Vorlesung Analysis II — SoSe '25 hhu
Teil 1: Differentialrechnung im $\mathbb{R}^n$ — K. Halupczok

an5: Partielle und totale Ableitungen

Stichworte: Funktionalmatrix, Gradient, Kettenregel, Richtungsableitungen

Literatur: [Hoff], Kapitel 9.4

5.1. **Einleitung:** Die totale Ableitung liefert einen einfachen Weg, Richtungsableitungen zu berechnen. Wir definieren für $m=1$ den Gradienten und beweisen die allgemeine Kettenregel.

5.2. **Konvention:** Betrachten Funktionen $f: U \to \mathbb{R}^m$ mit $U \subseteq \mathbb{R}^n$.

Sei $a \in U$ ein innerer Punkt von U, d.h. $\exists\, \delta > 0: U_a^\delta \subseteq U$.

Wir bezeichnen die Menge aller Richtungsvektoren $v \in \mathbb{R}^n$ mit

$$S^{n-1} := \{ v \in \mathbb{R}^n;\ \|v\|_2 = 1 \},$$ die $(n-1)$-dimensionale Sphäre im $\mathbb{R}^n$.

- Für einen inneren Punkt $a \in U$ und ein $v \in S^{n-1}$ haben wir in an4.8

$$D_v f(a) := \lim_{0 \neq t \to 0} \frac{1}{t}\big(f(a+tv) - f(a)\big)$$

als Richtungsableitung von f in a in Richtung v definiert.
(Diese Def. benutzt, dass $a + tv \in U$ ist für alle $t \in \mathbb{R}$ mit hinreichend kleinem $|t|$.)

- **Speziell:** Ist $v = e_i$ der i-te Einheitsvektor, so ist

$$D_i f(a) = \frac{\partial f}{\partial \xi_i}(a) = f_{\xi_i}(a) := (D_{e_i} f)(a)$$ die i-te partielle Ableitung.

5.3 **Satz:** Vor.: f in a diff'bar (d.h. total diff'bar), $v \in S^{n-1}$.

Beh.: f ist in Richtung v in a diff'bar und

$$(D_v f)(a) = \underbrace{(Df)(a)}_{\in \mathbb{R}^{m\times n}} \cdot \underbrace{v}_{\in \mathbb{R}^n} = \underbrace{f'(a)(v)}_{\text{mit } f'(a) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \text{ ausgedrückt, d.h. als lineare Abb.}}$$

Bew.: $\frac{1}{t}\big(f(a+tv) - f(a)\big) = \frac{1}{t}\Big(f'(a)(tv) + o(\|tv\|_2)\Big)$

$$\underset{f'(a) \text{ ist lin. Abb.}}{=} f'(a)(v) + \underbrace{\underbrace{\frac{|t| \cdot \|v\|_2}{t}}_{\text{beschränkt}} \cdot \underbrace{\frac{o(\|tv\|_2)}{\|tv\|_2}}_{\xrightarrow{t\to 0} 0}}_{\xrightarrow{t \to 0} 0}.$$

□

5.4. <u>Bsp.:</u> $f: \mathbb{R}^2 \to \mathbb{R}^3$, $f(x,y) = \begin{pmatrix} x^2-y \\ 3x \\ \sin(y) \end{pmatrix}$ gibt $f'(x,y) = \begin{pmatrix} 2x & -1 \\ 3 & 0 \\ 0 & \cos(y) \end{pmatrix}$, sei $a = \begin{pmatrix} 1 \\ \pi/4 \end{pmatrix} \in \mathbb{R}^2$.

Dann: $f'\left(\begin{pmatrix} 1 \\ \pi/4 \end{pmatrix}\right) = \begin{pmatrix} 2 & -1 \\ 3 & 0 \\ 0 & \frac{1}{2}\sqrt{2} \end{pmatrix}$ und sei $v = \frac{1}{5}\begin{pmatrix} 3 \\ 4 \end{pmatrix} \in \mathbb{R}^2$, haben $\|v\|_2 = \frac{1}{5}\cdot\sqrt{3^2+4^2} = 1$.

Die Ableitung von f in a in Richtung v berechnet sich dann als $f'\begin{pmatrix} 1 \\ \pi/4 \end{pmatrix} \cdot \frac{1}{5}\begin{pmatrix} 3 \\ 4 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 2\cdot 3 + (-1)\cdot 4 \\ 3\cdot 3 + 0\cdot 4 \\ 0\cdot 3 + \frac{1}{2}\sqrt{2}\, 4 \end{pmatrix} = \underline{\frac{1}{5}\begin{pmatrix} 2 \\ 9 \\ 2\sqrt{2} \end{pmatrix}}$.

Die partiellen Ableitungen sind $D_1 f(a) = f'\begin{pmatrix} 1 \\ \pi/4 \end{pmatrix} \cdot e_1 = \underline{\underline{\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}}}$, $D_2 f(a) = f'\begin{pmatrix} 1 \\ \pi/4 \end{pmatrix} \cdot e_2 = \underline{\underline{\begin{pmatrix} -1 \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix}}}$.

5.5. <u>Folgerungen:</u> Sei f in a diff'bar.

<u>Beh.:</u> a) f ist in a in jeder Koordinate partiell diff'bar,

b) $\underline{(D_i f)(a) = f'(a)\cdot e_i}$, und dies ist die i-te Spalte von $f'(a)$,

denn Sie wissen ja:

die Spalten einer Matrix sind genau die Bilder der Einheitsvektoren.

Also sind die Spalten von f' genau die partiellen Ableitungen von f.

c) Es ist

$$f'(a) = Df(a) = \begin{pmatrix} (D_1 f_1)(a) & \cdots & (D_n f_1)(a) \\ (D_1 f_2)(a) & \cdots & (D_n f_2)(a) \\ \vdots & & \vdots \\ (D_1 f_m)(a) & \cdots & (D_n f_m)(a) \end{pmatrix}$$

d) Es ist $\underline{(D_j f_i)(a) = pr_i(D_j f(a))}$, also $D_j f_i = pr_i \circ D_j f$ für alle $j \in \{1,\dots,n\}$, $i \in \{1,\dots,m\}$.

<u>Bew.:</u> a), b), c): Klar mit <u>5.3</u> und $v = e_i$.

Für d): Haben $f(a) = \begin{pmatrix} f_1(a) \\ \vdots \\ f_m(a) \end{pmatrix}$ und $D_j f(a) = \begin{pmatrix} D_j f_1(a) \\ D_j f_2(a) \\ \vdots \\ D_j f_m(a) \end{pmatrix}$,

also $pr_i(D_j f(a)) = D_j f_i(a)$. □

5.6. <u>Def.:</u> Man nennt $(D_j f_i(a))_{\substack{1\le i\le m \\ 1\le j\le n}} \in \mathbb{R}^{m\times n}$

die <u>Jacobimatrix/Funktionalmatrix</u> von f in a.

<u>Bem.:</u> „$\Leftarrow$" kann in <u>5.5</u> nicht gelten, die Existenz der partiellen Ableitungen reicht nicht zum Nachweis der Differenzierbarkeit! (Vgl. <u>4.15, 4.16</u>)

5.7. Fall: Sei $m = 1$, also f ein Skalarfeld.

Dann: $f'(a) = (D_1 f(a), \ldots, D_n f(a)) =: (\operatorname{grad} f(a))^T$ "Gradient"

$=: (\nabla f(a))^T$ "Nabla"

Wir nennen den Spaltenvektor $\operatorname{grad} f(a) = \begin{pmatrix} D_1 f(a) \\ \vdots \\ D_n f(a) \end{pmatrix} \in \mathbb{R}^n$ den Gradient von f in a.

Mit $\nabla := \begin{pmatrix} D_1 \\ \vdots \\ D_n \end{pmatrix}$ bezeichnen wir den Nabla-Operator.

Somit: $f(x) = f(a) + (\operatorname{grad} f(a))^T \cdot (x-a) + o(\|x-a\|)$

$\Rightarrow f(x) = f(a) + \langle \operatorname{grad} f(a), x-a \rangle + o(\|x-a\|)$.

Sei $\operatorname{grad} f(a) \neq 0$, und betrachte alle $v \in S^{n-1}$.

Dann gilt: $|D_v f(a)| = |f'(a)(v)| = |(\operatorname{grad} f(a))^T \cdot v|$

$$= |\langle \operatorname{grad} f(a), v \rangle| \underset{\substack{\uparrow \\ \text{Cauchy-Schwarz-Ungl.} \\ \text{(Anhang 7 in an1)}}}{\leq} \|\operatorname{grad} f(a)\|_2 \cdot \underbrace{\|v\|_2}_{=1},$$

wobei „=" genau dann gilt, wenn v parallel zu $\operatorname{grad} f(a)$,

d.h. ex. $t \in \mathbb{R} \setminus \{0\}$: $\operatorname{grad} f(a) = t v$,

in diesem Fall wird für $D_v f(a)$ der maximale Wert angenommen.

Für $\tilde{v} := \frac{1}{\|\operatorname{grad} f(a)\|_2} \operatorname{grad} f(a)$ gilt demnach:

$$|D_{\tilde{v}} f(a)| = \|\operatorname{grad} f(a)\|_2 .$$

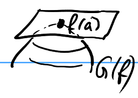
5.8. FAZIT: $\operatorname{grad} f(a)$ ist die Richtung maximaler Steigung von f in a (welche dann $\|\operatorname{grad} f(a)\|_2$ beträgt).

5.9. Veranschaulichung: Der Graph von f, nämlich $G(f) := \{ \binom{x}{f(x)} \in \mathbb{R}^{n+1} \}$ (falls $U = \mathbb{R}^n$), wird in $\binom{a}{f(a)}$ approximiert durch

$$\xi_{n+1} = f(a) + \langle \operatorname{grad} f(a), x-a \rangle,$$

und dies ist die Glg. für eine n-dim. Hyperebene im $\mathbb{R}^{n+1}$!

Diese heißt Tangentialhyperebene von f im Punkte a.



5.10. Bsp. mit $n=2$: $f(x,y) = x^2 - \frac{3}{2} y^2$, $a = \binom{\alpha_1}{\alpha_2}$ $\rightsquigarrow \operatorname{grad} f(a) = (2\alpha_1, -3\alpha_2)^T$, und

$\xi_3 = \alpha_1^2 - \frac{3}{2}\alpha_2^2 + 2\alpha_1(\xi_1 - \alpha_1) - 3\alpha_2(\xi_2 - \alpha_2)$ ist die Glg. der Tangentialhyperebene im $\mathbb{R}^3$.

5.11. <u>Fall</u>: Sei $m=n$, also f ein <u>Vektorfeld</u>.

Dann heißt $\operatorname{div} f(a) = \langle \nabla, f \rangle (a) := \sum_{i=1}^{n} D_i f_i(a) = \operatorname{spur} f'(a) \in \mathbb{R}$

die <u>Divergenz von f in a</u>.

[Erinnerung an lineare Algebra: $\operatorname{spur} A = \sum_{i=1}^{n} a_{ii}$, wenn $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, heißt <u>Spur</u> der Matrix A.

Die Funktionalmatrix ist quadratisch: $Df(a) = f'(a) = \begin{pmatrix} D_1 f_1(a) & D_2 f_1(a) & \cdots & D_n f_1(a) \\ \vdots & \vdots & & \vdots \\ D_1 f_n(a) & D_2 f_n(a) & \cdots & D_n f_n(a) \end{pmatrix} \in \mathbb{R}^{n\times n}$.

Ihre Determinante heißt <u>Funktionaldeterminante</u> bzw. <u>Jacobideterminante</u>.

Eine wichtige Rechenregel für das Ableiten verketteter Funktionen im Mehrdimensionalen ist die (allgemeine)

5.12. <u>Kettenregel</u>: <u>Vor.</u>: $a \in U \xrightarrow{f} U_1 \xrightarrow{g} \mathbb{R}^k$, $U \subset \mathbb{R}^n$, $U_1 \subset \mathbb{R}^m$, f in a diff'bar, g in $f(a)$ diff'bar

(insb. a innerer Punkt von U, $f(a)$ innerer Punkt von U_1).

<u>Beh.</u>: $g\circ f$ in a diff'bar,

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

<u>Bew.</u>: Setze $b := f(a)$.

Dann gilt für $r_1(x) = o(\|x-a\|)$, dass $f(x) = f(a) + f'(a)(x-a) + r_1(x)$, ⊛

und für $r_2(y) = o(\|y-b\|)$, dass $g(y) = g(b) + g'(b)(y-b) + r_2(y)$.

$\underset{\substack{y=f(x)\\ b=f(a)}}{\Rightarrow} g(f(x)) = g(f(a)) + g'(f(a))\,(f(x)-f(a)) + r_2(f(x))$

$\underset{⊛}{\Rightarrow} g\circ f(x) = g\circ f(a) + g'(f(a))\cdot(f'(a)(x-a)) + g'(f(a))(r_1(x)) + r_2(f(x))$

$\Rightarrow g\circ f(x) = g\circ f(a) + \underbrace{(Dg)(f(a))\cdot(Df)(a)}_{= D(g\circ f)(a)}\cdot(x-a) + g'(f(a))(r_1(x)) + r_2(f(x))$ $\leadsto$ Beh.

Noch z.z.: $g'(f(a))(r_1(x)) + r_2(f(x)) = o(\|x-a\|)$. (haben $\|Ax\|_\infty \le m\|A\|_\infty \|x\|_\infty$)

(Def. für eine Matrix $A \in \mathbb{R}^{m\times m}$ den Wert $\|A\|_\infty := \max_{i,j} |\alpha_{ij}|$, wenn $A = (\alpha_{ij})_{ij}$.)

- Es gilt:

$$\|\underbrace{g'(f(a))}_{\in \mathbb{R}^{k\times m}} \cdot \underbrace{r_1(x)}_{\in \mathbb{R}^m}\|_\infty \overset{\text{Ü}}{\le} m \underbrace{\|g'(f(a))\|_\infty}_{\substack{\text{maximaler Eintrag} \\ \text{der Matrix } g'(f(a)) \\ \text{im Betrag, unabh. von } x}} \cdot \|r_1(x)\|_\infty = o(\|x-a\|)$$

nach Vor. an r_1.

- Bleibt, z.z.: $r_2(f(x)) = o(\|x-a\|)$.

Haben $r_2(y) = o(\|y-b\|)$, d.h. $\frac{r_2(y)}{\|y-b\|_\infty} \xrightarrow{y\to b} 0$.

Wähle $\eta > 0$. Dann ist für y nahe b: $r_2(y) < \eta \cdot \|y-b\|_\infty$.

Es folgt: $r_2(f(x)) < \eta \|f(x)-f(a)\|_\infty = \eta \|f'(a)(x-a) + r_1(x)\|_\infty$

$$\le \underbrace{\eta \cdot n \underbrace{\|f'(a)\|_\infty}_{\substack{\text{Konstant,} \\ \text{d.h. unabh. von } x}} \cdot \|x-a\|_\infty}_{\le \tilde{\eta} \|x-a\|_\infty} + \underbrace{\eta \|r_1(x)\|_\infty}_{= o(\|x-a\|)},$$

also $\frac{r_2(f(x))}{\|x-a\|_\infty} \xrightarrow{x\to a} 0$, da $\eta > 0$ beliebig.

$\square$

5.13. **Illustration der Kettenregelformel:** $D(g\circ f)(a) = (Dg)(f(a)) \cdot (Df)(a)$

$$\begin{pmatrix} D_1(g\circ f)_1(a) & \cdots & D_n(g\circ f)_1(a) \\ \vdots & & \vdots \\ D_1(g\circ f)_k(a) & \cdots & D_n(g\circ f)_k(a) \end{pmatrix} = \begin{pmatrix} D_1 g_1(f(a)) & \cdots & D_m g_1(f(a)) \\ \vdots & & \vdots \\ D_1 g_k(f(a)) & \cdots & D_m g_k(f(a)) \end{pmatrix} \cdot \begin{pmatrix} D_1 f_1(a) & \cdots & D_n f_1(a) \\ \vdots & & \vdots \\ D_1 f_m(a) & \cdots & D_n f_m(a) \end{pmatrix}.$$

5.14. • Bsp.: Sei $a \in U$, $T_a : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto x+a$, ist diff'bar. (Translation um a)

Dann: f in a diff'bar $\Leftrightarrow$ $f \circ T_{-a}$ in 0 diff'bar

$\Leftrightarrow$ $T_{-f(a)} \circ f$ in a diff'bar.

• Bsp.: $\mathbb{R}^3 \xrightarrow{f} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$, $n = 3$, $m = 2$, $g = 1$, $f(x,y,z) = \begin{pmatrix} x^2+y \\ 2y-z \end{pmatrix}$, $g(u,v) = uv$.

Betr. $a = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^3$. Dann: $D(g\circ f)(a) \underset{\text{K.R.}}{=} (Dg)(\underbrace{f(a)}_{=\binom{3}{1}}) \cdot (Df)(\underbrace{a}_{=\left(\begin{smallmatrix}1\\2\\3\end{smallmatrix}\right)}) = (1,\ 3) \cdot \begin{pmatrix} 2\cdot 1 & 1 & 0 \\ 0 & 2 & -1 \end{pmatrix} = (2, 7, -3) \in \mathbb{R}^{1\times 3}$.

$\rightsquigarrow f(a) = \begin{pmatrix} 1+2 \\ 2\cdot 2-3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$

5.15. <u>Bsp.</u>: $n=k=1$, $m=3$: $f(t)=\begin{pmatrix}\varphi(t)\\ \psi(t)\\ \chi(t)\end{pmatrix}$ mit $\varphi,\psi,\chi: U\to\mathbb{R}$, $a\in U\subseteq\mathbb{R}$,
φ,ψ,χ diff'bar in $a\in\mathbb{R}$.

Sei $b:=f(a)$, $g: U_1\to\mathbb{R}$ diff'bar, $U_1\subseteq\mathbb{R}^3$, b innerer Punkt von U_1, $h:=g\circ f$.
Somit zeigt die Kettenregel, dass

$$h'(a)=g'(b)\cdot f'(a)=\left(D_1g(f(a)),\ D_2g(f(a)),\ D_3g(f(a))\right)\cdot\begin{pmatrix}\varphi'(a)\\ \psi'(a)\\ \chi'(a)\end{pmatrix}$$

$$=D_1g(f(a))\varphi'(a)+D_2g(f(a))\psi'(a)+D_3g(f(a))\chi'(a)\in\mathbb{R}$$

In der Literatur wird dafür oft geschrieben: (mit $x(t)=\varphi(t)$, $y(t)=\psi(t)$, $z(t)=\chi(t)$)

$$\frac{dh}{dt}=\frac{\partial g}{\partial x}\frac{dx}{dt}+\frac{\partial g}{\partial y}\frac{dy}{dt}+\frac{\partial g}{\partial z}\frac{dz}{dt}$$ oder auch $$dh=\frac{\partial g}{\partial x}dx+\frac{\partial g}{\partial y}dy+\frac{\partial g}{\partial z}dz.$$

5.16. Spezialfall $k=1$ der Kettenregel, Verallgemeinerung von <u>5.15</u>:

$$\frac{\partial g}{\partial t_j}(a)=\sum_{i=1}^{m}\frac{\partial g}{\partial x_i}(f_1(a),\dots,f_m(a))\cdot\frac{\partial f_i}{\partial t_j}(a)\qquad \text{für alle } j\in\{1,\dots,n\},$$

bzw. schreibbar als $D_j(g\circ f)(a)=(D_1g(f(a)),\dots,D_mg(f(a)))\cdot(D_jf_i(a))_{1\le i\le m}$.
Bildet man rechts das Matrixprodukt, so ist dies $=\sum_{i=1}^{m}D_ig(f(a))\cdot D_jf_i(a)$.
Ist $k=n=1$, folgt $D(g\circ f)(a)=\langle(\operatorname{grad} g)\circ f, f'\rangle(a)$, vgl. <u>5.15</u>.

5.17 Berechnung von Richtungsableitungen im Fall $\underline{\underline{m=1}}$:
<u>Satz 5.3</u> kann mit der Kettenregel bewiesen werden:
⌈ Haben $D_vf(a)=\lim\limits_{h\to 0}\frac{1}{h}(f(a+hv)-f(a))=g'(0)$
für die Funktion $g(h):=f(a+hv)=f\circ s(h)$,
wo $s(h):=a+hv$, $s:\mathbb{R}\to\mathbb{R}^n$.

Die <u>Kettenregel</u> liefert $D_vf(a)=g'(0)=D(f\circ s)(0)\underset{K.R.}{=}(Df)(s(0))\cdot s'(0)=Df(a)\cdot v$ ⌋

5.18. <u>Bem.</u>: Die Voraussetzung "f total diff'bar" in <u>5.3</u> ist notwendig!
Betr. z.B. $f:\mathbb{R}^2\to\mathbb{R}$, $f(x,y)=\begin{cases}\frac{x^2y}{x^4+y^2} & \text{für } (x,y)\neq(0,0),\\ 0 & \text{sonst.}\end{cases}$

Dann ist $\Big(\underbrace{D_{(u,v)}}_{\text{Richtungsvektor}}f\Big)(0,0)=\frac{u^2}{v}$ für $v\neq 0$, denn $\frac{1}{t}\Big(f(\underbrace{(0,0)+t(u,v)}_{=(tu,tv)})-f(0,0)\Big)=\frac{u^2v}{t^2u^4+v^2}\xrightarrow[t\to 0]{}\underbrace{\frac{u^2}{v}}_{\neq 0}$,

und $D_1f(0,0)=D_2f(0,0)=0$, also $\operatorname{grad} f(0,0)=\binom{0}{0}$, und die r.S. in <u>5.3</u> ist $=0$.