(91) Erinnerung (Kettenregel)

$f: \mathbb{R}^n \to \mathbb{R}^m$, $g: \mathbb{R}^m \to \mathbb{R}^k$

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

Jetzt: Spezialfall: $f: \mathbb{R}^n \to \mathbb{R}^m$, $g: \mathbb{R}^m \to \mathbb{R}$

Dann $g\circ f: \mathbb{R}^n \to \mathbb{R}$

$$Dg(y) = (D_1 g(y), \dots, D_m g(y)) \in \mathbb{R}^{1\times m}$$

$$\nabla g(y) = (D_1 g(y), \dots, D_m g(y)) \in \mathbb{R}^m$$

$$D(g\circ f)(x) = (D_1(g\circ f)(x), \dots, D_n(g\circ f)(x))$$

$$= \underbrace{(Dg)(f(x))}_{\mathbb{R}^{1\times m}} \cdot \underbrace{Df(x)}_{\mathbb{R}^{m\times n}}$$

Achtung: Für die Gradienten gilt

$$\nabla(g\circ f)(x) = (Df(x))^T (\nabla g)(f(x))$$

Bem: Im Allgemeinen betrachten wir Vektoren als Spaltenvektoren

D.h. eigentlich gilt:

$$\nabla g(y) = \begin{pmatrix} D_1 g(y) \\ \vdots \\ D_m g(y) \end{pmatrix}$$

Streng genommen: $\nabla g(y) = Dg(y)^T$

$$\nabla(g\circ f)(x) = D(g\circ f)(x)^T$$
$$= (Dg(f(x)) \cdot Df(x))^T$$
$$= Df(x)^T \cdot (Dg)(f(x))^T$$
$$= (Df(x))^T \cdot \nabla g(f(x))$$

(9.2) **Def 6.7** (Richtungsableitungen)
Sei $U \subseteq \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$. Sei ferner $x \in U$ und $v \in \mathbb{R}^n$ mit $\|v\| = 1$. Dann heißt f richtungs diff.bar im Punkt x in Richtung v, wenn

$$D_v f(x) := \frac{\partial f}{\partial v}(x) := \lim_{h \to 0} \frac{f(x+hv) - f(x)}{h} = \frac{d}{dt} f(x+tv)\Big|_{t=0}$$

existiert. $D_v f(x)$ heißt Richtungsableitung von f in x in Richtung v.

Bem $\frac{\partial f}{\partial x_i}(x) = D_i f(x) = D_{e_i} f(x) = \frac{\partial f}{\partial e_i}(x)$.

Satz 6.8 Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \to \mathbb{R}$ stetig diff.bar. Dann ist f für jedes $v \in \mathbb{R}^n$ mit $\|v\| = 1$ richtungs-diff.bar und es gilt

$$D_v f(x) = \langle \nabla f(x), v \rangle.$$

Beweis: Sei $\varepsilon > 0$ klein genug, so dass $F: (-\varepsilon, \varepsilon) \to \mathbb{R}$, $F(t) = f(x+tv)$ wohldef. ist.

z.z. F diff.bar in $t=0$ und $F'(0) = \langle \nabla f(x), v \rangle$.

Betrachte $\varphi: (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ mit $\varphi(t) = x + tv \in U$.
Dann ist φ diff.bar mit $D\varphi(t) = v \in \mathbb{R}^{n \times 1}$ und $F = f \circ \varphi$. Somit liefert die Kettenregel

$$F'(t) = DF(t) = (Df)(\varphi(t))\, D\varphi(t) = (\nabla f)(x+tv)^T v = \langle \nabla f(x+tv), v \rangle,$$

also $F'(0) = \langle \nabla f(x), v \rangle$. ∎

$[a, b \in \mathbb{R}^n, \quad a^T b = \sum a_i b_i = \langle a, b \rangle = a \cdot b]$

Erinnerung CS-Ungleichung

$$\langle a, b \rangle \leq \|a\| \cdot \|b\|$$

mit „=" gdw. $a = tb$, $t \geq 0$

93) **Bem** Sei $\nabla f(x) \neq 0$, $\|v\| = 1$,

$D_v f(x) = \langle \nabla f(x), v \rangle \overset{(*)}{\leq} \|\nabla f(x)\| \cdot \|v\| = \|\nabla f(x)\|$

d.h. $|D_v f(x)| \leq \|\nabla f(x)\|$.

In $(*)$ gilt „$=$" gdw. $\nabla f(x) = t v$, $t \geq 0$,
d.h. gdw. $\nabla f(x) = \|\nabla f(x)\| v$, bzw. wenn

$$v = \frac{\nabla f(x)}{\|\nabla f(x)\|}.$$

D.h. $D_v f(x)$ wird maximal für $v = \frac{\nabla f(x)}{\|\nabla f(x)\|}$.

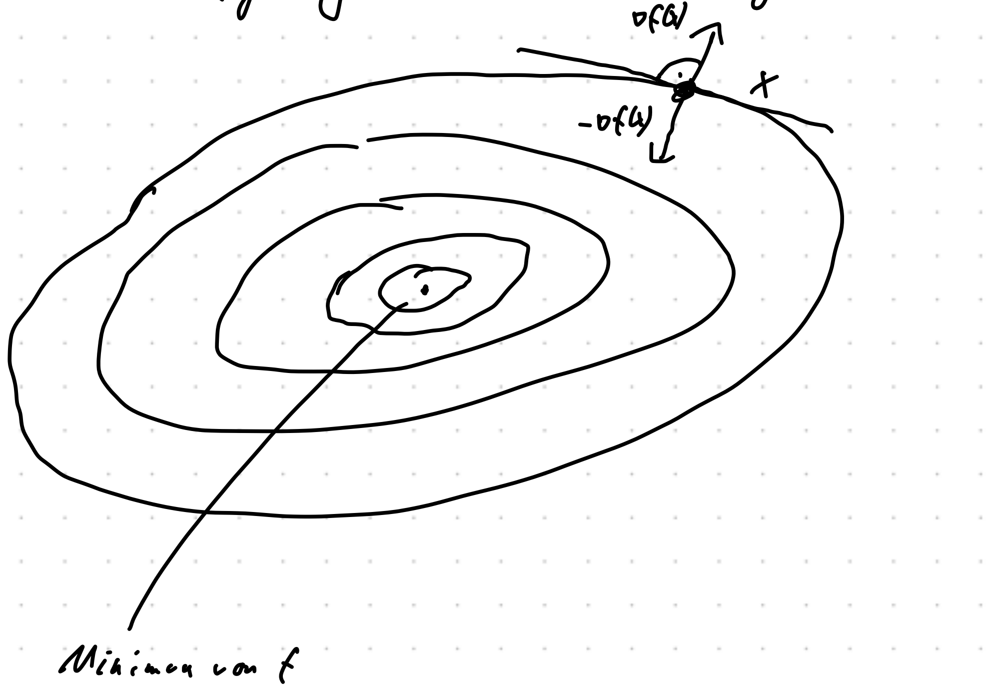
Man sagt $\nabla f(x)$ ist die <u>Richtung des steilsten Anstiegs</u> von f in x.

<u>Analog</u> gilt <u>$-\nabla f(x)$</u> ist die Richtung des <u>steilsten Abstiegs</u>.

<u>Anwendung</u>: Gradientenverfahren um Minima von f zu berechnen

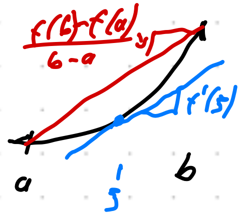
$$x^{k+1} = x^k - \sigma_k \nabla f(x^k)$$

Bem Sei $v \perp \nabla f(x)$, d.h. $\langle \nabla f(x), v \rangle = 0$, dann gilt $D_v f(x) = 0$, d.h. die Steigung von f in Richtung v ist 0.



Genau genommen gilt:
$\nabla f(x)$ steht senkrecht auf
$N_f(f(x)) = \{y \in U \mid f(y) = f(x)\}$

(94) Mittelwert



Erinnerung (Ana I) Sei $f:[a,b]\to\mathbb{R}$ diff.bar in $[a,b]$. Dann gilt

(1) $\exists \xi\in(a,b): \ f(b)-f(a) = f'(\xi)(b-a)$

bzw.

(2) $\exists \theta\in(0,1): \ f(b)-f(a) = f'(a+\theta(b-a))\,(b-a)$

Falls f zusätzlich stetig diff.bar

(3) $$f(b)-f(a) = \int_a^b f'(t)\,dt$$

bzw.

(4) $$f(b)-f(a) = \int_0^1 f'(a+t(b-a))\,dt\cdot(b-a)$$

D.h. in (4) wird $f'(a+\theta(b-a))$ aus (2) durch den Mittelwert

$$\int_0^1 f'(a+t(b-a))\,dt = f'(a+\theta(b-a))$$

ersetzt.

Wir verallgemeinern nun (4) auf $f: \mathbb{R}^n\to\mathbb{R}^m$.

Def 6.9 Sei $f:[a,b]\to\mathbb{R}^{m\times n}$. Dann heißt f Riemann-integrierbar, wenn jedes $f_{ij}:[a,b]\to\mathbb{R}$ Riemann-int.bar ist und wir setzen:

$$\int_a^b f(t)\,dt := \left(\int_a^b f_{ij}(t)\,dt\right)_{i,j}\in\mathbb{R}^{m\times n}.$$

Bem Für $f:[a,b]\to\mathbb{R}^n$ analog.

Lemma 6.10 Sei $f:[a,b]\to\mathbb{R}^n$ stetig, dann

$$\left\|\int_a^b f(t)\,dt\right\| \le \int_a^b \|f(t)\|\,dt$$

Beweis Sei $z=\int_a^b f(t)\,dt$. Dann gilt

$$\|z\|^2 = \langle z,z\rangle = \left\langle \int_a^b f(t)\,dt, z\right\rangle \overset{(*)}{=} \int_a^b \langle f(t),z\rangle\,dt$$
$$\le \int_a^b \|f(t)\|\cdot\|z\|\,dt = \|z\|\cdot\int_a^b \|f(t)\|\,dt.$$ ☒

(teile durch $\|z\|$).

(95) zu (*)

$$\left\langle \int_a^b f(t)\,dt, z \right\rangle = \sum_{i=1}^n \left(\int_a^b f(t)\,dt \right)_i z_i$$

$$= \sum_{i=1}^n \int_a^b f_i(t)\,dt\, z_i = \sum_{i=1}^n \int_a^b f_i(t) z_i\,dt$$

$$= \int_a^b \underbrace{\sum_{i=1}^n f_i(t) z_i}_{= \langle f(t), z \rangle}\,dt.$$

Satz 6.11 (Mittelwertsatz)

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \to \mathbb{R}^n$ stetig diff.bar. Seien $a, b \in U$, sodass $a + t(b-a) \in U \ \forall t \in [0,1]$. Dann gilt

$$(5) \quad f(b) - f(a) = \left(\int_0^1 Df(a + t(b-a))\,dt \right) \cdot (b-a).$$

Beweis Sei $g: [0,1] \to \mathbb{R}^n$, $g(t) = f(a + t(b-a))$.

Dann ist jedes $g_i: [0,1] \to \mathbb{R}$ stetig diff.bar.

Also gilt nach MWS (4) $\forall i$

$$f_i(b) - f_i(a) = g_i(1) - g_i(0) = \int_0^1 g_i'(0 + t(1-0))\,dt\,(1-0)$$

$$= \int_0^1 g_i'(t)\,dt$$

$$= \int_0^1 \langle Df_i(a + t(b-a)), b-a \rangle\,dt$$

$$= \left\langle \int_0^1 Df_i(a + t(b-a))\,dt, b-a \right\rangle.$$

Das ist gerade die i-te Komponente von (5). ∎