D \to \mathbb{R}$ be a $\mathcal{C}^2$–function, then for the Laplacian we have

$$\Delta f = \operatorname{div}(\nabla f)$$

Definition: Let $D \subset \mathbb{R}^3$ open and $\mathrm{f} : D \to \mathbb{R}^3$ a partial differentiable vector field. We define the rotation as

$$\operatorname{rot} \mathrm{f}(\mathrm{x}^0) := \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right)^T \Bigg|_{\mathrm{x}^0}$$

Alternative notations and additional rules.

$$\text{rot}\, f(x) = \nabla \times f(x) = \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ f_1 & f_2 & f_3 \end{vmatrix}$$

Remark: The following rules hold true:

$$\begin{aligned} \text{rot}\,(\alpha\, f + \beta\, g) &= \alpha\, \text{rot}\, f + \beta\, \text{rot}\, g \\ \text{rot}\,(\varphi \cdot f) &= (\nabla\varphi) \times f + \varphi\, \text{rot}\, f \end{aligned}$$

Remark: Let $D \subset \mathbb{R}^3$ and $\varphi : D \to \mathbb{R}$ be a $\mathcal{C}^2$–function. Then

$$\text{rot}\,(\nabla\varphi) = 0\,,$$

using the exchangeability theorem of Schwarz. I.e. gradient fileds are rotation-free everywhere.

1.2 The total differential

Definition: Let $D \subset \mathbb{R}^n$ open, $x^0 \in D$ and $f : D \to \mathbb{R}^m$. The function $f(x)$ is called differentiable in x^0 (or totally differentiable in x_0), if there exists a linear map

$$l(x, x^0) := A \cdot (x - x^0)$$

with a matrix $A \in \mathbb{R}^{m \times n}$ which satisfies the following approximation property

$$f(x) = f(x^0) + A \cdot (x - x^0) + o(\|x - x^0\|)$$

i.e.

$$\lim_{x \to x^0} \frac{f(x) - f(x^0) - A \cdot (x - x^0)}{\|x - x^0\|} = 0.$$

The total differential and the Jacobian matrix.

Notation: We call the linear map l the differential or the total differential of f(x) at the point x^0. We denote l by $df(x^0)$.

The related matrix A is called Jacobi–matrix of f(x) at the point x^0 and is denoted by $J f(x^0)$ (or $Df(x^0)$ or $f'(x^0)$).

Remark: For $m = n = 1$ we obtain the well known relation

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(|x - x_0|)$$

for the derivative $f'(x_0)$ at the point x_0.

Remark: In case of a scalar function ($m = 1$) the matrix $A = a$ is a row vextor and $a(x - x^0)$ a scalar product $\langle a^T, x - x^0 \rangle$.

Total and partial differentiability.

Theorem: Let $\mathrm{f} : D \to \mathbb{R}^m$, $\mathrm{x}^0 \in D \subset \mathbb{R}^n$, D open.

a) If $\mathrm{f}(\mathrm{x})$ is differentiable in x^0, then $\mathrm{f}(\mathrm{x})$ is continuous in x^0.

b) If $\mathrm{f}(\mathrm{x})$ is differentiable in x^0, then the (total) differential and thus the Jacobi–matrix are uniquely determined and we have

$$\mathrm{J\,f}(\mathrm{x}^0) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathrm{x}^0) & \dots & \dfrac{\partial f_1}{\partial x_n}(\mathrm{x}^0) \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(\mathrm{x}^0) & \dots & \dfrac{\partial f_m}{\partial x_n}(\mathrm{x}^0) \end{pmatrix} = \begin{pmatrix} Df_1(\mathrm{x}^0) \\ \vdots \\ Df_m(\mathrm{x}^0) \end{pmatrix}$$

c) If $\mathrm{f}(\mathrm{x})$ is a $\mathcal{C}^1$–function on D, then $\mathrm{f}(\mathrm{x})$ is differentiable on D.

Proof of a).

If f is differentiable in x^0, then by definition

$$\lim_{x \to x^0} \frac{f(x) - f(x^0) - A \cdot (x - x^0)}{\|x - x^0\|} = 0$$

Thus we conclude

$$\lim_{x \to x^0} \|f(x) - f(x^0) - A \cdot (x - x^0)\| = 0$$

and we obtain

$$\begin{aligned} \|f(x) - f(x^0)\| &\leq \|f(x) - f(x^0) - A \cdot (x - x^0)\| + \|A \cdot (x - x^0)\| \\ &\to 0 \qquad \text{as } x \to x^0 \end{aligned}$$

Therefore the function f is continuous at x^0.

Proof of b).

Let $x = x^0 + te_i$, $|t| < \varepsilon$, $i \in \{1, \dots, n\}$. Since f in differentiable at x^0, we have

$$\lim_{x \to x^0} \frac{f(x) - f(x^0) - A \cdot (x - x^0)}{\|x - x^0\|_\infty} = 0$$

We write

$$\begin{aligned} \frac{f(x) - f(x^0) - A \cdot (x - x^0)}{\|x - x^0\|_\infty} &= \frac{f(x^0 + te_i) - f(x^0)}{|t|} - \frac{tAe_i}{|t|} \\ &= \frac{t}{|t|} \cdot \left(\frac{f(x^0 + te_i) - f(x^0)}{t} - Ae_i \right) \\ &\to 0 \qquad \text{as } t \to 0 \end{aligned}$$

Thus

$$\lim_{t \to 0} \frac{f(x^0 + te_i) - f(x^0)}{t} = Ae_i \qquad i = 1, \dots, n$$

Examples.

- Consider the scalar function $f(x_1, x_2) = x_1 e^{2x_2}$. Then the Jacobian is given by:
$$Jf(x_1, x_2) = Df(x_1, x_2) = e^{2x_2}(1, 2x_1)$$
- Consider the function $\mathrm{f} : \mathbb{R}^3 \to \mathbb{R}^2$ defined by
$$\mathrm{f}(x_1, x_2, x_3) = \begin{pmatrix} x_1 x_2 x_3 \\ \sin(x_1 + 2x_2 + 3x_3) \end{pmatrix}$$
The Jacobian is given by
$$\mathrm{Jf}(x_1, x_2, x_3) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \end{pmatrix} = \begin{pmatrix} x_2 x_3 & x_1 x_3 & x_1 x_2 \\ \cos(s) & 2\cos(s) & 3\cos(s) \end{pmatrix}$$
with $s = x_1 + 2x_2 + 3x_3$.

Further examples.

- Let $f(x) = Ax$, $A \in \mathbb{R}^{m\times n}$ and $x \in \mathbb{R}^n$. Then

$$Jf(x) = A \qquad \text{for all } x \in \mathbb{R}^n$$

- Let $f(x) = x^T Ax = \langle x, Ax\rangle$, $A \in \mathbb{R}^{n\times n}$ and $x \in \mathbb{R}^n$. Then we have

$$\begin{aligned} \frac{\partial f}{\partial x_i} &= \langle e_i, Ax\rangle + \langle x, Ae_i\rangle \\ &= e_i^T Ax + x^T Ae_i \\ &= x^T(A^T + A)e_i \end{aligned}$$

  We conclude

$$Jf(x) = \text{grad} f(x) = x^T(A^T + A)$$

Rules for the differentiation.

Theorem:

a) **Linearität:** LET $\mathrm{f}, \mathrm{g} : D \to \mathbb{R}^m$ be differentiable in $\mathrm{x}^0 \in D$, D open. Then $\alpha\,\mathrm{f}(\mathrm{x}^0) + \beta\,\mathrm{g}(\mathrm{x}^0)$, and $\alpha, \beta \in \mathbb{R}$ are differentiable in x^0 and we have

$$\begin{aligned} \mathrm{d}(\alpha \mathrm{f} + \beta \mathrm{g})(\mathrm{x}^0) &= \alpha\,\mathrm{df}(\mathrm{x}^0) + \beta\,\mathrm{dg}(\mathrm{x}^0) \\ \mathrm{J}(\alpha \mathrm{f} + \beta \mathrm{g})(\mathrm{x}^0) &= \alpha\,\mathrm{Jf}(\mathrm{x}^0) + \beta\,\mathrm{Jg}(\mathrm{x}^0) \end{aligned}$$

b) **Chain rule:** Let $\mathrm{f} : D \to \mathbb{R}^m$ be differentiable in $\mathrm{x}^0 \in D$, D open. Let $\mathrm{g} : E \to \mathbb{R}^k$ be differentiable in $\mathrm{y}^0 = f(\mathrm{x}^0) \in E \subset \mathbb{R}^m$, E open. Then $g \circ f$ is differentiable in x^0.

For the differentials it holds

$$\mathrm{d}(\mathrm{g} \circ \mathrm{f})(\mathrm{x}^0) = \mathrm{dg}(\mathrm{y}^0) \circ \mathrm{df}(\mathrm{x}^0)$$

and analogously for the Jacobian matrix

$$\mathrm{J}(\mathrm{g} \circ \mathrm{f})(\mathrm{x}^0) = \mathrm{Jg}(\mathrm{y}^0) \cdot \mathrm{Jf}(\mathrm{x}^0)$$

Examples for the chain rule.

Let $I \subset \mathbb{R}$ be an intervall. Let $\mathrm{h} : I \to \mathbb{R}^n$ be a curve, differentiable in $t_0 \in I$ with values in $D \subset \mathbb{R}^n$, D open. Let $f : D \to \mathbb{R}$ be a scalar function, differentiable in $\mathrm{x}^0 = \mathrm{h}(t_0)$.

Then the composition

$$(f \circ \mathrm{h})(t) = f(h_1(t), \ldots, h_n(t))$$

is differentiable in t_0 and we have for the derivative:

$$\begin{aligned}(f \circ \mathrm{h})'(t_0) &= \mathrm{J}f(\mathrm{h}(t_0)) \cdot \mathrm{Jh}(t_0) \\ &= \mathrm{grad} f(\mathrm{h}(t_0)) \cdot \mathrm{h}'(t_0) \\ &= \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathrm{h}(t_0)) \cdot h_k'(t_0)\end{aligned}$$

Directional derivative.

Definition: Let $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open, $\mathrm{x}^0 \in D$, and $\mathrm{v} \in \mathbb{R} \setminus \{0\}$ a vector. Then

$$D_{\mathrm{v}} f(\mathrm{x}^0) := \lim_{t \to 0} \frac{f(\mathrm{x}^0 + t\mathrm{v}) - f(\mathrm{x}^0)}{t}$$

is called the directional derivative (Gateaux–derivative) of $f(\mathrm{x})$ in the direction of v.

Example: Let $f(x, y) = x^2 + y^2$ and $\mathrm{v} = (1, 1)^T$. Then the directional derivative in the direction of v is given by:

$$\begin{aligned} D_{\mathrm{v}} f(x, y) &= \lim_{t \to 0} \frac{(x + t)^2 + (y + t)^2 - x^2 - y^2}{t} \\ &= \lim_{t \to 0} \frac{2xt + t^2 + 2yt + t^2}{t} \\ &= 2(x + y) \end{aligned}$$

Remarks.

- For $\mathrm{v} = \mathrm{e}_i$ the directional derivative in the direction of v is given by the partial derivative with respect to x_i:

$$D_{\mathrm{v}} f(\mathrm{x}^0) = \frac{\partial f}{\partial x_i}(\mathrm{x}^0)$$

- If v is a unit vector, i.e. $\|\mathrm{v}\| = 1$, then the directional derivative $D_{\mathrm{v}} f(\mathrm{x}^0)$ describes the slope of $f(\mathrm{x})$ in the direction of v.

- If $f(\mathrm{x})$ is differentiable in x^0, then all directional derivatives of $f(\mathrm{x})$ in x^0 exist. With $\mathrm{h}(t) = \mathrm{x}^0 + t\mathrm{v}$ we have

$$D_{\mathrm{v}} f(\mathrm{x}^0) = \frac{d}{dt}(f \circ \mathrm{h})|_{t=0} = \mathrm{grad}\, f(\mathrm{x}^0) \cdot \mathrm{v}$$

This follows directely applying the chain rule.

Properties of the gradient.

Theorem: Let $D \subset \mathbb{R}^n$ open, $f : D \to \mathbb{R}$ differentiable in $\mathrm{x}^0 \in D$. Then we have

a) The gradient vector $\operatorname{grad} f(\mathrm{x}^0) \in \mathbb{R}^n$ is orthogonal in the level set

$$N_{\mathrm{x}^0} := \{\mathrm{x} \in D \mid f(\mathrm{x}) = f(\mathrm{x}^0)\}$$

In the case of $n = 2$ we call the level sets contour lines, in $n = 3$ we call the level sets equipotential surfaces.

2) The gradient $\operatorname{grad} f(\mathrm{x}^0)$ gives the direction of the steepest slope of $f(\mathrm{x})$ in x^0.

Idea of the proof:

a) application of the chain rule.

b) for an arbitrary direction v we conclude with the Cauchy–Schwarz inequality

$$|D_{\mathrm{v}} f(\mathrm{x}^0)| = |(\operatorname{grad} f(\mathrm{x}^0), \mathrm{v})| \;\leq\; \|\operatorname{grad} f(\mathrm{x}^0)\|_2$$

Equality is obtained for $\mathrm{v} = \operatorname{grad} f(\mathrm{x}^0)/\|\operatorname{grad} f(\mathrm{x}^0)\|_2$.

Curvilinear coordinates.

Definition: Let $U, V \subset \mathbb{R}^n$ be open and $\Phi : U \to V$ be a $\mathcal{C}^1$-map, for which the Jacobimatrix $\mathrm{J}\Phi(\mathrm{u}^0)$ is regular (invertible) at every $\mathrm{u}^0 \in U$.

In addition there exists the inverse map $\Phi^{-1} : V \to U$ and the inverse map is also a $\mathcal{C}^1$–map.

Then $\mathrm{x} = \Phi(\mathrm{u})$ defines a coodinate transformation from the coordinates u to x.

Example: Consider for $n = 2$ the polar coordinates $\mathrm{u} = (r, \varphi)$ with $r > 0$ and $-\pi < \varphi < \pi$ and set

$$\begin{aligned} x &= r\cos\varphi \\ y &= r\sin\varphi \end{aligned}$$

with the cartesian coordinates $\mathrm{x} = (x, y)$.

Calculation of the partial derivatives.

For all $u \in U$ with $x = \Phi(u)$ the following relations hold

$$\begin{aligned} \Phi^{-1}(\Phi(u)) &= u \\ J\,\Phi^{-1}(x) \cdot J\,\Phi(u) &= I_n \qquad \text{(chain rule)} \\ J\,\Phi^{-1}(x) &= (J\,\Phi(u))^{-1} \end{aligned}$$

Let $\tilde{f} : V \to \mathbb{R}$ be a given function. Set

$$f(u) := \tilde{f}(\Phi(u))$$

the by using the chain rule we obtain

$$\frac{\partial f}{\partial u_i} = \sum_{j=1}^{n} \frac{\partial \tilde{f}}{\partial x_j} \frac{\partial \Phi_j}{\partial u_i} =: \sum_{j=1}^{n} g^{ij} \frac{\partial \tilde{f}}{\partial x_j}$$

with

$$g^{ij} := \frac{\partial \Phi_j}{\partial u_i}, \qquad G(u) := (g^{ij}) = (J\,\Phi(u))^T$$

Notations.

We use the short notation

$$\frac{\partial}{\partial u_i} = \sum_{j=1}^{n} g^{ij} \frac{\partial}{\partial x_j}$$

Analogously we can express the partial derivatives with respect to x_i by the partial derivatives with respect to u_j

$$\frac{\partial}{\partial x_i} = \sum_{j=1}^{n} g_{ij} \frac{\partial}{\partial u_j}$$

where

$$(g_{ij}) := (g^{ij})^{-1} = (\mathrm{J}\,\Phi)^{-T} = (\mathrm{J}\,\Phi^{-1})^{T}$$

We obtain these relations by applying the chain rule on Φ^{-1}.

Example: polar coordinates.

We consider polar coordinates

$$\mathrm{x} = \Phi(\mathrm{u}) = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}$$

We calculate

$$\mathrm{J}\,\Phi(\mathrm{u}) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

and thus

$$(g^{ij}) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ & \\ -r\sin\varphi & r\cos\varphi \end{pmatrix} \qquad (g_{ij}) = \begin{pmatrix} \cos\varphi & -\dfrac{1}{r}\sin\varphi \\ & \\ \sin\varphi & \dfrac{1}{r}\cos\varphi \end{pmatrix}$$

Partial derivatives for polar coordinates.

The calculation of the partial derivatives gives

$$\frac{\partial}{\partial x} = \cos\varphi \frac{\partial}{\partial r} - \frac{1}{r}\sin\varphi \frac{\partial}{\partial \varphi}$$

$$\frac{\partial}{\partial y} = \sin\varphi \frac{\partial}{\partial r} + \frac{1}{r}\cos\varphi \frac{\partial}{\partial \varphi}$$

Example: Calculation of the Laplacian–operator in polar coordinates

$$\frac{\partial^2}{\partial x^2} = \cos^2\varphi \frac{\partial^2}{\partial r^2} - \frac{\sin(2\varphi)}{r}\frac{\partial^2}{\partial r \partial \varphi} + \frac{\sin^2\varphi}{r^2}\frac{\partial^2}{\partial \varphi^2} + \frac{\sin(2\varphi)}{r^2}\frac{\partial}{\partial \varphi} + \frac{\sin^2\varphi}{r}\frac{\partial}{\partial r}$$

$$\frac{\partial^2}{\partial y^2} = \sin^2\varphi \frac{\partial^2}{\partial r^2} + \frac{\sin(2\varphi)}{r}\frac{\partial^2}{\partial r \partial \varphi} + \frac{\cos^2\varphi}{r^2}\frac{\partial^2}{\partial \varphi^2} - \frac{\sin(2\varphi)}{r^2}\frac{\partial}{\partial \varphi} + \frac{\cos^2\varphi}{r}\frac{\partial}{\partial r}$$

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2} + \frac{1}{r}\frac{\partial}{\partial r}$$

Example: spherical coordinates.

We consider spherical coordinates

$$x = \Phi(u) = \begin{pmatrix} r\cos\varphi\cos\theta \\ r\sin\varphi\cos\theta \\ r\sin\theta \end{pmatrix}$$

The Jacobian–matrix is given by:

$$J\,\Phi(u) = \begin{pmatrix} \cos\varphi\cos\theta & -r\sin\varphi\cos\theta & -r\cos\varphi\sin\theta \\ \sin\varphi\cos\theta & r\cos\varphi\cos\theta & -r\sin\varphi\sin\theta \\ \sin\theta & 0 & r\cos\theta \end{pmatrix}$$

Partial derivatives for spherical coordinates.

Calculating the partial derivatives gives

$$\frac{\partial}{\partial x} = \cos\varphi \cos\theta \frac{\partial}{\partial r} - \frac{\sin\varphi}{r\cos\theta}\frac{\partial}{\partial\varphi} - \frac{1}{r}\cos\varphi \sin\theta \frac{\partial}{\partial\theta}$$

$$\frac{\partial}{\partial y} = \sin\varphi \cos\theta \frac{\partial}{\partial r} + \frac{\cos\varphi}{r\cos\theta}\frac{\partial}{\partial\varphi} - \frac{1}{r}\sin\varphi \sin\theta \frac{\partial}{\partial\theta}$$

$$\frac{\partial}{\partial z} = \sin\theta \frac{\partial}{\partial r} + \frac{1}{r}\cos\theta \frac{\partial}{\partial\theta}$$

Example: calculation of the Laplace–operator in spherical coordinates

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2\cos^2\theta}\frac{\partial^2}{\partial\varphi^2} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{\tan\theta}{r^2}\frac{\partial}{\partial\theta}$$

1.3 Mean value theorems and Taylor expansion

Theorem (Mean value theorem): Let $f : D \to \mathbb{R}$ be a scalar differentiable function on an open set $D \subset \mathbb{R}^n$. Let $\mathrm{a}, \mathrm{b} \in D$ be points in D such that the connecting line segment

$$[\mathrm{a}, \mathrm{b}] := \{\mathrm{a} + t(\mathrm{b} - \mathrm{a}) \mid t \in [0, 1]\}$$

lies entirely in D. Then there exits a number $\theta \in (0, 1)$ with

$$f(\mathrm{b}) - f(\mathrm{a}) = \operatorname{grad} f(\mathrm{a} + \theta(\mathrm{b} - \mathrm{a})) \cdot (\mathrm{b} - \mathrm{a})$$

Proof: We set

$$h(t) := f(\mathrm{a} + t(\mathrm{b} - \mathrm{a}))$$

with the mean value theorem for a single variable and the chain rules we conclude

$$\begin{aligned} f(\mathrm{b}) - f(\mathrm{a}) &= h(1) - h(0) = h'(\theta) \cdot (1 - 0) \\ &= \operatorname{grad} f(\mathrm{a} + \theta(\mathrm{b} - \mathrm{a})) \cdot (\mathrm{b} - \mathrm{a}) \end{aligned}$$

Definition and example.

Definition: If the condition $[a, b] \subset D$ holds true for **all** points $a, b \in D$, then the set D is called convex.

Example for the mean value theorem: Given a scalar function

$$f(x, y) := \cos x + \sin y$$

It is

$$f(0, 0) = f(\pi/2, \pi/2) = 1 \quad \Rightarrow \quad f(\pi/2, \pi/2) - f(0, 0) = 0$$

Applying the mean value theorem there exists a $\theta \in (0, 1)$ with

$$\operatorname{grad} f \left(\theta \begin{pmatrix} \pi/2 \\ \pi/2 \end{pmatrix} \right) \cdot \begin{pmatrix} \pi/2 \\ \pi/2 \end{pmatrix} = 0$$

Indeed this is true for $\theta = \frac{1}{2}$.

Mean value theorem is only true for scalar functions.

Attention: The mean value theorem for multivariate functions is only true for scalar functions but in general not for vector–valued functions!

Examples: Consider the **vector–valued** Function

$$\mathrm{f}(t) := \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \qquad t \in [0, \pi/2]$$

It is

$$\mathrm{f}(\pi/2) - \mathrm{f}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

and

$$\mathrm{f}'\left(\theta\,\frac{\pi}{2}\right) \cdot \left(\frac{\pi}{2} - 0\right) = \frac{\pi}{2} \begin{pmatrix} -\sin(\theta\pi/2) \\ \cos(\theta\pi/2) \end{pmatrix}$$

BUT: the vectors on the right hand side have lenght $\sqrt{2}$ and $\pi/2$!

A mean value estimate for vector–valued functions.

Theorem: Let $f : D \to \mathbb{R}^m$ be differentiable on an open set $D \subset \mathbb{R}^n$. Let a, b bei points in D with $[a, b] \subset D$. Then there exists a $\theta \in (0, 1)$ with

$$\|f(b) - f(a)\|_2 \leq \|J f(a + \theta(b - a)) \cdot (b - a)\|_2$$

Idea of the proof: Application of the mean value theorem to the scalar function $g(x)$ definid as

$$g(x) := (f(b) - f(a))^T f(x) \qquad \text{(scalar product!)}$$

Remark: Another (weaker) for of the mean value estimate is

$$\|f(b) - f(a)\| \leq \sup_{\xi \in [a,b]} \|J f(\xi))\| \cdot \|(b - a)\|$$

where $\| \cdot \|$ denotes an arbitrary vector norm with related matrix norm.

Taylor series: notations.

We define the multi–index $\alpha \in \mathbb{N}_0^n$ as

$$\alpha := (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$$

Let

$$|\alpha| := \alpha_1 + \cdots + \alpha_n \qquad \alpha! := \alpha_1! \cdot \cdots \cdot \alpha_n!$$

Let $f : D \to \mathbb{R}$ be $|\alpha|$ times continuous differentiable. Then we set

$$D^\alpha = D_1^{\alpha_1} D_2^{\alpha_2} \ldots D_n^{\alpha_n} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}},$$

where $D_i^{\alpha_i} = \underbrace{D_i \ldots D_i}_{\alpha_i\text{–mal}}$. We write

$$\mathrm{x}^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} \qquad \text{for } \mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

The Taylor theorem.

Theorem: (Taylor)
Let $D \subset \mathbb{R}^n$ be open and convex. Let $f : D \to \mathbb{R}$ be a $\mathcal{C}^{m+1}$–function and $\mathrm{x}_0 \in D$. Then the Taylor–expansion holds true in $\mathrm{x} \in D$

$$\begin{aligned} f(\mathrm{x}) &= T_m(\mathrm{x};\mathrm{x}_0) + R_m(\mathrm{x};\mathrm{x}_0) \\ T_m(\mathrm{x};\mathrm{x}_0) &= \sum_{|\alpha|\le m} \frac{D^\alpha f(\mathrm{x}_0)}{\alpha!}(\mathrm{x}-\mathrm{x}_0)^\alpha \\ R_m(\mathrm{x};\mathrm{x}_0) &= \sum_{|\alpha|=m+1} \frac{D^\alpha f(\mathrm{x}_0+\theta(\mathrm{x}-\mathrm{x}_0))}{\alpha!}(\mathrm{x}-\mathrm{x}_0)^\alpha \end{aligned}$$

for an appropriate $\theta \in (0,1)$.

Notation: In the Taylor–expansion we denote $T_m(\mathrm{x};\mathrm{x}_0)$ Taylor–polynom of degree m and $R_m(\mathrm{x};\mathrm{x}_0)$ Lagrange–remainder.

Derivation of the Taylor expansion.

We define a scalar function in one single variable $t \in [0, 1]$ as

$$g(t) := f(\mathrm{x}_0 + t(\mathrm{x} - \mathrm{x}_0))$$

and calculate the (univariate) Taylor–expansion at $t = 0$. It is

$$g(1) = g(0) + g'(0) \cdot (1 - 0) + \frac{1}{2} g''(\xi) \cdot (1 - 0)^2 \quad \text{for a } \xi \in (0, 1).$$

The calculation of $g'(0)$ is given by the chain rule

$$\begin{aligned} g'(0) &= \frac{d}{dt} f(x_1^0 + t(x_1 - x_1^0), x_2^0 + t(x_2 - x_2^0), \ldots, x_n^0 + t(x_n - x_n^0))\Big|_{t=0} \\ &= D_1 f(\mathrm{x}_0) \cdot (x_1 - x_1^0) + \ldots + D_n f(\mathrm{x}_0) \cdot (x_n - x_n^0) \\ &= \sum_{|\alpha|=1} \frac{D^\alpha f(\mathrm{x}_0)}{\alpha!} \cdot (\mathrm{x} - \mathrm{x}_0)^\alpha \end{aligned}$$

Continuation of the derivation.

Calculation of $g''(0)$ gives

$$g''(0) = \frac{d^2}{dt^2} f(\mathrm{x}_0 + t(\mathrm{x} - \mathrm{x}_0))\Big|_{t=0} = \frac{d}{dt} \sum_{k=1}^{n} D_k f(\mathrm{x}^0 + t(\mathrm{x} - \mathrm{x}^0))(x_k - x_k^0)\Big|_{t=0}$$

$$= D_{11} f(\mathrm{x}_0)(x_1 - x_1^0)^2 + D_{21} f(\mathrm{x}_0)(x_1 - x_1^0)(x_2 - x_2^0)$$

$$+ \ldots + D_{ij} f(\mathrm{x}_0)(x_i - x_i^0)(x_j - x_j^0) + \ldots +$$

$$+ D_{n-1,n} f(\mathrm{x}_0)(x_{n-1} - x_{n-1}^0)(x_n - x_n^0) + D_{nn} f(\mathrm{x}_0)(x_n - x_n^0)^2)$$

$$= \sum_{|\alpha|=2} \frac{D^\alpha f(\mathrm{x}_0)}{\alpha!} (\mathrm{x} - \mathrm{x}_0)^\alpha \qquad \text{(exchange theorem of Schwarz!)}$$

Continuation: Proof of the Taylor–formula by (mathematical) induction!

Proof of the Taylor theorem.

The function

$$g(t) := f(\mathrm{x}^0 + t(\mathrm{x} - \mathrm{x}^0))$$

is $(m+1)$–times continuous differentiable and we have

$$g(1) = \sum_{k=0}^{m} \frac{g^{(k)}(0)}{k!} + \frac{g^{(m+1)}(\theta)}{(m+1)!} \quad \text{for a } \theta \in [0,1].$$

In addition we have (by induction over k)

$$\frac{g^{(k)}(0)}{k!} = \sum_{|\alpha|=k} \frac{D^\alpha f(\mathrm{x}^0)}{\alpha!} (\mathrm{x} - \mathrm{x}^0)^\alpha$$

and

$$\frac{g^{(m+1)}(\theta)}{(m+1)!} = \sum_{|\alpha|=m+1} \frac{D^\alpha f(\mathrm{x}^0 + \theta(\mathrm{x} - \mathrm{x}^0))}{\alpha!} (\mathrm{x} - \mathrm{x}^0)^\alpha$$

Examples for the Taylor–expansion.

1. Calculate the Taylor–polynom $T_2(\mathrm{x};\mathrm{x}_0)$ of degree 2 of the function

$$f(x,y,z) = x\,y^2\,\sin z$$

at $(x,y,z) = (1,2,0)^T$.

2. The calculation of $T_2(\mathrm{x};\mathrm{x}_0)$ requires the partial derivatives up to order 2.

3. These derivatives have to be evaluated at $(x,y,z) = (1,2,0)^T$.

4. The result is $T_2(\mathrm{x};\mathrm{x}_0)$ in the form

$$T_2(\mathrm{x};\mathrm{x}_0) = 4z(x+y-2)$$

5. Details on extra slide.

Remarks to the remainder of a Taylor–expansion.

Remark: The remainder of a Taylor–expansion contains **all** partial derivatives of order $(m+1)$:

$$R_m(\mathrm{x};\mathrm{x}_0) = \sum_{|\alpha|=m+1} \frac{D^\alpha f(\mathrm{x}_0 + \theta(\mathrm{x}-\mathrm{x}_0))}{\alpha!} (\mathrm{x}-\mathrm{x}_0)^\alpha$$

If all these derivative are bounded by aconstant C in a neighborhood of x_0 then the estimate for the remainder hold true

$$|R_m(\mathrm{x};\mathrm{x}_0)| \leq \frac{n^{m+1}}{(m+1)!} C \, \|\mathrm{x}-\mathrm{x}_0\|_\infty^{m+1}$$

We conlude for the quality of the approximation of a $\mathcal{C}^{m+1}$–function by the Taylor–polynom

$$f(\mathrm{x}) = T_m(\mathrm{x};\mathrm{x}_0) + O\left(\|\mathrm{x}-\mathrm{x}_0\|^{m+1}\right)$$

Special case $m=1$: For a $\mathcal{C}^2$–function $f(\mathrm{x})$ we obtain

$$f(\mathrm{x}) = f(\mathrm{x}^0) + \operatorname{grad} f(\mathrm{x}^0)\cdot(\mathrm{x}-\mathrm{x}^0) + O(\|\mathrm{x}-\mathrm{x}^0\|^2).$$