Funktionen in mehreren Veränderlichen

Ziel: Wir wollen „schöne" Funktionen

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

verstehen. Manchmal ist f nicht auf ganz $\mathbb{R}^n$ definiert sondern nur auf einer Teilmenge $U \subseteq \mathbb{R}^n$.

Für so eine Funktion $f : U \to \mathbb{R}$ suchen wir insbesondere:

1) Näherungsformeln für das Verhalten in der Nähe eines Punktes $\underline{u}_0 \in U$,

2) Maxima und Minima sowie die Stellen $\underline{u} \in U$, an denen sie angenommen werden,

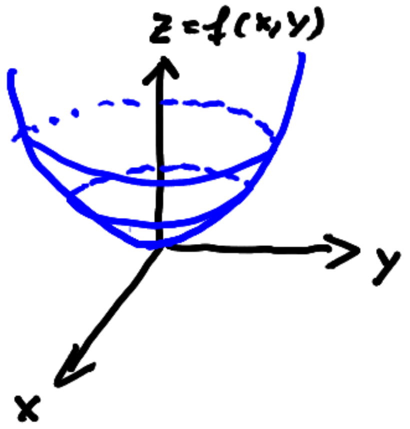
3) Den „Rauminhalt unter dem Graph", also das höherdimensionale Integral.

Bsp1:
$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$$
$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto x^2 + y^2$$

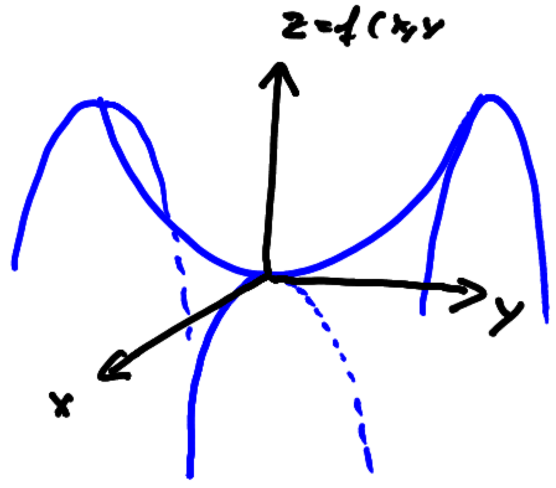
Rotationsparaboloid

$$g : \mathbb{R}^2 \longrightarrow \mathbb{R}$$
$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto y^2 - x^2$$

Sattelfläche



$z = f(x,y)$



$z = g(x,y)$

Wir bilden die <u>partiellen Ableitungen</u>

$$\frac{df}{dx} = 2x \qquad \frac{df}{dy} = 2y$$

$$\frac{dg}{dx} = -2x \qquad \frac{dg}{dy} = 2y$$

Die partiellen Ableitungen ergeben sich, wenn wir die anderen Variablen konstant halten. Diese formalen Ableitungen hatten wir schon im letzten Semester bei der Fehlerrechnung nach Gauß kennen gelernt.

Def: $f: U \to \mathbb{R} \qquad U \subseteq \mathbb{R}^n$

Die $1 \times n$-Matrix

$$D_{\underline{u}} f := \left(\frac{df}{dx_1}\Big|_{\underline{u}} \; \cdots \; \frac{df}{dx_n}\Big|_{\underline{u}} \right)$$

heißt Ableitung von f an der Stelle $\underline{u} \in U$.

Bsp 2

$$D_{\binom{x}{y}} f = (2x \quad 2y) \qquad f, g \text{ wie in Bsp 1.}$$

$$D_{\binom{x}{y}} g = (-2x \quad 2y)$$

Erinnerung: Für eine differenzierbare / analytische Funktion $f(x)$ in einer Variablen gilt die Näherung

$$f(x_0 + h) \approx f(x_0) + f'(x_0)\, h$$

Wir hatten die Gültigkeit dieser Näherung als Definition der Differenzierbarkeit genommen.

Def: $f: U \to \mathbb{R}$, $U \in \mathbb{R}^n$ heißt differenzierbar an der Stelle $\underline{u} \in U$, wenn die partiellen Ableitungen $\frac{df}{dx_i}$ an der Stelle $\underline{u} \in U$ existieren und die Näherung

$$f(\underline{u}+\underline{h}) = f(\underline{u}) + (D_{\underline{u}} f)(\underline{h}) + \text{Terme höherer Ordnung}$$

gilt.

Interpretation: Sei f differenzierbar an der Stelle $u_0 \in U$. Die Abbildung

$$\mathbb{R}^n \longrightarrow \mathbb{R}$$

$$\underline{u} \mapsto f(\underline{u}_0) + (D_{\underline{u}_0} f)(\overbrace{\underline{u}-\underline{u}_0}^{\underline{h}})$$

hat als Graph eine „n-dimensionale Ebene":

- $n=1$: Graph ist eine Gerade
- $n=2$: Graph ist eine Ebene

$\vdots$

Diese Ebene berührt den Graphen

$$\{(\underline{u}, f(\underline{u})) \in \mathbb{R}^{n+1} \mid \underline{u} \in U\}$$

im Punkt $(y_0, f(y_0))$. Es handelt sich gerade um die Tangentialebene an den Graphen von f. Es gilt:

$$\boxed{\text{f ist differenzierbar an der Stelle y_0 genau dann, wenn der Graph von f bei $(y_0, f(y_0))$ eine Tangentialebene hat.}}$$

<u>Bsp</u>:

$$g(x,y) := \begin{cases} \dfrac{xy(x+y)}{x^2+y^2} & x^2+y^2 \neq 0 \\ 0 & x=y=0 \end{cases}$$

1) $(x,y) \to 0 \Rightarrow g(x,y) \Rightarrow 0$

also g ist stetig

2) $g(t,0) = 0$ für alle t

$\Rightarrow$ Die Gerade $t \to (t,0,0)$ ist tangential im Punkte $(0,0)$

3) $g(0,t) = 0$ für alle t

$\Rightarrow$ Die Gerade $t \to (0,t,0)$ ist tangential im Punkte $(0,0)$

4) $g(t,t) = \frac{t^2\, 2t}{2t^2} = t$ für alle t

$\Rightarrow$ Die Gerade $t \to (t,t,0)$ ist tangential im Punkt $(0,0)$

<u>Aber</u>: Diese drei Geraden liegen nicht in einer Ebene.

<u>Also</u>: g ist <u>nicht</u> differenzierbar im Punkt $(0,0)$.

<u>Fakt</u>: Existieren in einer ganzen Umgebung von $\underline{u}_0 \in U$ alle partiellen Ableitungen $\left.\frac{df}{dx_i}\right|_{\underline{u}}$ und hängen sie stetig vom Punkt $\underline{u}$ ab, so ist f differenzierbar in dieser Umgebung.

<u>Bem</u>: Die partiellen Ableitungen lassen sich gut ausrechnen, weil wir das Ableiten von Funktionen in einer Variablen mit den Ableitungsregeln gut beherrschen.

<u>Bsp</u>: $h(x,y) = \frac{\sin(x)}{y^2} = y^{-2}\sin(x)$

$$\frac{dh}{dx} = \frac{\cos(x)}{y^2} \qquad \frac{dh}{dy} = -2y^{-3}\sin(x)$$

Die Richtungsableitungen

Sei $\underline{u}_0 \in U \subseteq \mathbb{R}^n$ und $f: U \to \mathbb{R}$.
Sei ferner ein Einheitsvektor $\underline{h} \in \mathbb{R}^n$ gegeben. Die Funktion

$$\gamma : \mathbb{R} \longrightarrow \mathbb{R}^n$$
$$t \longmapsto \underline{u}_0 + t\underline{h}$$

beschreibt die gleichförmige Bewegung eines Teilchens auf einer Geraden durch den Punkt $\underline{u}_0$. Deuten wir f z.B. als Temperatur, so ist

$$f(\gamma(t))$$

die Temperatur, die das Teilchen zur Zeit t erfährt (eben weil es zu dieser Zeit am Ort $\underline{u} = \gamma(t)$ ist und dort die Temperatur $f(\underline{u}) = f(\gamma(t))$ herrscht.

Die Ableitung $\frac{d}{dt} f(\gamma(t)) \Big|_{t=0}$ heißt

Richtungsableitung von f an der Stelle $\underline{u}_0$ in Richtung $\underline{h}$:

$$\left.\frac{df}{d\underline{h}}\right|_{\underline{u}_0} := \left.\frac{d}{dt} f(\underline{u}_0 + t\underline{h})\right|_{t=0}$$

Beob: Die partiellen Ableitungen lassen sich als Richtungsableitungen nach den Standardrichtungen deuten:

$$\underline{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i^{te} \text{ Stelle}$$

$$\frac{df}{dx_i} = \frac{df}{d\underline{e}_i}$$

denn: $\underline{u}_0 = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$

$$\left.\frac{df}{d\underline{e}_i}\right|_{\underline{u}_0} = \frac{d}{dt} f(\underline{u}_0 + t\underline{e}_i)$$

$$= \frac{d}{dt} f(u_1, \dots, u_i + t, \dots, u_n)$$

$$= \lim_{h \to 0} \frac{f(u_1, \dots, u_i + h, \dots, u_n) - f(u_1, \dots, u_n)}{h}$$

$$= \left.\frac{df}{dx_i}\right|_{\underline{u}_0}$$

Beob: Für eine differenzierbare Abbildung

$$f : U \to \mathbb{R} \qquad U \subseteq \mathbb{R}^n$$

enthält die Ableitung (das totale Differential) $D_{\underline{y}_0} f$ die vollständige Information über <u>alle</u> Richtungsableitungen:

$$\left.\frac{df}{d\underline{h}}\right|_{\underline{y}_0} = \left(D_{\underline{y}_0} f\right) \underline{h}$$

Denn:

$$f(\underline{y}_0 + t\underline{h}) \approx f(\underline{y}_0) + \left.\frac{df}{d\underline{h}}\right|_{\underline{y}_0} \cdot t + \text{höhere Ordnung in } t$$

$$f(\underline{y}_0 + t\underline{h}) \approx f(\underline{y}_0) + t\left(D_{\underline{y}_0} f\right) \underline{h} + \text{höhere Ordnung in } t \text{ und } \underline{h}$$

Die Näherungen erster Ordnung stimmen überein und es folgt für $t = 1$

$$\left(D_{\underline{y}_0} f\right) \underline{h} = \left.\frac{df}{d\underline{h}}\right|_{\underline{y}_0} \qquad \boxtimes$$

Kettenregel I

$U \subseteq \mathbb{R}^n$, $\quad I \subseteq \mathbb{R}$ Intervall

$g : U \to \mathbb{R}$ differenzierbar

$f_1, \dots, f_n : I \to \mathbb{R}$ differenzierbar

<u>Annahme</u>: $\begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix} \in U \quad \forall\, t \in I$

$$\rightsquigarrow \quad f : I \to U$$
$$t \mapsto \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}$$

<u>Regel</u>: Dann ist $g \circ f : I \to \mathbb{R}$ differenzierbar, und es ist

$$\frac{d}{dt}(g \circ f) = \left(D_{f(t)} g\right)\left(D_t f\right) \text{''}$$

$$\text{''} \quad = \sum_{i=1}^{n} \left.\frac{dg}{dx_i}\right|_{f(t)} \frac{df}{dt}$$

<u>Notation</u>: Oft findet man:

$$\frac{dy}{dt} = \frac{dg}{dx_1}\frac{dx_1}{dt} + \dots + \frac{dg}{dx_n}\frac{dx_n}{dt}$$

Hier liegt die Auffassung zugrunde, daß g von $x_1, \dots, x_n$ abhängt und jedes x_i selbst eine Funktion von t ist.

<u>Bsp:</u> $g(x,y) = x^2 + y^2 \qquad D_{\binom{x}{y}} g = \begin{pmatrix} 2x & 2y \end{pmatrix}$

$$x(t) = \sin(t) \qquad \frac{dx}{dt} = \cos(t)$$

$$y(t) = \cos(t) \qquad \frac{dy}{dt} = -\sin(t)$$

$$\frac{dg}{dt} = 2x\Big|_{x=\sin(t)} \cdot \frac{dx}{dt} + 2y\Big|_{y=\cos(t)} \frac{dy}{dt}$$

$$= 2\sin(t)\cos(t) + 2\cos(t)(-\sin(t))$$

$$= 0$$

$$g(t) = \sin(t)^2 + \cos(t)^2 = 1$$

<u>also</u> $\frac{dg}{dt} = 0$ ✓

Der Gradient

Def: Sei $f: U \to \mathbb{R}$, $U \subseteq \mathbb{R}^n$ differenzierbar an der Stelle $\underline{u} \in U$. Dann heißt

$$\nabla_{\underline{u}} f := (D_{\underline{u}} f)^T \in \mathbb{R}^n$$

der Gradient von f an der Stelle $\underline{u}$.

Fakt: Ein Gradient $\neq 0$ zeigt in die Richtung $\underline{h} := \frac{\nabla_{\underline{u}} f}{|\nabla_{\underline{u}} f|}$ mit der größten Richtungsableitung $\left.\frac{df}{d\underline{h}}\right|_{\underline{u}}$.

D.h.: Der Gradient zeigt die Richtung, entlang der die Funktion f am schnellsten wächst.

Fakt: Ist $\nabla_{\underline{u}} f \neq 0$, so sind die Richtungen $\underline{h}$ orthogonal zum Gradienten gerade die Richtungen, entlang derer f ungefähr konstant bleibt (nur Terme der Ordnung ≥ 2

bleiben in der Näherungsformel stehen).

<u>Satz</u>: Hat $f : U \to \mathbb{R}$, $U \subseteq \mathbb{R}^n$ ein lokales Maximum / Minimum an der Stelle $\underline{y} \in U$, so ist

$$\nabla_{\underline{y}} f = 0 \qquad D_{\underline{y}} f = 0$$

(<u>Notwendiges</u> <u>Kriterium</u> für lokale Extrema.)

$\mathbb{R}^m$-wertige Funktionen

$$U \subseteq \mathbb{R}^n \qquad V \subseteq \mathbb{R}^m$$

$$f : U \to V$$
$$\underline{u} \mapsto f(\underline{u}) = \begin{pmatrix} f_1(\underline{u}) \\ \vdots \\ f_m(\underline{u}) \end{pmatrix}$$

Def: $f_i : U \to \mathbb{R}$ heißt <u>Koordinatenfunktion</u> von $f : U \to \mathbb{R}^m$.

Die Funktion $f : U \to \mathbb{R}^m$ heißt differenzierbar an der Stelle $\underline{u}_0 \in U$, wenn alle Koordinatenfunktionen $f_i : U \to \mathbb{R}$ an der Stelle $\underline{u}_0 \in U$ differenzierbar sind. Die <u>Ableitung</u> von f bei $\underline{u}_0$ ist die $m \times n$-Matrix

$$D_{\underline{u}_0} f = \left(\frac{df_i}{dx_j}\bigg|_{\underline{u}_0} \right)_{i=1,\dots,m;\; j=1,\dots,n}$$

Die Einträge $\frac{df_i}{dx_j}$ heißen <u>partielle Ableitungen</u>.

Bem: Die Ableitung ist wieder die lineare Bestapproximation von f in der Nähe von $\underline{u}_0$:

$$f(\underline{u}) = \underbrace{f(\underline{u}_0)}_{m\text{-Spalte}} + \underbrace{\overbrace{(D_{\underline{u}_0} f)}^{m\times n} \overbrace{(\underline{u}-\underline{u}_0)}^{n\text{-Spalte}}}_{m\text{-Spalte}} + \text{Terme höherer Ordnung}$$

Kettenregel II

$$\begin{array}{ccccc} \mathbb{R}^n & & \mathbb{R}^m & & \mathbb{R}^l \\ \text{IU} & & \text{IU} & & \text{IU} \\ U & \xrightarrow{f} & V & \xrightarrow{g} & W \end{array}$$

f differenzierbar an der Stelle $\underline{u}_0 \in U$

g differenzierbar an der Stelle $\underline{v}_0 := f(\underline{u}_0) \in V$

$\Rightarrow h := g \circ f : U \to W$ ist differenzierbar an der Stelle $\underline{u}_0 \in U$ und die Ableitung ergibt sich durch Matrixmultiplikation:

$$D_{\underline{u}_0} h = (D_{\underline{v}_0} g)(D_{\underline{u}_0} f)$$

$$D_{\underline{u}_0}(g \circ f) = (D_{f(\underline{u}_0)} g)(D_{\underline{u}_0} f)$$

Übung: Erkenne Kettenregel I als Spezialfall.

<u>Bem</u>: Ist $f : U \to \mathbb{R}$ eine reellwertige auf ganz U differenzierbare Funktion, so läßt sich die Ableitung als Funktion

$$Df : U \longrightarrow \mathbb{M}_{1\times n}(\mathbb{R}) \cong \mathbb{R}^n$$
$$\underline{u} \longmapsto D_{\underline{u}} f$$

auffassen. Dieselbe Funktion wird direkt durch den Gradienten gegeben:

$$\nabla f : U \longrightarrow \mathbb{R}^n$$
$$\underline{u} \longmapsto \nabla_{\underline{u}} f = \left(D_{\underline{u}} f\right)^T$$

<u>Def</u>: Eine auf U differenzierbare Funktion $f : U \to \mathbb{R}$ heißt <u>zweimal</u> differenzierbar, wenn $\nabla f : U \to \mathbb{R}^n$ auf ganz U differenzierbar ist. Die <u>zweite Ableitung</u>/ <u>Hessematrix</u> von f an der Stelle $\underline{u}_0 \in U$ ist die $n\times n$-Matrix

$$D^2_{\underline{u}} f = H_{\underline{u}} f := \left(\frac{d^2 f}{dx_i\, dx_j}\Big|_{\underline{u}_0} \right)_{i,j}$$

Satz von Schwarz:

Ist $f: U \to \mathbb{R}$ zweimal differenzierbar und sind alle partiellen Ableitungen zweiter Ordnung stetig, so ist die Hessematrix symmetrisch, d.h.:

$$\boxed{\frac{d^2 f}{dx_i\, dx_j} = \frac{d^2 f}{dx_j\, dx_i}}$$

Bsp: $f(x, y, z) = x \sin(yz) + \ln(xy)$

$$\frac{df}{dx} = \sin(yz) + y\frac{1}{xy}$$

$$\frac{df}{dz} = x \cos(yz)\, y$$

$$\frac{d^2 f}{dx\, dz} = \cos(yz)\, y$$

$$\frac{d^2 f}{dz\, dx} = \cos(yz)\, y$$

stimmen überein!

Taylor-Entwicklung:

Ist $f: U \to \mathbb{R}$ analytisch, so wird f in der Nähe eines Punktes $\underline{u}_0 \in U$ durch die Reihenentwicklung

$$\boxed{f\left(\underline{u}_0 + \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}\right) = f(\underline{u}_0) + \sum_{i=1}^{n} \frac{df}{dx_i} h_i + \frac{1}{2} \sum_{i,j=1}^{n} \frac{df}{dx_i dx_j} h_i h_j + \cdots}$$

beschrieben. In Vektorschreibweise stellen sich die ersten Glieder wiefolgt dar:

$$f(\underline{u}_0 + \underline{h}) = f(\underline{u}_0) + (D_{\underline{u}_0} f)\,\underline{h} + \frac{1}{2} \underline{h}^T (H_{\underline{u}_0} f)\,\underline{h} + \cdots$$

Extrema: $f : U \to \mathbb{R}$ „schön" (z.B. analytisch)

Erinnerung: Eine notwendige Bedingung dafür, daß f an der Stelle $\underline{u}_0$ ein Minimum hat, ist:

$$D_{\underline{u}_0} f = 0 \qquad (\text{äquivalent: } \nabla_{\underline{u}_0} f = 0)$$

Dann ist die Taylorentwicklung um $\underline{u}_0$:

$$f(\underline{u}_0 + \underline{h}) \approx f(\underline{u}_0) + \underline{h}^T (D^2_{\underline{u}_0} f)\, \underline{h}$$

Hinreichendes Kriterium:

1) $D_{\underline{u}_0} f = 0$
2) $H_{\underline{u}_0} f$ sei positiv definit

Dann hat f an der Stelle $\underline{u}_0$ ein lokales Minimum.

Analog:

1) $D_{\underline{u}_0} f = 0$
2) $-H_{\underline{u}_0} f$ ist pos. def. ($H_{\underline{u}_0} f$: neg. def.)

Dann hat f an der Stelle $\underline{u}_0$ ein lokales Maximum.

Def: Eine symmetrische Matrix $B \in M_{n\times n}(\mathbb{R})$ heißt <u>indefinit</u>, wenn es Vektoren $\underline{u}, \underline{v} \in \mathbb{R}^n$ mit

$$\underline{u}^T B \underline{u} < 0 < \underline{v}^T B \underline{v}$$

gibt.

<u>Kriterium</u>: Für $f: U \to \mathbb{R}$ und $\underline{u}_0 \in U$ sei

1) $\nabla_{\underline{u}_0} f = 0$ und
2) $H_{\underline{u}_0} f$ indefinit.

Dann hat f bei $\underline{u}_0$ weder ein Maximum noch ein Minimum, sondern einen <u>Sattelpunkt</u>, d.h. es gibt in jeder noch so kleinen Umgebung von $\underline{u}_0$ Punkte $\underline{v}$ und $\underline{w}$ mit:

$$f(\underline{v}) < f(\underline{u}_0) < f(\underline{w})$$