15.1.14

Repetitorium WiSe 2013/14

Analysis-Teil:

Stichworte:

(a) Mehrdimensionales Ableiten: partielle und totale Differenzierbarkeit, Höhere Ableitungen, Kettenregel in mehreren Veränderlichen, Richtungsableitungen, Schrankensatz

(b) Mehrdimensionale Taylorformel Taylorformel, Hessematrix

(a) Mehrdimensionales Ableiten

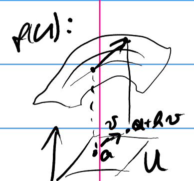
Betr. Funktionen $f : U \to \mathbb{R}^m$ mit $U \subset \mathbb{R}^n$ (offene Teilmenge),   Differenzierbarkeit? Stetigkeit vgl. Termin 11

- Fall $n = m$: f heißt Vektorfeld
- Fall $m = 1$, $n > 1$: f heißt Skalarfeld
- Fall $n = 1$: ist f stetig und U ein IV, so ist f eine Kurve.

  In diesem Fall ist $f'(a) = (f_1'(a), \dots, f_m'(a))$, wenn $f = (f_1, \dots, f_m)$,

  wobei f_j die j-te Koordinatenfunktion von f ist, d.h. $f_j := \underbrace{pr_j}_{\text{Projektion auf } j\text{-te Komponente}} \circ f$

  $pr_j : \mathbb{R}^n \to \mathbb{R}$, $(x_1, \dots, x_n) \mapsto x_j$



- Richtungsableitung: Ist $v \in K^n$ mit $\|v\| = 1$,

  so def. $D_v f(a) := \lim\limits_{h \to 0} \frac{1}{h}\left(f(a + hv) - f(a)\right)$

  $=$ "Wachstum von f bei a in Richtung v" $=$ "Ableitung in Richtung v"

- partielle Ableitung: für $1 \le i \le n$ ist

  $$D_i f(a) = \frac{\partial f}{\partial x_i}(a) := D_{e_i} f(a) \quad \text{definiert als}$$

  die Richtungsableitung in Richtung des kanon. Einheitsvektors $e_i = (0, \dots, 0, \underset{\text{Stelle } i}{1}, 0, \dots, 0)$

  Ist $f = (f_1, \dots, f_m)$ mit $f_j : U \to \mathbb{R}$, so ist $\frac{\partial f}{\partial x_i} = \left(\frac{\partial f_1}{\partial x_i}, \dots, \frac{\partial f_m}{\partial x_i}\right)$,

  und $\frac{\partial f_j}{\partial x_i}$ erhält man durch Ableiten nach der Variablen i

  [Denn $D_{e_i} f_j(a) = \lim\limits_{h \to 0} \frac{1}{h}\left(f_j(a + he_i) - f_j(a)\right) = \lim\limits_{h \to 0} \frac{1}{h}\big(f_j(\dots, a_{i-1}, a_i + h, a_{i+1}, \dots)$

  $- f_j(\dots, a_{i-1}, a_i, a_{i+1}, \dots)\big)$]

- Der Gradient von $f: U \to \mathbb{R}^1$, $U \subset \mathbb{R}^n$, im Punkt $a \in U$ ist gleich
  $\operatorname{grad} f(a) := (D_1 f(a), \ldots, D_n f(a)) \in \mathbb{R}^{1\times n}$ bzw. $\in \mathbb{R}^n$. [Zeilenvektor!]
  Die Funktion $\operatorname{grad}: U \to \mathbb{R}^n$, $x \mapsto \operatorname{grad} f(x)$ heißt Gradient von f.

- Ist $f: U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$, so heißt die Matrix
$$Df(a) := \begin{pmatrix} D_1 f_1(a) & \cdots & D_n f_1(a) \\ \vdots & & \vdots \\ D_1 f_m(a) & \cdots & D_n f_m(a) \end{pmatrix} \in \mathbb{R}^{m\times n}$$
  die Jacobi-Matrix bzw. Funktionalmatrix von f in a.
  Ihre Determinante heißt Funktionaldeterminante bzw. Jacobi-Determinante.
  In der j-ten Zeile steht der Gradient der j-ten Koordinatenfunktion f_j.
  In der i-ten Spalte steht die i-te partielle Ableitung der Funktion f: $D_i f(a) = Df(a) \cdot e_i$

Definitionen: Sei $f: U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$.
- f heißt partiell differenzierbar, falls alle partiellen Ableitungen existieren
- f heißt stetig partiell differenzierbar, falls alle partiellen Ableitungen existieren und stetig sind
- Höhere partielle Ableitungen: $D_i^k f(a) := D_i(D_i^{k-1} f)(a)$, $D_i^0 f(a) = f(a)$.
  Gemischte: $D_i D_j f(a)$ für: erst nach x_j, dann nach x_i ableiten, usw.
  (und dann a einsetzen)

Sätze:
- alle Richtungsableitungen existieren $\Rightarrow$ f partiell db
- ⊗ möglich, dass alle Richtungsabl. ex., ohne dass f stetig ist!
  z.B.: $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) := 1$ für $0 < y < x^2$ und $f(x,y) :\underset{\text{sonst}}{=} 0 \Rightarrow$ alle $D_v f(0,0) = 0$
- Satz von Schwarz: f 2mal st. partiell db. $\Rightarrow$ $D_i D_j f(a) = D_j D_i f(a)$ $\leadsto$ Hessematrix ist symmetrisch (s.u.)

Motivation totale Diff'barkeit: Bei $f: U \to \mathbb{R}$, $U \subset \mathbb{R}^1$: $f(x) = \underbrace{f(a) + f'(x)\cdot(x-a)}_{\text{lin. Approx. an } f \text{ in } a \leadsto \text{Tangentenglg. } y = f(a) + f'(x)\cdot(x-a)} + \text{Fehler}$

Wollen Lin. Approximation/Tangentenglg. übertragen auf $f: U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$.

Definition: Geg. $f: U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$. Die i-te Komponentenfkt. ist $f_i := pr_i \circ f$, $1 \le i \le m$.

f heißt in $a \in U$ diff'bar (total diff'bar)

$$:\Leftrightarrow \exists A \in \underbrace{Hom(\mathbb{R}^n, \mathbb{R}^m)}_{\text{bzw. } \mathbb{R}^{m \times n}} : f(x) = f(a) + A(x-a) + o(\|x-a\|)$$

wobei

$o(\|\xi\|) = \varphi(\xi)$ eine Fkt. $\varphi: V \to \mathbb{R}^m$, $0 \in V \subset \mathbb{R}^n$, ist mit $\boxed{\lim_{\xi \to 0} \frac{\varphi(\xi)}{\|\xi\|} = 0}$ ← "φ ist Fehler", ist eigentliche Aussage der Diff'barkeit

- Dabei ist $A \in \mathbb{R}^{m \times n}$ die Jacobimatrix $A = Df(a)$, die wir Ableitung von f in a nennen, und $Df(a) \cdot (x-a)$ ist "Matrix mal Vektor $x-a$"

- Die Glg. $y = f(a) + Df(a) \cdot (x-a)$ bestimmt die Tangential"hyper"ebene durch $(a, f(a))$, insb. für $m = 1$

Bsp.: $f: \mathbb{R}^2 \to \mathbb{R}$, $f\binom{x_1}{x_2} = x_1^2 + x_2^2 - 1 \to \operatorname{grad} f\binom{x_1}{x_2} = (2x_1, 2x_2)$, insb. $\operatorname{grad} f\binom{1}{2} = (2,4)$

Dann: $y = f(1,2) + (2,4) \cdot (x_1 - 1, x_2 - 2)^T = 4 + 2(x_1 - 1) + 4(x_2 - 2) = 2x_1 + 4x_2 - 6$ ist Tangentialebene von f in $a = \binom{1}{2}$ im $\mathbb{R}^3$

Definitionen: $f: U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$. Dann:

- f heißt diff'bar, falls $Df(a)$ in jedem $a \in U$ ex.
- f heißt stetig diff'bar, falls $Df: U \to \mathbb{R}^{m \times n}$, $x \mapsto Df(x)$ stetig ist.

Sätze:
- f stetig diff'bar $\Leftrightarrow$ f stetig partiell diff'bar
- f diff'bar $\Rightarrow$ f stetig, nicht "$\Leftarrow$"
- f diff'bar $\Rightarrow$ f partiell diff'bar, nicht "$\Leftarrow$" vgl. ⊛ auf S. -2-
- f diff'bar in a $\Leftrightarrow$ alle Koordinatenfunktionen f_i in a diff'bar
- f diff'bar $\not\Rightarrow$ f stetig partiell diff'bar, z.B. $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \begin{cases} y - x^2, & y \ge x^2 \\ y^2/x^2 - y, & 0 \le y < x^2 \end{cases}$, $f(x,-y) = -f(x,y)$, in $a = (0,0)$
- Produktregel: $f, g: U \to \mathbb{R}^1$, $U \subset \mathbb{R}^n$ diff'bar $\Rightarrow$ $D(fg) = \operatorname{grad}(fg) = \operatorname{grad}(f) g + f \operatorname{grad}(g)$
- $Df(x) = 0$ $\forall x \in U \Rightarrow f$ konstant auf jeder Zusammenhangskomponente von U
- f stetig und f partiell db $\not\Rightarrow$ f diff'bar, z.B. $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \frac{xy}{\sqrt{x^2+y^2}}$ für $(x,y) \ne (0,0)$, $f(0,0) = 0$

• **Kettenregel**: Vor.: $f : \underset{\mathbb{R}^n}{\overset{\cap}{U}} \longrightarrow \underset{\mathbb{R}^m}{\overset{\cap}{f(U)}} \subseteq V \xrightarrow{g} \mathbb{R}^k$, f in $a \in U$ db, g in $f(a) \in V$ db

Beh.: $g \circ f$ in a db, $\underbrace{D(g\circ f)(a)}_{\in \mathbb{R}^{k\times n}} = \underbrace{(Dg)(f(a))}_{\in \mathbb{R}^{k\times m}} \cdot \underbrace{(Df)(a)}_{\in \mathbb{R}^{m\times n}}$

• Spezialfall $k=1$: $\quad \frac{\partial(g\circ f)}{\partial t_i}(a) = \sum_{j=1}^{m} \frac{\partial g}{\partial x_j}(f_1(a),\dots,f_m(a)) \cdot \frac{\partial f_j}{\partial t_i}(a)$

für alle $1 \le i \le n$,

bzw. schreibbar als: $D_i(g\circ f)(a) = (D_1 g(f(a)),\dots,D_m g(f(a))) \cdot (D_i f_j(a))_{1\le j\le m}$,

bildet man rechts das Matrixprodukt,

so ist dies $= \sum_{j=1}^m D_j g(f(a)) \cdot D_i f_j(a)$.

Ist $k=m=1$, folgt $D(g\circ f)(a) = \langle (\operatorname{grad} g)\circ f, f' \rangle (a)$

• MWS: $f: U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$, f st. db., $a \in U$, $x \in \mathbb{R}^n$, $\{a+tx \,;\, t\in[0,1]\} \subseteq U$.

Dann: $f(a+x) - f(x) = \left(\int_0^1 Df(a+tx)\,dt\right) \cdot x$. (Spezialfall $m=1$: wie "alte" MWS)

• **Schrankensatz**: f st. db, $\|Df(a)\| \le L$ für alle $a\in U$, so gilt $\|f(x)-f(a)\| \le L \cdot \|x-a\|$ für alle $a,x \in U$

• Satz zur Berechnung von Richtungsableitungen: ($m=1$)

ist f (total) diff'bar in a, so gilt die Formel

$\boxed{D_v f(a) = \langle \operatorname{grad} f(a), v\rangle = Df(a)\cdot v}$ für alle $v \in \mathbb{R}^n$ mit $\|v\|=1$.

⌈ Herleitung: $D_v f(a) = \lim_{h\to 0} \frac{1}{h}(f(a+hv) - f(a)) = g'(0)$ für die Fkt. $g(h) := f(a+hv) = f\circ s$ mit $s(h) := a + hv$, $s: \mathbb{R}\to\mathbb{R}^n$.

Kettenregel $\leadsto$ $D_v f(a) = g'(0) = D(f\circ s)(0) = (Df)(s(0)) \cdot s'(0) = Df(a)\cdot v$. ⌋

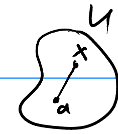
Bem.: • Vor. "total db" notwendig! z.B. $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \frac{x^2 y}{x^4+y^2}$ für $(x,y)\neq(0,0)$, $f(0,0)=0$,

es ist $D_{(u,v)} f(0,0) = \frac{u^2}{v}$ für $v\neq 0$ und $D_1 f(0,0) = D_2 f(0,0) = 0$, also $\operatorname{grad} f(0,0) = 0$.

• $|D_v f(a)| = |\langle \operatorname{grad} f(a), v\rangle| \overset{CS}{\le} \|\operatorname{grad} f(a)\| \cdot \|v\|$ und "=", falls $\operatorname{grad} f(a)$ in gleiche Richtung wie v zeigt

$\leadsto$ Also: Gradient von f zeigt in Richtung des stärksten Wachstums von f (= geom. Bedeutung von $\operatorname{grad} f$)

(0) Mehrdimensionale Taylorformel



- Taylor-Satz/-Formel: Sei $U \subset \mathbb{R}^n$, $a \in U$, $x \in \mathbb{R}^n$, $\forall t \in [0,1]: a+tx \in U$.
  Sei $f: U \to \mathbb{R}^1$ $(k+1)$-mal st. db. Dann ex. $\theta \in [0,1]$:

$$f(a+x) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(a)}{\alpha!} \cdot x^\alpha + \sum_{|\alpha| = k+1} \frac{D^\alpha f(a+\theta x)}{\alpha!} \cdot x^\alpha$$

- Summation über $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| := \alpha_1 + \ldots + \alpha_n$, $\alpha! := \alpha_1! \cdots \alpha_n!$,
  "multiindex" $\underbrace{x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}}_{\text{Monom vom Grad } |\alpha| \in \mathbb{N}_0}$, $D^\alpha f := D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n} f$
- a ist der Entwicklungspunkt der Taylorentwicklung.
- Alternative Version der Formel, wenn z für $a+x$ geschrieben wird:

$$f(z) = \underbrace{\sum_{|\alpha| \leq k} \frac{D^\alpha f(a)}{\alpha!} \cdot (z-a)^\alpha}_{\text{Taylorpolynom in } a \text{ vom Grad} \leq k} + \sum_{|\alpha| = k} \frac{D^\alpha f(a+\theta(z-a))}{\alpha!} \cdot (z-a)^\alpha,$$

  für alle z mit $\forall t \in [0,1]: a + t(z-a) \in U$,
  d.h. $\underbrace{\mathrm{str}(a,z)}_{\text{Strecke, die } a \text{ mit } z \text{ verbindet}} \subseteq U$
- Monom = irgendein Produkt aus Unbestimmten
- Grad eines Polynoms P = höchster Grad eines Monoms in P, das vorkommt
- P homogen: alle Monome in P haben denselben Grad
  Bsp.: x^2y ist Monom vom Grad 3, $x^2 + xy^2$ ist Pol. vom Grad 3, kein Monom,
  $xy^2 + x^2y - x^3$ ist homogenes Pol. vom Grad 3, da alle Monome darin Grad 3 haben
- Liegt jede Strecke $\mathrm{str}(a,z)$ für $a, z \in U$ ganz in U, so heißt U konvex.

- Hessematrix: $(\mathrm{Hess} f)(x) := (D_i D_j f(x))_{ij} \in \mathbb{R}^{n \times n}$, symmetrisch (nach S.v.Schwarz) falls f 2mal st. db. (partiell)
  Hess$f$ kommt in der Formel für das zweite Taylorpolynom vor:
  0. Taylorpolynom: $T_0(z) = f(a)$
  1. Taylorpolynom: $T_1(z) = T_0(z) + \sum_{|\alpha|=1} \frac{D^\alpha f(a)}{\alpha!}(z-a)^\alpha$
  $= f(a) + D_1 f(a) \cdot (z_1 - a_1) + \ldots + D_n f(a)(z_n - a) = f(a) + \langle \mathrm{grad} f(a), z-a \rangle$
  2. Taylorpolynom: $T_2(z) = T_1(z) + \sum_{|\alpha|=2} \frac{D^\alpha f(a)}{\alpha!}(z-a)^\alpha$
  $= f(a) + \langle \mathrm{grad} f(a), z-a \rangle + \frac{1}{2} \sum_{i,j=1}^{n} D_i D_j f(a) \cdot (x_i - a_i) \cdot (x_j - a_j)$
  $= T_1(z) + \frac{1}{2}(x-a)^T \cdot \mathrm{Hess} f(a) \cdot (x-a)$