

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 Wintersemester 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 representation of curves and surfaces.
- 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.

Chapter 1. Multi variable differential calculus

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 (temperature) can be expressed as a function of the others (R is the gas constant)

$$p = p(V, T) = \frac{RT}{V}$$

$$V = V(p, T) = \frac{RT}{p}$$

$$T = T(p, V) = \frac{pV}{R}$$

1.1. Partial derivatives

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

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

$$\begin{aligned}\frac{\partial f}{\partial x_i}(x^0) &:= \lim_{t \rightarrow 0} \frac{f(x^0 + te_i) - f(x^0)}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(x_1^0, \dots, x_i^0 + t, \dots, x_n^0) - f(x_1^0, \dots, x_i^0, \dots, x_n^0)}{t}\end{aligned}$$

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 $x_i, i = 1, \dots, n$ 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 $x^0 \in \mathbb{R}^2$ there exist both partial derivatives and both partial derivatives are continuous:

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

Thus f is a \mathcal{C}^1 -function.

- The function

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

at $x^0 = (0, 0)^T$ 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

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

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

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

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

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

Gradient and nabla-operator.

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

- We denote the **row vector**

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

as **gradient** of f at x^0 .

- We denote the symbolic vector

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

as **nabla-operator**.

- Thus we obtain the **column vector**

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

More rules on differentiation.

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

$$\text{grad}(\alpha f + \beta g) = \alpha \cdot \text{grad} f + \beta \cdot \text{grad} g$$

$$\text{grad}(f \cdot g) = g \cdot \text{grad} f + f \cdot \text{grad} g$$

$$\text{grad}\left(\frac{f}{g}\right) = \frac{1}{g^2}(g \cdot \text{grad} f - f \cdot \text{grad} g), \quad g \neq 0$$

Examples:

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

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

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

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

where $x = (x_1, \dots, x_n)$ 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 $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined as

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

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

$$f_x(0, 0) = f_y(0, 0) = 0$$

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

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

Example (continuation).

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

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

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

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

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

and thus we have

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

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 $f : D \rightarrow \mathbb{R}$ be partial differentiable in a neighborhood of $x^0 \in D$ and let the partial derivatives $\frac{\partial f}{\partial x_i}$, $i = 1, \dots, n$, 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:

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

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

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

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

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

Using the boundedness of the partial derivatives

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

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

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

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

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

Higher order derivatives.

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

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

exist on D .

Alternative notation:

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

If all the derivatives of k -th order are continuous the function f is called k -times continuous partial differentiable or called a C^k -function on D . Continuous functions f are called C^0 -functions.

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

Partial derivatives are not arbitrarily exchangeable.

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

Example: For the function

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

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 exchangeability.

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

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

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

Idea of the proof:

Apply the mean 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 **arbitrarily!**

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 exchangeable 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):

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

- 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

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

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

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

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

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

Vector valued functions.

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

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

$$\frac{\partial f}{\partial x_i}(x^0) = \lim_{t \rightarrow 0} \frac{f(x^0 + te_i) - f(x^0)}{t}$$

exist. The calculation is done componentwise

$$\frac{\partial f}{\partial x_i}(x^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_i} \\ \frac{\partial f_2}{\partial x_i} \\ \vdots \\ \frac{\partial f_m}{\partial x_i} \end{pmatrix} \quad \text{for } i = 1, \dots, n$$

Vectorfields.

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

Examples of vectorfields:

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

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

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

or

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

Rules of computation and the rotation.

The following rules hold true:

$$\operatorname{div}(\alpha f + \beta g) = \alpha \operatorname{div} f + \beta \operatorname{div} g \quad \text{for } f, g : D \rightarrow \mathbb{R}^n$$

$$\operatorname{div}(\varphi \cdot f) = (\nabla \varphi, f) + \varphi \operatorname{div} f \quad \text{for } \varphi : D \rightarrow \mathbb{R}, f : D \rightarrow \mathbb{R}^n$$

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

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

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

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

Alternative notations and additional rules.

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

Remark: The following rules hold true:

$$\operatorname{rot}(\alpha f + \beta g) = \alpha \operatorname{rot} f + \beta \operatorname{rot} g$$

$$\operatorname{rot}(\varphi \cdot f) = (\nabla \varphi) \times f + \varphi \operatorname{rot} f$$

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

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

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

1.2 The total differential

Definition: Let $D \subset \mathbb{R}^n$ open, $x^0 \in D$ and $f : D \rightarrow \mathbb{R}^m$. 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.

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

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 $Jf(x^0)$ (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 vector and $a(x - x^0)$ a scalar product $\langle a^T, x - x^0 \rangle$.

Total and partial differentiability.

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

- a) If $f(x)$ is differentiable in x^0 , then $f(x)$ is continuous in x^0 .
- b) If $f(x)$ is differentiable in x^0 , then the (total) differential and thus the Jacobi-matrix are uniquely determined and we have

$$Jf(x^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^0) & \dots & \frac{\partial f_1}{\partial x_n}(x^0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x^0) & \dots & \frac{\partial f_m}{\partial x_n}(x^0) \end{pmatrix} = \begin{pmatrix} Df_1(x^0) \\ \vdots \\ Df_m(x^0) \end{pmatrix}$$

- c) If $f(x)$ is a C^1 -function on D , then $f(x)$ is differentiable on D .

Proof of a).

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

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

Thus we conclude

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

and we obtain

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

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 is differentiable at x^0 , we have

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

We write

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

Thus

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

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 $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ defined by

$$f(x_1, x_2, x_3) = \begin{pmatrix} x_1 x_2 x_3 \\ \sin(x_1 + 2x_2 + 3x_3) \end{pmatrix}$$

The Jacobian is given by

$$Jf(x_1, x_2, x_3) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \end{pmatrix} = \begin{pmatrix} x_2 x_3 & x_1 x_3 & x_1 x_2 \\ \cos(s) & 2 \cos(s) & 3 \cos(s) \end{pmatrix}$$

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

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

- 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 $f, g : D \rightarrow \mathbb{R}^m$ be differentiable in $x^0 \in D$, D open. Then $\alpha f(x^0) + \beta g(x^0)$, and $\alpha, \beta \in \mathbb{R}$ are differentiable in x^0 and we have

$$d(\alpha f + \beta g)(x^0) = \alpha df(x^0) + \beta dg(x^0)$$

$$J(\alpha f + \beta g)(x^0) = \alpha Jf(x^0) + \beta Jg(x^0)$$

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

For the differentials it holds

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

and analogously for the Jacobian matrix

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

Examples for the chain rule.

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

Then the composition

$$(f \circ h)(t) = f(h_1(t), \dots, h_n(t))$$

is differentiable in t_0 and we have for the derivative:

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

Directional derivative.

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

$$D_v f(x^0) := \lim_{t \rightarrow 0} \frac{f(x^0 + tv) - f(x^0)}{t}$$

is called the **directional derivative (Gateaux-derivative)** of $f(x)$ in the direction of v .

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

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

Remarks.

- For $v = e_i$ the directional derivative in the direction of v is given by the partial derivative with respect to x_i :

$$D_v f(x^0) = \frac{\partial f}{\partial x_i}(x^0)$$

- If v is a unit vector, i.e. $\|v\| = 1$, then the directional derivative $D_v f(x^0)$ describes the **slope** of $f(x)$ in the direction of v .
- If $f(x)$ is differentiable in x^0 , then all directional derivatives of $f(x)$ in x^0 exist. With $h(t) = x^0 + tv$ we have

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

This follows directly applying the chain rule.

Properties of the gradient.

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

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

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

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 $\text{grad } f(x^0)$ gives the direction of the steepest slope of $f(x)$ 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

$$|D_v f(x^0)| = |(\text{grad } f(x^0), v)| \leq \|\text{grad } f(x^0)\|_2$$

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

Curvilinear coordinates.

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

Then $x = \Phi(u)$ defines a **coordinate transformation** from the coordinates u to x .

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

$$x = r \cos \varphi$$

$$y = r \sin \varphi$$

with the **cartesian coordinates** $x = (x, y)$.

Calculation of the partial derivatives.

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

$$\Phi^{-1}(\Phi(u)) = u$$

$$J\Phi^{-1}(x) \cdot J\Phi(u) = I_n \quad (\text{chain rule})$$

$$J\Phi^{-1}(x) = (J\Phi(u))^{-1}$$

Let $\tilde{f} : V \rightarrow \mathbb{R}$ 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

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

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

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

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

Example: polar coordinates.

We consider polar coordinates

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

We calculate

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

and thus

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

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

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

The Jacobian–matrix is given by:

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

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

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

Chapter 1. Multivariate differential calculus

1.3 Mean value theorems and Taylor expansion

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

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

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

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

Proof: We set

$$h(t) := f(a + t(b - a))$$

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

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

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

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

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

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

It is

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

and

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

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

A mean value estimate for **vector-valued** functions.

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

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

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

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

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

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

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

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

Let

$$|\alpha| := \alpha_1 + \dots + \alpha_n \quad \alpha! := \alpha_1! \cdots \alpha_n!$$

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

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

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

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

The Taylor theorem.

Theorem: (Taylor)

Let $D \subset \mathbb{R}^n$ be open and convex. Let $f : D \rightarrow \mathbb{R}$ be a C^{m+1} -function and $x_0 \in D$. Then the Taylor-expansion holds true in $x \in D$

$$f(x) = T_m(x; x_0) + R_m(x; x_0)$$

$$T_m(x; x_0) = \sum_{|\alpha| \leq m} \frac{D^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha$$

$$R_m(x; x_0) = \sum_{|\alpha|=m+1} \frac{D^\alpha f(x_0 + \theta(x - x_0))}{\alpha!} (x - x_0)^\alpha$$

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

Notation: In the Taylor-expansion we denote $T_m(x; x_0)$ Taylor-polynom of degree m and $R_m(x; x_0)$ Lagrange-remainder.

Derivation of the Taylor expansion.

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

$$g(t) := f(x_0 + t(x - x_0))$$

$$g(0) = f(x_0)$$

$$g(1) = f(x)$$

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

$$g(1) = \underbrace{g(0)}_{T_1[1;0]} + \underbrace{g'(0) \cdot (1-0)}_{T_1[1;0]} + \underbrace{\frac{1}{2} g''(\xi) \cdot (1-0)^2}_{R_1[1;0]} \text{ for a } \xi \in (0, 1).$$

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

$$g'(0) = \left. \frac{d}{dt} f(x_1^0 + t(x_1 - x_1^0), x_2^0 + t(x_2 - x_2^0), \dots, x_n^0 + t(x_n - x_n^0)) \right|_{t=0}$$

$$= D_1 f(x_0) \cdot (x_1 - x_1^0) + \dots + D_n f(x_0) \cdot (x_n - x_n^0) = \text{grad } f(x_0) \cdot (x - x^0)$$

$$D^\alpha f(x_0) = \frac{\partial}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} f(x_0)$$

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$$

$$\alpha! = 1! \cdot 0! \cdot \dots \cdot 0! = 1 \quad (x - x^0)^\alpha = x_1 - x_1^0 \dots$$

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \quad D^\alpha f(x_0) = \frac{\partial}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} f(x_0)$$

$$\alpha! = 1 \quad (x - x^0)^\alpha = x_1 - x_1^0 \dots$$

Continuation of the derivation.

Calculation of $g''(0)$ gives

$$g''(0) = \frac{d^2}{dt^2} f(x_0 + t(x - x_0)) \Big|_{t=0} = \frac{d}{dt} \sum_{k=1}^n D_k f(x_0 + t(x - x_0)) (x_k - x_k^0) \Big|_{t=0}$$

$$\sum_{j=1}^n \sum_{i=1}^n (x_j - x_j^0) \frac{\partial^2}{\partial x_i \partial x_j} f(x_0 + t(x - x_0)) (x_i - x_i^0) \Big|_{t=0} = *$$
$$= D_{11} f(x_0) (x_1 - x_1^0)^2 + D_{21} f(x_0) (x_1 - x_1^0) (x_2 - x_2^0)$$

$$+ \dots + D_{ij} f(x_0) (x_i - x_i^0) (x_j - x_j^0) + \dots +$$

$$+ D_{n-1,n} f(x_0) (x_{n-1} - x_{n-1}^0) (x_n - x_n^0) + D_{nn} f(x_0) (x_n - x_n^0)^2$$

$$A_{ij} = e_{ij}$$

$$(*) = \sum_{i,j=1}^n (x_j - x_j^0) \frac{\partial^2 f(x_0)}{\partial x_j \partial x_i} (x_i - x_i^0) = (x - x^0)^T A (x - x^0)$$

$$= \sum_{|\alpha|=2} \frac{D^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha \quad (\text{exchange theorem of Schwarz!})$$

α_{ij}

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

Proof of the Taylor theorem.

The function

$$g(t) := f(x^0 + t(x - x^0))$$

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)

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

and

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

Examples for the Taylor–expansion.

- 1 Calculate the Taylor–polynom $T_2(x; x_0)$ 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 $T_2(x; x_0)$ 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 $T_2(x; x_0)$ in the form

$$T_2(x; x_0) = 4z(x + y - 2)$$

- 5 Details on extra slide.

Examples for the Taylor–expansion.

- 1 Calculate the Taylor–polynom $T_2(x; x_0)$ 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 $T_2(x; x_0)$ 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 $T_2(x; x_0)$ in the form

$$T_2(x; x_0) = 4z(x + y - 2)$$

- 5 Details on extra slide.

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

$$x^0 = (1, 2, 0)$$

$$f(x^0) = f(1, 2, 0) = 0$$

$$\text{grad } f = (y^2 \sin z, 2xy \sin z, xy^2 \cos z)$$

$$\text{grad } f(x^0) \cdot (x - x^0) = (0, 0, 4) \begin{pmatrix} x-1 \\ y-2 \\ z \end{pmatrix} = 4z$$

$$f(x, y, z) = \underbrace{0}_{T_0} + \underbrace{4z}_{T_1}$$

$$f \in C^2 \\ f \in C^\infty$$

$$\begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{pmatrix}_{x=x^0}$$

$$= \begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 4 \\ 4 & 4 & 0 \end{pmatrix}$$

$$f_{zx} = f_{xz} = y^2 \cos z$$

$$f_{zy} = f_{yz} = 2xy \cos z$$

$$(x - x^0) \begin{pmatrix} 1 \\ x - x^0 \end{pmatrix} =$$

$$= (x - x^0) \begin{pmatrix} 4z \\ 4z \\ 4(x-1) + 4(y-2) \end{pmatrix}$$

=

Remarks to the remainder of a Taylor–expansion.

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

$$R_m(x; x_0) = \sum_{|\alpha|=m+1} \frac{D^\alpha f(x_0 + \theta(x - x_0))}{\alpha!} (x - x_0)^\alpha$$

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

$$|R_m(x; x_0)| \leq \frac{n^{m+1}}{(m+1)!} C \|x - x_0\|_\infty^{m+1}$$

We conclude for the quality of the approximation of a C^{m+1} –function by the Taylor–polynom

$$f(x) = T_m(x; x_0) + O(\|x - x_0\|^{m+1})$$

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

$$f(x) = \underbrace{f(x^0) + \text{grad } f(x^0) \cdot (x - x^0)}_{T_1} + O(\|x - x^0\|^2).$$

The Hesse-matrix.

The matrix

$$Hf(x_0) := \begin{pmatrix} f_{x_1 x_1}(x_0) & \dots & f_{x_1 x_n}(x_0) \\ \vdots & & \vdots \\ f_{x_n x_1}(x_0) & \dots & f_{x_n x_n}(x_0) \end{pmatrix}$$

$$n=1 \\ Hf(x_0) = f''(x_0)$$

is called **Hesse-matrix** of f at x_0 .

Hesse-matrix = Jacobi-matrix of the gradient ∇f

The Taylor-expansion of a C^3 -function can be written as

$$f(x) = \underbrace{f(x_0)}_{T_0} + \underbrace{\text{grad } f(x_0)(x - x_0)}_{T_1} + \underbrace{\frac{1}{2}(x - x_0)^T Hf(x_0)(x - x_0)}_{T_2} + O(\|x - x_0\|^3)$$

The Hesse-matrix of a C^2 -function is symmetric.

from \mathbb{R}_2

Chapter 2. Applications of multivariate differential calculus

2.1 Extrem values of multivariate functions

Definition: Let $D \subset \mathbb{R}^n$, $f : D \rightarrow \mathbb{R}$ and $x^0 \in D$. Then at x^0 the function f has

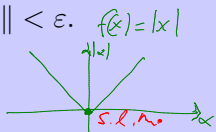
- a **global maximum** if $f(x) \leq f(x^0)$ for all $x \in D$.
- a **strict global maximum** if $f(x) < f(x^0)$ for all $x \in D$.
- a **local maximum** if there exists an $\varepsilon > 0$ such that

$$f(x) \leq f(x^0) \quad \text{for all } x \in D \text{ with } \|x - x^0\| < \varepsilon.$$

- a **strict local maximum** if there exists an $\varepsilon > 0$ such that

$$f(x) < f(x^0) \quad \text{for all } x \in D \text{ with } \|x - x^0\| < \varepsilon.$$

Analogously we define the different forms of minima.



Necessary conditions for local extrem values.

Theorem: If a C^1 -function $f(x)$ has a local extrem value (minimum or maximum) at $x^0 \in D^0$, then

$$\text{grad } f(x^0) = 0 \in \mathbb{R}^n$$

Proof: For an arbitrary $v \in \mathbb{R}^n$, $v \neq 0$ the function

$$\varphi(t) := f(x^0 + tv)$$

$x = x^0 + tv$

is differentiable in a neighborhood of $t^0 = 0$.

$\varphi(t)$ has a local extrem value at $t^0 = 0$. We conclude:

Charakter

$$0 = \varphi'(0) = \text{grad } f(x^0) \cdot v \neq 0$$

*Max. at $t=0$
Min.*

$$\Rightarrow \text{grad } f(x_0) \cdot v = 0$$

$\forall v \neq 0$

Since this holds true for all $v \neq 0$ we obtain

$$\text{grad } f(x^0) = (0, \dots, 0)^T$$

$$v = \text{grad } f(x_0)$$

$$|\text{grad } f(x_0)|^2 = 0$$

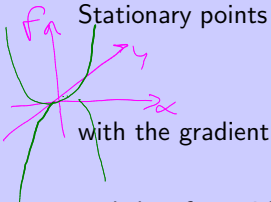
$f(x) = |x|$, $f \notin C^1$
 $x^0 = 0$ global Minimum
but $f'(0) \nexists$

Remarks to local extrem values.

$$f: D \subset \mathbb{R}^n \rightarrow \mathbb{R} \quad \text{grad } f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \text{rectn. fields}$$

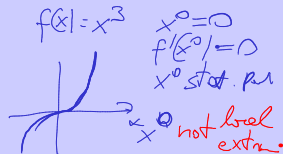
Bemerkungen:

- Typically the condition $\text{grad } f(x^0) = 0$ gives a **non-linear** system of n **equations** for n unknowns for the calculation of $x = x^0$.
- The points $x^0 \in D^0$ with $\text{grad } f(x^0) = 0$ are called **stationary points** of f . Stationary points are **not** necessarily local extrem values. As an example take



$$f(x, y) := x^2 - y^2$$

$$\text{grad } f(x, y) = 2(x, -y)$$



and therefore with the only stationary point $x^0 = (0, 0)^T$. However, the point x^0 is a **saddel point** of f , i.e. in every neighborhood of x^0 there exist two points x^1 and x^2 with

$$f(x^1) < f(x^0) < f(x^2).$$

Classification of stationary points.

Theorem: Let $f(x)$ be a C^2 -function on D^0 and let $x^0 \in D^0$ be a stationary point of $f(x)$, i.e. $\text{grad } f(x^0) = 0$.

a) necessary condition

If x^0 is a local extrem value of f , then:

x^0 local minimum $\Rightarrow H f(x^0)$ positiv semidefinit

x^0 local maximum $\Rightarrow H f(x^0)$ negativ semidefinit

b) sufficient condition

If $H f(x^0)$ is positiv definit (negativ definit) then x^0 is a strict local minimum (maximum) of f .

If $H f(x^0)$ is indefinit then x^0 is a saddle point, i.e. in every neighborhood of x^0 there exist points x^1 and x^2 with $f(x^1) < f(x^0) < f(x^2)$.

$$f \in C^2, \quad \text{grad} f(x_0) = 0$$

x_0 strict local minimum $\not\Rightarrow$ $Hf(x_0)$ is positive definite

$$f(x) = x^4$$



$x^0 = 0$ stat. point $f'(0) = 0$
 x^0 strict local mini (global)

$$f''(x^0) = f''(0) = 0$$

Proof of the theorem, part a).

Let x^0 be a local minimum. For $v \neq 0$ and $\varepsilon > 0$ sufficiently small we conclude from the Taylor-expansion

$$f(x) = f(x^0) + \underbrace{\text{grad} f(x^0)(x-x^0)}_{\cancel{0}} + \underbrace{\frac{1}{2}(x-x^0)^T H f(x^0 + \theta(x-x^0))(x-x^0)}_{R_1} + \dots$$

$x = x^0 + \varepsilon v$ $x - x^0 = \varepsilon v$

$$f(x^0 + \varepsilon v) - f(x^0) = \frac{1}{2}(\varepsilon v)^T H f(x^0 + \theta \varepsilon v)(\varepsilon v) \geq 0 \quad (1)$$

with $\theta = \theta(\varepsilon, v) \in (0, 1)$.

The gradient in the Taylor expansion $\text{grad} f(x^0) = 0$ vanishes since x^0 is stationary.

From (1) it follows

$$v^T H f(x^0 + \theta \varepsilon v) v \geq 0 \quad (2)$$

Since f is a \mathcal{C}^2 -function, the Hesse-matrix is a continuous map. In the limit $\varepsilon \rightarrow 0$ we conclude from (2),

$$v^T H f(x^0) v \geq 0 \quad \forall v \in \mathbb{R}^n, v \neq 0$$

i.e. $H f(x^0)$ is positiv semidefinit.

Proof of the theorem, part b).

If $Hf(x^0)$ is positiv definit, then $Hf(x)$ is positiv definit in a sufficiently small neighborhood $x \in K_\varepsilon(x^0) \subset D$ around x^0 . This follows from the continuity of the second partial derivatives.

For $x \in K_\varepsilon(x^0)$, $x \neq x^0$ we have

$$f(x) - f(x^0) = \underbrace{\frac{1}{2}(x - x^0)^T Hf(x^0 + \theta(x - x^0))}_{R_1} (x - x^0)$$

Handwritten note: $\text{grad } f(x^0) = 0$

with $\theta \in (0, 1)$, i.e. f has a strict local minimum at x^0 .

If $Hf(x^0)$ is indefinit, then there exist Eigenvectors v w for Eigenvalues of $Hf(x^0)$ with opposite sign with

$$v^T Hf(x^0) v > 0$$

$$w^T Hf(x^0) w < 0$$

and thus x^0 is a saddle point.

Remarks.

- A stationary point x^0 with $\det Hf(x^0) = 0$ is called **degenerate**. The Hesse-matrix has an Eigenvalue $\lambda = 0$.
- If x^0 is **not** degenerate, then there exist 3 cases for the Eigenvalues of $Hf(x^0)$:

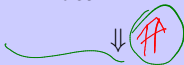
all Eigenvalues are strictly positive $\Rightarrow x^0$ is a strict local minimum

all Eigenvalues are strictly negative $\Rightarrow x^0$ is a strict local maximum

there are strictly positive and negative Eigenvalues $\Rightarrow x^0$ saddle point

- The following implications are true (**but not the inverse**)

x^0 local minimum \Leftarrow x^0 strict local minimum



$Hf(x^0)$ positiv semidefinit \Leftarrow $Hf(x^0)$ positiv definit

$f(x) = x^3$
 $f'(0) = 0$ semidef.
no local m.!!

Further remarks.

- If f is a C^3 -function, x^0 a stationary point of f and $Hf(x^0)$ positiv definit Then the following estimate is true:

$$\lambda_{\min} \|x - x^0\|^2 \geq (x - x^0)^T Hf(x^0) (x - x^0) \geq \lambda_{\min} \cdot \|x - x^0\|^2$$

where λ_{\min} denoted the **smallest** Eigenvalue of the Hesse-matrix.

Using the Taylor theorem we obtain:

$$f(x) - f(x^0) \geq \frac{1}{2} \lambda_{\min} \|x - x^0\|^2 + R_2(x; x^0)$$

$$\geq \|x - x^0\|^2 \left(\frac{\lambda_{\min}}{2} - C \|x - x^0\| \right)$$

with an appropriate constant $C > 0$.

The function f grows at least quadratically around x^0 .

Example .

We consider the function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ $f \in C^2(\mathbb{C})$

$$f(x, y) := y^2(x - 1) + x^2(x + 1)$$

$$y^2 + x(3x + 2) = 0$$

$$2y(x - 1) = 0$$

$$x = 1 \vee y = 0$$

and look for stationary points :

*2 non-linear equations
2 unknowns*

$$\text{grad } f(x, y) = (y^2 + x(3x + 2), 2y(x - 1))^T$$

no sol!

$$y^2 + 5 = 0 \quad \downarrow \quad x = 0 \vee x = -\frac{2}{3}$$

The condition $\text{grad } f(x, y) = 0$ gives two stationary points

$$x^0 = (0, 0)^T \quad \text{und} \quad x^1 = (-2/3, 0)^T.$$

The related Hesse-matrices of f at x^0 and x^1 are

$$Hf(x^0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{and} \quad Hf(x^1) = \begin{pmatrix} -2 & 0 \\ 0 & -10/3 \end{pmatrix}$$

The matrix $Hf(x^0)$ is indefinit, therefore x^0 is a saddle point. $Hf(x^1)$ is negativ definit and thus x^1 is a strict local ein strenges maximum of f .

2.2 Implicitly defined functions

Aim: study the set of solutions of the system of *non-linear* equations of the form

$$g(x) = 0$$

with $g : D \rightarrow \mathbb{R}^m$, $D \subset \mathbb{R}^n$. I.e. we consider m equations for n unknowns with

$$m < n.$$

Thus: there are **less** equations than unknowns.

We call such a system of equations **underdetermined** and the set of solutions $G \subset \mathbb{R}^n$ contains typically *infinitely* many points.

Solvability of (non-linear) equations.

Question: can we **solve** the system $g(x) = 0$ with respect to certain unknowns, i.e. with respect to the last m variables x_{n-m+1}, \dots, x_n ?

In other words: is there a function $f(x_1, \dots, x_{n-m})$ with

$$g(x) = 0 \iff (x_{n-m+1}, \dots, x_n)^T = f(x_1, \dots, x_{n-m})$$

Terminology: "solve" means express the last m variables by the first $n - m$ variables?

Other question: with respect to which m variables can we solve the system? Is the solution possible *globally* on the domain of definition D ? Or only *locally* on a subdomain $\tilde{D} \subset D$?

Geometrical interpretation: The set of solution G of $g(x) = 0$ can be expressed (at least locally) as graph of a function $f : \mathbb{R}^{n-m} \rightarrow \mathbb{R}^m$.

Example.

The equation for a circle

$$g(x, y) = x^2 + y^2 - r^2 = 0 \quad \text{mit } r > 0$$

defines an **underdetermined** non-linear system of equations since we have **two** unknowns (x, y) , but only **one** scalar equation.

The equation for the circle can be solved **locally** and defines the four functions :

$$y = \sqrt{r^2 - x^2}, \quad -r \leq x \leq r$$

$$y = -\sqrt{r^2 - x^2}, \quad -r \leq x \leq r$$

$$x = \sqrt{r^2 - y^2}, \quad -r \leq y \leq r$$

$$x = -\sqrt{r^2 - y^2}, \quad -r \leq y \leq r$$

Example.

Let g be an affin-linear function, i.e. g has the form

$$g(x) = Cx + b \quad \text{for } C \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$$

We split the variables x into two vectors

$$x^{(1)} = (x_1, \dots, x_{n-m})^T \in \mathbb{R}^{n-m} \quad \text{and} \quad x^{(2)} = (x_{n-m+1}, \dots, x_n)^T \in \mathbb{R}^m$$

Splitting of the matrix $C = [B, A]$ gives the form

$$g(x) = Bx^{(1)} + Ax^{(2)} + b$$

with $B \in \mathbb{R}^{m \times (n-m)}$, $A \in \mathbb{R}^{m \times m}$.

The system of equations $g(x) = 0$ can be solved (uniquely) with respect to the variables $x^{(2)}$, if A is regular. Then

$$g(x) = 0 \quad \iff \quad x^{(2)} = -A^{-1}(Bx^{(1)} + b) = f(x^{(1)})$$

Continuation of the example.

Question: How can we write the matrix A as dependent of g ?

From the equation

$$g(x) = Bx^{(1)} + Ax^{(2)} + b$$

we see that

$$A = \frac{\partial g}{\partial x^{(2)}}(x^{(1)}, x^{(2)})$$

holds, i.e. A is the Jacobian of the map

$$x^{(2)} \rightarrow g(x^{(1)}, x^{(2)})$$

for fixed $x^{(1)}$!

We conclude: Solvability is given if the Jacobian is regular (invertible).

Satz über implizite Funktionen.

Satz: Sei $g : D \rightarrow \mathbb{R}^m$ eine \mathcal{C}^1 -Funktion, $D \subset \mathbb{R}^n$ offen. Die Variablen in D seien (x, y) mit $x \in \mathbb{R}^{n-m}$ und $y \in \mathbb{R}^m$. Der Punkt $(x^0, y^0) \in D$ sei eine Lösung von $g(x^0, y^0) = 0$.

Falls die Jacobi-Matrix

$$\frac{\partial g}{\partial y}(x^0, y^0) := \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(x^0, y^0) & \dots & \frac{\partial g_1}{\partial y_m}(x^0, y^0) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial y_1}(x^0, y^0) & \dots & \frac{\partial g_m}{\partial y_m}(x^0, y^0) \end{pmatrix}$$

regulär ist, so gibt es Umgebungen U von x^0 und V von y^0 , $U \times V \subset D$ und eine eindeutig bestimmte stetig differenzierbare Funktion $f : U \rightarrow V$ mit

$$f(x^0) = y^0 \quad \text{und} \quad g(x, f(x)) = 0 \quad \text{für alle } x \in U$$

und

$$Jf(x) = - \left(\frac{\partial g}{\partial y}(x, f(x)) \right)^{-1} \left(\frac{\partial g}{\partial x}(x, f(x)) \right)$$

Beispiel.

Für die Kreisgleichung $g(x, y) = x^2 + y^2 - r^2 = 0$, $r > 0$ findet man im Punkt $(x^0, y^0) = (0, r)$

$$\frac{\partial g}{\partial x}(0, r) = 0, \quad \frac{\partial g}{\partial y}(0, r) = 2r \neq 0$$

Man kann also in einer Umgebung von $(0, r)$ die Kreisgleichung nach y auflösen:

$$f(x) = \sqrt{r^2 - x^2}$$

Die Ableitung $f'(x)$ kann man durch **implizite Differentiation** berechnen:

$$g(x, y(x)) = 0 \quad \implies \quad g_x(x, y(x)) + g_y(x, y(x))y'(x) = 0$$

Also

$$2x + 2y(x)y'(x) = 0 \quad \implies \quad y'(x) = f'(x) = -\frac{x}{y(x)}$$

Ein weiteres Beispiel.

Betrachte die Gleichung $g(x, y) = e^{y-x} + 3y + x^2 - 1 = 0$.

Es gilt

$$\frac{\partial g}{\partial y}(x, y) = e^{y-x} + 3 > 0 \quad \text{für alle } x \in \mathbb{R}.$$

Die Gleichung ist also für jedes $x \in \mathbb{R}$ nach $y =: f(x)$ auflösbar und $f(x)$ ist eine stetig differenzierbare Funktion. Implizite Differentiation liefert

$$e^{y-x}(y' - 1) + 3y' + 2x = 0 \quad \implies \quad y' = \frac{e^{y-x} - 2x}{e^{y-x} + 3}$$

Erneute Differentiation liefert

$$e^{y-x}y'' + e^{y-x}(y' - 1)^2 + 3y'' + 2 = 0 \quad \implies \quad y' = -\frac{2 + e^{y-x}(y' - 1)^2}{e^{y-x} + 3}$$

Aber: Explizites Auflösen nach y (mit Hilfe elementarer Funktionen) ist in diesem Fall nicht möglich!

Allgemeine Bemerkung.

Implizites Differenzieren einer durch

$$g(x, y) = 0, \quad \frac{\partial g}{\partial y} \neq 0$$

implizit definierten Funktion $y = f(x)$, mit $x, y \in \mathbb{R}$, ergibt

$$f'(x) = -\frac{g_x}{g_y}$$

$$f''(x) = -\frac{g_{xx}g_y^2 - 2g_{xy}g_xg_y + g_{yy}g_x^2}{g_y^3}$$

Daher ist der Punkt x^0 ein **stationärer** Punkt von $f(x)$, falls gilt

$$g(x^0, y^0) = g_x(x^0, y^0) = 0 \quad \text{und} \quad g_y(x^0, y^0) \neq 0$$

Weiter ist x^0 ein **lokales Maximum** (bzw. **Minimum**), falls

$$\frac{g_{xx}(x^0, y^0)}{g_y(x^0, y^0)} > 0 \quad \left(\text{bzw.} \quad \frac{g_{xx}(x^0, y^0)}{g_y(x^0, y^0)} < 0 \right)$$

Implizite Darstellung ebener Kurven.

Betrachte die Lösungsmenge einer skalaren Gleichungen

$$g(x, y) = 0$$

Falls gilt

$$\text{grad } g = (g_x, g_y) \neq 0$$

so definiert $g(x, y)$ lokal eine Funktion $y = f(x)$ oder $x = \bar{f}(y)$.

Definition: Ein Lösungspunkt (x^0, y^0) der Gleichung $g(x, y) = 0$ mit

- $\text{grad } g(x^0, y^0) \neq 0$ heißt **regulärer** Punkt,
- $\text{grad } g(x^0, y^0) = 0$ heißt **singulärer** Punkt.

Beispiel: Betrachte wieder die Kreisgleichung

$$g(x, y) = x^2 + y^2 - r = 0 \quad \text{mit } r > 0.$$

Auf der Kreislinie liegen **keine** singulären Punkte!

Horizontale und vertikale Tangenten.

Bemerkung:

a) Gilt für einen regulären Punkt (x^0, y^0)

$$g_x(x^0) = 0 \quad \text{und} \quad g_y(x^0) \neq 0$$

so besitzt die Lösungskurve eine **horizontale Tangente** in x^0 .

b) Gilt für einen regulären Punkt (x^0, y^0)

$$g_x(x^0) \neq 0 \quad \text{und} \quad g_y(x^0) = 0$$

so besitzt die Lösungskurve eine **vertikale Tangente** in x^0 .

c) Ist x^0 ein **singulärer Punkt**, so wird die Lösungsmenge bei x^0 "in zweiter Näherung" durch folgende **quadratische Gleichung** approximiert.

$$g_{xx}(x^0)(x - x^0)^2 + 2g_{xy}(x^0)(x - x^0)(y - y^0) + g_{yy}(x^0)(y - y^0)^2 = 0$$

Bemerkungen.

Wegen c) erhält man für $g_{xx}, g_{xy}, g_{yy} \neq 0$:

$\det Hg(x^0) > 0$: x^0 ist ein **isolierter Punkt** der Lösungsmenge

$\det Hg(x^0) < 0$: x^0 ist ein **Doppelpunkt**

$\det Hg(x^0) = 0$: x^0 ist ein **Rückkehrpunkt** bzw. eine **Spitze**

Geometrische Interpretation:

- Gilt $\det Hg(x^0) > 0$, so sind beide Eigenwerte von $Hg(x^0)$ entweder strikt positiv oder strikt negativ, d.h. x^0 ist ein strenges lokales **Minimum** oder **Maximum** von $g(x)$.
- Gilt $\det Hg(x^0) < 0$, so haben die beiden Eigenwerte von $Hg(x^0)$ ein unterschiedliches Vorzeichen, d.h. x^0 ist ein **Sattelpunkt** von $g(x)$.
- Gilt $\det Hg(x^0) = 0$, so ist der stationäre Punkt x^0 von $g(x)$ **ausgeartet**.

Beispiel 1.

Betrachte den singulären Punkt $x^0 = 0$ der impliziten Gleichung

$$g(x, y) = y^2(x - 1) + x^2(x - 2) = 0$$

Berechnung der partiellen Ableitungen bis zur Ordnung 2:

$$g_x = y^2 + 3x^2 - 4x$$

$$g_y = 2y(x - 1)$$

$$g_{xx} = 6x - 4$$

$$g_{xy} = 2y$$

$$g_{yy} = 2(x - 1)$$

$$Hg(0) = \begin{pmatrix} -4 & 0 \\ 0 & -2 \end{pmatrix}$$

Also ist $x^0 = 0$ ein **isolierter Punkt**.

Beispiel 2.

Betrachte den singulären Punkt $x^0 = 0$ der impliziten Gleichung

$$g(x, y) = y^2(x - 1) + x^2(x + q^2) = 0$$

Berechnung der partiellen Ableitungen bis zur Ordnung 2:

$$g_x = y^2 + 3x^2 + 2xq^2$$

$$g_y = 2y(x - 1)$$

$$g_{xx} = 6x + 2q^2$$

$$g_{xy} = 2y$$

$$g_{yy} = 2(x - 1)$$

$$Hg(0) = \begin{pmatrix} 2q^2 & 0 \\ 0 & -2 \end{pmatrix}$$

Also ist $x^0 = 0$ für $q \neq 0$ ein **Doppelpunkt**.

Beispiel 3.

Betrachte den singulären Punkt $x^0 = 0$ der impliziten Gleichung

$$g(x, y) = y^2(x - 1) + x^3 = 0$$

Berechnung der partiellen Ableitungen bis zur Ordnung 2:

$$g_x = y^2 + 3x^2$$

$$g_y = 2y(x - 1)$$

$$g_{xx} = 6x$$

$$g_{xy} = 2y$$

$$g_{yy} = 2(x - 1)$$

$$Hg(0) = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}$$

Also ist $x^0 = 0$ eine **Spitze** (bzw. ein **Rückkehrpunkt**).

Implizite Darstellung von Flächen.

- Die Lösungsmenge einer skalaren Gleichung $g(x, y, z) = 0$ ist für $\text{grad } g \neq 0$ lokal eine **Fläche** im \mathbb{R}^3 .
- Für die **Tangentialebene** in $x^0 = (x^0, y^0, z^0)^T$ mit $g(x^0) = 0$ und $\text{grad } g(x^0) \neq 0^T$ bekommen wir für $\Delta x^0 = x - x^0$ mit Taylor-Entwicklung

$$\text{grad } g \cdot \Delta x^0 = g_x(x^0)(x - x^0) + g_y(x^0)(y - y^0) + g_z(x^0)(z - z_0) = 0$$

d.h. der Gradient steht senkrecht auf der Fläche $g(x, y, z) = 0$.

- Ist zum Beispiel $g_z(x^0) \neq 0$, so gibt es lokal bei x^0 eine Darstellung der Form

$$z = f(x, y)$$

und für die **partielle Ableitungen** von $f(x, y)$ bekommt man

$$\text{grad } f(x, y) = (f_x, f_y) = -\frac{1}{g_z}(g_x, g_y) = \left(-\frac{g_x}{g_z}, \frac{g_y}{g_z} \right)$$

mit dem Satz über implizite Funktionen.

Das Umkehrproblem.

Frage: Lässt sich ein vorgegebenes Gleichungssystem

$$y = f(x)$$

mit $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen, nach x auflösen, also **invertieren**?

Satz: (Umkehrsatz)

Sei $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen, eine \mathcal{C}^1 -Funktion. Ist für ein $x^0 \in D$ die Jacobi-Matrix $Jf(x^0)$ regulär, so gibt es Umgebungen U und V von x^0 und $y^0 = f(x^0)$, so dass f den Bereich U **bijektiv** auf V abbildet.

Die Umkehrfunktion $f^{-1} : V \rightarrow U$ ist ebenfalls eine \mathcal{C}^1 -Funktion und es gilt für alle $x \in U$:

$$Jf^{-1}(y) = (Jf(x))^{-1}, \quad y = f(x)$$

Bemerkung: Man nennt dann f lokal einen \mathcal{C}^1 -Diffeomorphismus.