

**Merkblatt zur Differentiation in normierten Räumen**

- **Definition:** Seien  $(V, \|\cdot\|_V), (W, \|\cdot\|_W)$  normierte Räume. Dann heißt  $f : V \rightarrow W$  differenzierbar in  $x \in V$ , wenn eine lineare, stetige Abbildung  $Df|_x \in L(V, W)$  existiert mit

$$\|f(x+h) - f(x) - Df|_x h\|_W = o(\|h\|_V), \quad h \rightarrow 0$$

Hinreichend dafür ist z.B.

$$\lim_{h \rightarrow 0} \frac{\|f(x+h) - f(x) - Df|_x h\|_W}{\|h\|_V} = 0$$

Beachte also:  $Df : V \rightarrow L(V, W)$ . Wir führen  $Df$  (und die höheren Ableitungen) in dieser abstrakten Form ein, um einen *koordinatenfreien* Ableitungsbegriff zur Verfügung zu haben, der für die Fälle  $V = \mathbb{R}^n, W = \mathbb{R}^m$  die dort üblichen Ableitungsbegriffe  $\{\text{Gradient, Jacobimatrix, Hessematrix}\}$  induziert.

- **Beispiel:**  $f : \mathbb{R}^d \rightarrow \mathbb{R}, f(x) = \|x\|_2^2$ .

$$f(x+h) - f(x) = \|x+h\|_2^2 - \|x\|_2^2 = \langle x+h, x+h \rangle - \langle x, x \rangle = 2\langle x, h \rangle + \langle h, h \rangle$$

Also  $Df|_x h = 2\langle x, h \rangle$  (linear in  $h$ , stetig als Abb. endlich-dim. Räume), denn

$$|f(x+h) - f(x) - 2\langle x, h \rangle| = |\langle h, h \rangle| = \|h\|_2^2 = o(\|h\|_2).$$

- **Zusammenhang zwischen  $Df$  und Jacobimatrix  $J_f$ :** Ist  $\dim V = n < \infty$  mit Basis  $\{v_1, \dots, v_n\} \subset V$ , dann ist die lineare, stetige Abbildung  $Df|_x : V \rightarrow W$  ja schon durch Angeben von

$$(Df|_x v_1, \dots, Df|_x v_n) \in W^n$$

eindeutig festgelegt (algebraisch ausgedrückt:  $L(V, W) \cong W^n$ ). Für  $V = \mathbb{R}^n$  und  $W = \mathbb{R}^m$  besitzt dann  $Df|_x$  eine *Matrixdarstellung*  $Df|_x h = Jh$  mit  $J \in \mathbb{R}^{m \times n}$ . Bezüglich der kanonischen Basen  $\{e_i\}$  ist  $J$  genau die Jacobimatrix  $J_f(x)$ , also  $J_f(x) = (Df|_x e_i)_{i=1, \dots, n} \in \mathbb{R}^{m \times n}$ .

- **Beispiele:**

- $f : \mathbb{R}^d \rightarrow \mathbb{R}, f(x) = \|x\|_2^2, J_f(x) = \text{grad } f(x) = 2x^T$ , aber  $Df|_x h = 2\langle x, h \rangle = 2x^T h = \text{grad } f(x)h$ . Man sieht  $\text{grad } f(x) = (Df|_x e_i)_{i=1, \dots, d}$ .
- $f : \mathbb{R}^n \rightarrow \mathbb{R}^m, f(x) = Ax, A \in \mathbb{R}^{m \times n}, J_f(x) = A$ , aber  $Df|_x h = Ah (= f(h)), J_f(x) = (Df|_x e_i)_{i=1, \dots, n}$ .

- **Höhere Ableitungen (z.B. Hessematrix  $H_f$ ):**  $L(V, W)$  ist ein normierter Raum mit der Operatornorm  $\|u\| := \|u\|_{L(V, W)} := \sup_{0 \neq x \in V} \frac{\|u(x)\|_W}{\|x\|_V}$ . Möglicherweise ist also bei differenzierbarem  $f : V \rightarrow W$  auch  $Df : V \rightarrow L(V, W)$  wieder differenzierbar. Falls ja, so wäre

$$D^2 f := D(Df) : V \rightarrow L(V, L(V, W)) \cong L_2(V, W).$$

Hier ist  $L_2(V, W)$  der Raum aller stetigen Bilinearformen  $\varphi : V \times V \rightarrow W$  (zur Isomorphie vgl. Lemma 12.3.10). Falls  $\dim V = n < \infty$  mit Basis  $\{v_1, \dots, v_n\} \subset V$ , so  $L(V, L(V, W)) \cong L(V, W^n) \cong W^{n^2}$ . Falls  $V = \mathbb{R}^n$  und  $W = \mathbb{R}$ , so besitzt  $D^2 f$  eine *Matrixdarstellung*  $D^2 f|_x(y, z) = \langle y, Hz \rangle$  mit einer quadratischen Matrix  $H \in \mathbb{R}^{n \times n}$ . Bezüglich der kanonischen Basen ist dies genau die Hessematrix  $H = H_f(x)$ . Wie bei der Jacobimatrix gilt  $H_f(x) = (D^2 f|_x(e_i, e_j))_{i, j=1, \dots, n} \in \mathbb{R}^{n \times n}$ .

Die höheren Ableitungen  $D^k f$  sind analog Abbildungen von  $V$  nach  $L(V, L(V, \dots, L(V, W) \dots)) \cong L_k(V, W)$ , dem Raum der stetigen  $k$ -Linearformen  $\varphi : V \times \dots \times V \rightarrow W$ . Ist  $\dim V = n < \infty$  mit Basis  $\{v_1, \dots, v_n\} \subset V$ , so gilt wie oben  $L_k(V, W) \cong W^{n^k}$ .

- **Beispiele:**

- $f : \mathbb{R}^d \rightarrow \mathbb{R}, f(x) = \|x\|_2^2, Df|_x h = 2\langle x, h \rangle$  wie oben. Für  $D^2 f$  rechnet man:  $Df|_{x+hy} - Df|_x y = 2\langle h, y \rangle$ , also  $|Df|_{x+hy} - Df|_x y - 2\langle h, y \rangle| = 0$ , Supremum über alle  $0 \neq y \in \mathbb{R}^d$  liefert  $\|Df|_{x+hy} - Df|_x - 2\langle h, \cdot \rangle\|_{L(\mathbb{R}^d, \mathbb{R})} = 0 = o(\|h\|_2)$  und damit  $D^2 f|_x h = 2\langle h, \cdot \rangle$ . Man beachte  $D^2 f|_x h \in L(\mathbb{R}^d, \mathbb{R})$  und  $D^2 f|_x \in L(\mathbb{R}^d, L(\mathbb{R}^d, \mathbb{R})) \cong L_2(\mathbb{R}^d, \mathbb{R})$ ! Und offenbar  $H_f(x) = (D^2 f|_x(e_i, e_j))_{i, j=1, \dots, d} = 2I$ .

- Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, f(x) = \begin{pmatrix} x_1^2 - x_2^2 \\ 2x_1 x_2 \end{pmatrix}$ ; das entspricht  $z \mapsto z^2$  in  $\mathbb{C} \cong \mathbb{R}^2$ . Es ist  $J_f(x) = \begin{pmatrix} 2x_1 & -2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}$  bzw.  $Df|_x h = J_f(x)h$ . Für  $D^2 f$  rechnet man  $Df|_{x+hy} - Df|_x y = (J_f(x+h) - J_f(x))y = 2 \begin{pmatrix} h_1 y_1 - h_2 y_2 \\ h_2 y_1 + h_1 y_2 \end{pmatrix} = J_f(h)y$ , also  $\|Df|_{x+hy} - Df|_x y - J_f(h)y\|_2 = 0$  für jedes  $y \in \mathbb{R}^2$ . Supremum über  $y$  liefert schließlich  $\|Df|_{x+hy} - Df|_x - J_f(h) \cdot \|_{L(\mathbb{R}^2, \mathbb{R}^2)} = 0 = o(\|h\|_2)$ , also  $D^2 f|_x y = J_f(\cdot)y$ . Beachte auch hier wieder  $D^2 f|_x y \in L(\mathbb{R}^2, \mathbb{R}^2)$  und  $D^2 f|_x \in L(\mathbb{R}^2, L(\mathbb{R}^2, \mathbb{R}^2)) \cong L_2(\mathbb{R}^2, \mathbb{R}^2)$ !