

# Extrema unter Nebenbedingungen (Extrema am Rand)

## 1 Problemstellung

**Beispiel 1** (Serie 8, Aufgabe 3 b)) Bestimme die globalen Extrema der folgenden Funktion  $f$  auf dem angegebenen Bereich  $B$ .  $f(x, y, z) := x - y - 2z$ ,

$$B := \{(x, y, z) \mid x \geq 0, y \geq 0, z \geq 0, x^2 + y^2 + z^2 \leq 1\}.$$

**Beispiel 2**  $f(x, y) = x - 2y$ . Bestimme die Extrema von  $f|_{S^1}$ .

Herangehensweise: Extrema in der Richtung des Gradienten von  $f$  suchen. Dazu gibt es zwei Sichtweisen

## 2 Notwendige Bedingung für Extrema

### Lemma 1

$f, g \in C^1(D \subset \mathbb{R}^n, \mathbb{R})$ .

$S := \{\mathbf{x} \in D \mid g(\mathbf{x}) = 0\}$ . Nehme an,  $\mathbf{x}_0 \in S$  ist ein lokales Extremum von  $f|_S$  mit  $\nabla f(\mathbf{x}_0), \nabla g(\mathbf{x}_0) \neq \mathbf{0}$ . Dann gilt

(1)  $\nabla f(\mathbf{x}_0) \parallel \nabla g(\mathbf{x}_0)$

(2)  $\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0)$  für ein  $\lambda \in \mathbb{R}$

(3)  $T_{\mathbf{x}_0}S = T_{\mathbf{x}_0}N$  für  $N = \{\mathbf{x} \in D \mid f(\mathbf{x}) = f(\mathbf{x}_0)\}$  (Niveaumenge von  $f$  zum Niveau  $f(\mathbf{x}_0)$ )

**Bemerkung 1** Die Bedingung  $\nabla g(\mathbf{x}_0) \neq \mathbf{0}$  sorgt dafür, dass wir überhaupt einen Tangentialraum an  $S$  im Punkt  $\mathbf{x}_0$  zur Verfügung haben.

**Bemerkung 2** Es ist klar, dass ein  $\mathbf{x}_0 \in S$  mit  $\nabla f(\mathbf{x}_0) = \mathbf{0}$  auch ein Kandidat für ein Extremum ist, das Lemma behandelt demgegenüber den neuen Fall, bei dem die Extrema Punkte mit  $\nabla f(\mathbf{x}_0) \neq \mathbf{0}$  sind.

**Beispiel 3** Es kann passieren, dass ein Sattelpunkt einer Funktion  $f$  unter einer Randbedingung zu einem Extremum wird:  $f(x, y) = x^2 - y^2$  mit einem

Sattelpunkt in  $\mathbf{0}$ . Mit  $S := \mathbb{R} \cdot \binom{0}{1} = \{(x, y) \in \mathbb{R}^2 | x = 0\}$  hat dann die Funktion  $f|_S$  ein Maximum bei  $\mathbf{0}$ .

### Beweis vom Lemma

(3): Nehme  $\mathbf{v} \in T_{\mathbf{x}_0}S$  mit zugehöriger Kurve  $\gamma : (-\varepsilon, \varepsilon) \rightarrow S$ , so dass  $\gamma(0) = \mathbf{x}_0$  und  $\dot{\gamma} = \mathbf{v}$ . Wenn  $\mathbf{x}_0$  ein Extremum ist, so muss gelten

$$0 = \left. \frac{d}{dt} \right|_{t=0} f \circ \gamma = \underbrace{\langle \nabla f(\gamma(0)), \dot{\gamma}(0) \rangle}_{\mathbf{x}_0 \quad \mathbf{v}}.$$

Da aber  $T_{\mathbf{x}_0}N = \{\mathbf{v} \in D | \langle \mathbf{v}, \nabla f \rangle = 0\}$  bedeutet dies, dass  $\mathbf{v} \in T_{\mathbf{x}_0}N$  und zwar für alle  $\mathbf{v} \in T_{\mathbf{x}_0}S$ . Also  $T_{\mathbf{x}_0}S \subseteq T_{\mathbf{x}_0}N$ . Beide Tangentialräume sind das orthogonale Komplement zu einem 1-dimensionalen Unterraum und haben somit die gleiche Dimension. Es muss also bereits gelten, dass  $T_{\mathbf{x}_0}S = T_{\mathbf{x}_0}N$ .

(1): Beide Gradienten stehen senkrecht auf dieselbe Ebene, müssen also parallel sein.

(2): Ist nur eine Umformulierung von (1). □

## 3 Verallgemeinerung

### Satz 1

$f, g_1, \dots, g_m \in C^1(D, \mathbb{R})$ ,

$S := \{\mathbf{x} \in D | g_1(\mathbf{x}) = \dots = g_m(\mathbf{x}) = 0\}$ ,

Nehme an,  $\mathbf{x}_0$  ist ein Extremum für  $f|_S$  und es gilt  $\nabla f(\mathbf{x}_0) \neq \mathbf{0}$ ,

$\nabla g_j(\mathbf{x}_0) \neq \mathbf{0}$  linear unabhängig  $j = 1, \dots, m$ . Dann gilt

(1)  $\nabla f(\mathbf{x}_0) \in (T_{\mathbf{x}_0}S)^\perp$

(2)  $\nabla f(\mathbf{x}_0) = \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}_0)$

(3)  $T_{\mathbf{x}_0}S \subseteq T_{\mathbf{x}_0}N$  für  $N = \{\mathbf{x} \in D | f(\mathbf{x}) = f(\mathbf{x}_0)\}$

**Bemerkung 3** Die Gleichung in (2) kann mit Hilfe der so genannten *La-grangefunktion*

$$L : D \times \mathbb{R}^m \rightarrow \mathbb{R}$$

$$L(\mathbf{x}, \boldsymbol{\lambda}) := f(\mathbf{x}) - \sum_{i=1}^m \lambda_i g_i(\mathbf{x})$$

umgeschrieben werden zu

$$\begin{cases} \nabla f(\mathbf{x}_0) &= \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}_0) \\ g_j(\mathbf{x}_0) &= 0 \end{cases} \Leftrightarrow \nabla L(\mathbf{x}_0) = \mathbf{0},$$

wobei  $\nabla L$  sowohl die  $x$ - als auch die  $\lambda$ -Ableitungen beinhaltet. Ein Extremum wird dadurch wieder zu einem kritischen Punkt und das Problem ist wieder wie im Fall ohne Randbedingungen darauf zurückgeführt, diese kritischen Punkte zu finden. Die  $\lambda_i$  werden *Lagrange-Multiplikatoren* genannt. Für fixes  $\lambda \in \mathbb{R}^m$  setzen wir auch

$$L_\lambda : D \rightarrow \mathbb{R}, \quad L_\lambda(\mathbf{x}) := L(\mathbf{x}, \lambda).$$

### Beweis

(1): Genau gleich wie (3) im Lemma (1)

(2):  $S = \bigcap_{i=1}^m g_i^{-1}(0)$  ist die Schnittmenge aller Niveaumengen zu den einzelnen Funktionen  $g_j$ .  $T_{\mathbf{x}_0}S$  muss also auf alle  $\nabla g_j(\mathbf{x}_0)$  senkrecht stehen, was bedeutet, dass  $(T_{\mathbf{x}_0}S)^\perp$  von den  $\nabla g_j(\mathbf{x}_0)$  aufgespannt wird. Da gemäss (1)  $\nabla f(\mathbf{x}_0) \in (T_{\mathbf{x}_0}S)^\perp$ , kann  $\nabla f(\mathbf{x}_0)$  als Linearkombination der  $\nabla g_j(\mathbf{x}_0)$  dargestellt werden.

(3): Wir wissen, dass  $(T_{\mathbf{x}_0}N)^\perp = \mathbb{R} \cdot \nabla f(\mathbf{x}_0)$  und dass für zwei Unterräume  $V, W \subseteq \mathbb{R}^n$  gilt:

$$V^\perp \subseteq W^\perp \Rightarrow W \subseteq V.$$

Aus diesen beiden Facts und (1) folgt die Behauptung sofort.  $\square$

### Beispiel 4 (Beispiel 2)

$f(x, y) = x - 2y$ . Bestimme die Extrema von  $f|_{S^1}$ .

Hier ist  $g(x, y) = x^2 + y^2 - 1$  und  $\nabla g(x, y) = 2\begin{pmatrix} x \\ y \end{pmatrix}$ ,  $\nabla f(x, y) = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ . Die notwendige Bedingung lautet

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix} = 2\lambda \begin{pmatrix} x \\ y \end{pmatrix}, \quad g(x, y) = 0.$$

Es folgt sofort  $\begin{pmatrix} x \\ y \end{pmatrix} = \pm \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ .

**Beispiel 5**  $f(x, y, z) = x - 2y + z$  auf  $S = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 = 1, z = 0\}$ . Hier sind

$$\begin{aligned} g_1(x, y, z) &= x^2 + y^2 - 1 & g_2(x, y, z) &= z \\ \nabla g_1(x, y, z) &= \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} & \nabla g_2(x, y, z) &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

und die Bedingung lautet

$$\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 2\lambda_1 \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Also  $(x, y, z) = \frac{1}{\sqrt{5}}(1, -2, 0)$ .

## 4 Hinreichende Bedingung

### Satz 2

$f, g_1, \dots, g_m \in C^{(2)}(D, \mathbb{R})$

$S = \{\mathbf{x} \in D \mid g_1(\mathbf{x}) = \dots = g_m(\mathbf{x}) = 0\}$

$\nabla f(\mathbf{x}_0) \neq \mathbf{0}, \nabla g_j(\mathbf{x}_0) \neq \mathbf{0}$  linear unabhängig.

Nehme zusätzlich an

$$\nabla L(\mathbf{x}_0, \lambda_0) = \mathbf{0},$$

das heisst die notwendige Bedingung dafür, dass bei  $\mathbf{x}_0$  ein lokales Extremum sein kann, ist erfüllt. Dann gilt

$$HL_{\lambda_0}|_{T_{\mathbf{x}_0}S} \begin{cases} \text{positiv definit} & \Rightarrow f(\mathbf{x}_0) \text{ lokales isoliertes Minimum} \\ \text{negativ definit} & \Rightarrow f(\mathbf{x}_0) \text{ lokales isoliertes Maximum} \\ \text{indefinit} & \Rightarrow f(\mathbf{x}_0) \text{ ist kein Extremum.} \end{cases}$$

**Bemerkung 4** Für eine lokale Parametrisierung  $\Phi: \mathbb{R}^{n-m} \rightarrow S$  von  $S$  um  $\mathbf{x}_0$  mit  $\Phi(\mathbf{0}) = x_0$  gilt

$$\begin{aligned} HL_{\lambda_0}|_{T_{\mathbf{x}_0}S} & \text{ positiv/negativ/in- definit} \\ \Leftrightarrow D\Phi^T(\mathbf{0})HL_{\lambda_0}D\Phi(\mathbf{0}) & \text{ positiv/negativ/in- definit,} \end{aligned}$$

da das Bild von  $D\Phi(\mathbf{0})$  gerade der Tangentialraum von  $S$  in  $\Phi(\mathbf{0}) = x_0$  ist:

$$\text{im } D\Phi(\mathbf{0}) = T_{\mathbf{x}_0}S.$$

**Bemerkung 5** Uns stehen zwei Beschreibungen einer Fläche  $S$  zur Verfügung.

1.  $S$  als  $k$ -dimensionale differenzierbare *Untermannigfaltigkeit* von  $\mathbb{R}^n$ :  
Für alle  $\mathbf{x} \in S$  existiert eine Umgebung  $U \subset \mathbb{R}^n$  von  $\mathbf{x}$  und ein Diffeomorphismus  $\psi: U \rightarrow V \subset \mathbb{R}^n$ , so dass  $\psi(U \cap S) \subset \mathbb{R}^k \times \{(0, \dots, 0) \in \mathbb{R}^{n-k}\}$
2.  $S$  als Niveaumenge einer Abbildung  $\mathbf{g} = (g_1, \dots, g_{n-k}): D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^{n-k}$ :  $S = \{\mathbf{x} \in D \mid g_1(\mathbf{x}) = \dots = g_{n-k}(\mathbf{x}) = 0\} = \mathbf{g}^{-1}(0) = \bigcap_{i=1}^{n-k} g_i^{-1}(0)$ , mit  $\nabla g_i$  linear unabhängig, d.h.  $D\mathbf{g}$  hat maximalen Rang.

Wird  $S$  auf eine dieser beiden Arten beschrieben, so wird  $S$  automatisch auch durch die andere beschrieben:

Für einen Diffeomorphismus  $\psi: U \rightarrow V \subset \mathbb{R}^n$  kann ein zugehöriges  $\mathbf{g}: D \rightarrow \mathbb{R}^{n-k}$  definiert werden als

$$\mathbf{g}(\mathbf{x}) := (\psi^{k+1}(\mathbf{x}), \dots, \psi^n(\mathbf{x})).$$

Umgekehrt existiert lokal (Satz über implizite Funktionen) zu gegebener Funktion  $\mathbf{g}: D \rightarrow \mathbb{R}^{n-k}$  unter den obigen Voraussetzungen für  $D\mathbf{g}$  (maximaler Rang) eine Funktion  $\Phi: \mathbb{R}^k \rightarrow \mathbb{R}^{n-k}$ , so dass (ordne  $x^1, \dots, x^n$  so,

dass  $\det(\partial_1 \mathbf{g}, \dots, \partial_{n-k} \mathbf{g}) \neq 0$

$$\mathbf{g}(x^1, \dots, x^k, \Phi(x^1, \dots, x^k)) = 0.$$

Der Diffeomorphismus  $\psi : U \rightarrow V \subset \mathbb{R}^n$  kann somit gewählt werden als

$$\psi(\mathbf{x}_1, \mathbf{x}_2) := (\mathbf{x}_1, \mathbf{x}_2 - \Phi(\mathbf{x}_1)),$$

für  $\mathbf{x}_1 \in \mathbb{R}^k$  und  $\mathbf{x}_2 \in \mathbb{R}^{n-k}$ . Auf  $S$  ist bekanntlich  $\mathbf{g}(\mathbf{x}_1, \mathbf{x}_2) = 0$  und somit  $\mathbf{x}_2 = \Phi(\mathbf{x}_1)$ . Also  $\psi(U \cap S) \subset \mathbb{R}^k \times \{\mathbf{0}\}$ .

### Beweis

Nach Bemerkung 5 existiert eine lokale Parametrisierung  $\Phi : \mathbb{R}^{n-m} \rightarrow S$  von  $S$  um  $\mathbf{x}_0$  mit  $\Phi(\mathbf{0}) = \mathbf{x}_0$ .

Dann hat  $f|_S \equiv L_{\lambda_0}|_S$  ein Extremum bei  $\mathbf{x}_0$  genau dann, wenn  $L_{\lambda_0} \circ \Phi$  ein Extremum bei  $\mathbf{0}$  hat.

Die Taylor-Entwicklung von  $L_{\lambda_0} \circ \Phi$  um  $\mathbf{0}$  ist

$$L_{\lambda_0} \circ \Phi(\boldsymbol{\xi}) = \underbrace{L_{\lambda_0} \circ \Phi(\mathbf{0})}_{L_{\lambda_0}(\mathbf{x}_0)} + D(L_{\lambda_0} \circ \Phi)(\mathbf{0})\boldsymbol{\xi} + \frac{1}{2}\boldsymbol{\xi}^T H(L_{\lambda_0} \circ \Phi)(\mathbf{0})\boldsymbol{\xi} + R(\boldsymbol{\xi}),$$

mit einem Rest  $R(\boldsymbol{\xi})$ , für den gilt

$$\lim_{\|\boldsymbol{\xi}\| \rightarrow 0} \frac{R(\boldsymbol{\xi})}{\|\boldsymbol{\xi}\|^2} = 0.$$

Da  $\nabla L_{\lambda_0}(\mathbf{x}_0) = 0$ , verschwindet der Term linear in  $\boldsymbol{\xi}$  (hätten wir  $f$  statt  $L_{\lambda_0}$  gewählt, wäre das nicht der Fall):

$$D(L_{\lambda_0} \circ \Phi)(\mathbf{0}) = DL(\Phi(\mathbf{0}))D\Phi(\mathbf{0}) = \nabla L_{\lambda_0}(\mathbf{x}_0)^T D\Phi(\mathbf{0}) = 0.$$

Weiter gilt

$$\begin{aligned} H(L_{\lambda_0} \circ \Phi)(\mathbf{0}) &= D^2(L_{\lambda_0} \circ \Phi)(\mathbf{0}) = D[(DL_{\lambda_0}) \circ \Phi D\Phi](\mathbf{0}) \\ &= D\Phi^T(\mathbf{0})HL_{\lambda_0}(\mathbf{x}_0)D\Phi(\mathbf{0}) + \underbrace{DL_{\lambda_0}(\mathbf{x}_0)}_0 D^2\Phi(\mathbf{0}). \end{aligned}$$

Zusammengefasst, mit der Abkürzung  $M := D\Phi^T(\mathbf{0})HL_{\lambda_0}(\mathbf{x}_0)D\Phi(\mathbf{0})$  und unter Verwendung von  $f|_S \equiv L_{\lambda_0}|_S$ :

$$f \circ \Phi(\boldsymbol{\xi}) = f(\mathbf{x}_0) + \frac{1}{2}\boldsymbol{\xi}^T M\boldsymbol{\xi} + R(\boldsymbol{\xi}).$$

Definiere  $m_- := \inf_{\|\boldsymbol{\xi}\|=1} \langle \boldsymbol{\xi}, M\boldsymbol{\xi} \rangle$  und  $m_+ := \sup_{\|\boldsymbol{\xi}\|=1} \langle \boldsymbol{\xi}, M\boldsymbol{\xi} \rangle$ , Falls  $M$  positiv definit ist, so ist  $m_- > 0$  und

$$f \circ \Phi(\boldsymbol{\xi}) \geq f(\mathbf{x}_0) + \|\boldsymbol{\xi}\|^2 \left( \frac{1}{2}m_- + \frac{R(\boldsymbol{\xi})}{\|\boldsymbol{\xi}\|^2} \right)$$

und falls  $M$  negativ definit ist, gilt  $m_+ < 0$  und

$$f \circ \Phi(\xi) \leq f(\mathbf{x}_0) + \|\xi\|^2 \left( \frac{1}{2} m_+ + \frac{R(\xi)}{\|\xi\|^2} \right).$$

In beiden Fällen gilt  $\lim_{\|\xi\| \rightarrow 0} \frac{R(\xi)}{\|\xi\|^2} = 0$  und es existiert ein  $r > 0$ , so dass  $\left| \frac{R(\xi)}{\|\xi\|^2} \right| < \frac{m_{\pm}}{2}$  für alle  $\|\xi\| < r$ . Auf  $B_r(0) \setminus \{\mathbf{0}\}$  ist also im einen Fall  $f \circ \Phi > f(\mathbf{x}_0)$ , im anderen Fall  $f \circ \Phi < f(\mathbf{x}_0)$  und es liegt ein lokales isoliertes Minimum bzw. Maximum vor.

Falls  $M$  indefinit ist, so existieren  $\xi_{\pm}$ , so dass  $\xi_+^T M \xi_+ =: a_+ > 0$  und  $\xi_-^T M \xi_- =: a_- < 0$ . Mit obiger Argumentation existiert ein  $r > 0$ , so dass für alle  $t \in (-r, r) \setminus \{0\}$  gilt:

$$\begin{aligned} f \circ \Phi(t\xi_+) &> 0 \\ f \circ \Phi(t\xi_-) &< 0. \end{aligned}$$

Das bedeutet, in jeder Umgebung von  $\mathbf{0}$  existieren sowohl positive, als auch negative Werte von  $f \circ \Phi$ .  $\square$

**Beispiel 6** (aus Zorich)  $f(x, y) = x^2 + y^2$ ,  $0 = g(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$ ,  $0 < a < b$  (Ellipsengleichung).

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y) = x^2 \left( 1 - \frac{\lambda}{a^2} \right) + y^2 \left( 1 - \frac{\lambda}{b^2} \right) - \lambda.$$

Die Bedingung lautet

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} = \frac{\partial L}{\partial \lambda} = 0,$$

was auf die Punkte  $(x, y, \lambda) = (\pm a, 0, a^2)$  und  $(x, y, \lambda) = (0, \pm b, b^2)$  führt.

$$\frac{1}{2} \mathbf{v}^T H L_{\lambda} \mathbf{v} = \left( 1 - \frac{\lambda}{a^2} \right) (v^1)^2 + \left( 1 - \frac{\lambda}{b^2} \right) (v^2)^2.$$

Da in den Punkten  $(x, y, \lambda) = (\pm a, 0, a^2)$  der Tangentialraum jeweils vertikale Geraden sind ( $v^1 = 0$ ), ist dort  $H L_{\lambda}|_{T_{x_0}} = \left( 1 - \frac{a^2}{b^2} \right) (v^2)^2 > 0$  und wir haben dort Minima. Bei  $(x, y, \lambda) = (0, \pm b, b^2)$  sind die Tangentialräume horizontale Geraden ( $v^2 = 0$ ) und  $H L_{\lambda}|_{T_{x_0}} = \left( 1 - \frac{b^2}{a^2} \right) (v^1)^2 < 0$ . Dort liegen also die Maxima.

## 5 Nachtrag: Klassifikation der Extrema mit Hilfe der Determinante

### Lemma 2

Ist in Satz 2  $S$  eine 1-dim Untermannigfaltigkeit, d.h.  $m = n - 1$ , so gilt

$$(-1)^{n-1} \det(HL) < 0 \Rightarrow f(\mathbf{x}_0) \text{ ist ein isoliertes lokales Maximum}$$

$$(-1)^{n-1} \det(HL) > 0 \Rightarrow f(\mathbf{x}_0) \text{ ist ein isoliertes lokales Minimum.}$$

Dabei ist  $HL$  die uneingeschränkte Hessematrix (mit Ableitungen nach  $\lambda_i$ ).

### Definition 1 (Verallgemeinertes Kreuzprodukt)

Sei  $\{\mathbf{e}_i\}_{1 \leq i \leq n}$  die Standardbasis von  $\mathbb{R}^n$ . Für  $n \in \mathbb{N}$  definiere den zur Determinante dualen Tensor oder das verallgemeinerte Kreuzprodukt

$$\begin{aligned} \epsilon : (\mathbb{R}^n)^{n-1} &\rightarrow \mathbb{R}^n \\ \epsilon(\mathbf{v}_1, \dots, \mathbf{v}_{n-1}) &:= \sum_{i=1}^n \det(\mathbf{v}_1, \dots, \mathbf{v}_{n-1}, \mathbf{e}_i) \mathbf{e}_i \end{aligned}$$

**Beispiel 7** Für  $n = 3$  ist

$$\epsilon(\mathbf{v}, \mathbf{w}) = \mathbf{v} \times \mathbf{w},$$

da  $\det(\mathbf{v}, \mathbf{w}, \mathbf{e}_i) = [\mathbf{v} \times \mathbf{w}]^i$ .

Für  $n = 2$  gilt mit  $\mathbf{v} = \begin{pmatrix} v^1 \\ v^2 \end{pmatrix}$

$$\epsilon(\mathbf{v}) = \begin{pmatrix} -v^2 \\ v^1 \end{pmatrix}.$$

**Bemerkung 6** Der Vektor  $\epsilon(\mathbf{v}_1, \dots, \mathbf{v}_{n-1})$  steht senkrecht auf  $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ :

$$\begin{aligned} \langle \epsilon(\mathbf{v}_1, \dots, \mathbf{v}_{n-1}), \mathbf{v}_k \rangle &= \sum_{i=1}^n \det(\mathbf{v}_1, \dots, \mathbf{v}_{n-1}, \mathbf{e}_i) \underbrace{\langle \mathbf{e}_i, \mathbf{v}_k \rangle}_{v_k^i} \\ &= \sum_{i=1}^n \det(\mathbf{v}_1, \dots, \mathbf{v}_{n-1}, \mathbf{e}_i v_k^i) \\ &= \det \left( \mathbf{v}_1, \dots, \mathbf{v}_{n-1}, \underbrace{\sum_{i=1}^n \mathbf{e}_i v_k^i}_{\mathbf{v}_k} \right) \\ &= 0. \end{aligned}$$

### Beweis (Lemma)

Die Hesse-Matrix  $HL$  hat beim kritischen Punkt  $(\mathbf{x}_0, \boldsymbol{\lambda}_0)$  die Form

$$HL(\mathbf{x}_0, \boldsymbol{\lambda}_0) = \begin{pmatrix} 0 & \mathbf{Dg}(\mathbf{x}_0) \\ \mathbf{Dg}^T(\mathbf{x}_0) & HL_{\boldsymbol{\lambda}_0}(\mathbf{x}_0) \end{pmatrix},$$

wobei  $D\mathbf{g}^T = (\nabla g_1, \dots, \nabla g_{n-1})$  eine  $(n-1) \times n$ -Matrix ist.

Die Determinante von  $HL(\mathbf{x}_0, \boldsymbol{\lambda}_0)$  muss dann in diesem Fall ein homogenes Polynom vom Grad 1 in den Einträgen von  $HL_{\boldsymbol{\lambda}_0}(\mathbf{x}_0)$ , denn

$$\frac{\partial}{\partial(HL_{\boldsymbol{\lambda}_0})_{ij}} HL = (-1)^{2(n-1)+i+j} \det(HL^{(n-1+i, n-1+j)}),$$

wobei  $HL^{(k,l)}$  den Minor von  $HL$  bezeichnet bei dem die  $k$ -te Spalte und die  $l$ -te Zeile gestrichen werden<sup>1</sup>.

Diese Minoren haben die Form

$$HL^{(k,l)} = \begin{pmatrix} 0 & A \\ B & C \end{pmatrix}$$

und deren Determinanten sind daher (alle beteiligten Matrizen  $A, B, C$  sind  $(n-1) \times (n-1)$ -Matrizen)

$$\det(HL^{(k,l)}) = (-1)^{n-1} \det(A) \det(B).$$

Wenn  $k = n-1+i$ , dann entsteht  $A$  aus  $D\mathbf{g}^T$  durch streichen der  $i$ -ten Zeile und wir haben (Entwicklung der Determinante nach der letzten Spalte, von rechts nach links angewendet):

$$\det(A) = (-1)^{n+i} \det(D\mathbf{g}^T, \mathbf{e}_i) =: t^i.$$

Hier haben wir

$$\mathbf{t} := \boldsymbol{\epsilon}(\nabla g_1, \dots, \nabla g_{n-1})$$

eingeführt. Analog folgt mit  $l = n-1+j$

$$\det(B) = (-1)^{n+j} t^j.$$

Somit ist

$$\frac{\partial}{\partial(HL_{\boldsymbol{\lambda}_0})_{ij}} HL = (-1)^{n-1} t^i t^j.$$

Da in der Determinante von  $HL$  keine Summanden ohne Faktoren  $[HL_{\boldsymbol{\lambda}_0}]_{ij}$  auftreten muss dann gelten:

$$\begin{aligned} \det(HL(\mathbf{x}_0, \boldsymbol{\lambda}_0)) &= (-1)^{n-1} \sum_{\substack{i=1 \\ j=1}}^n t^i [HL_{\boldsymbol{\lambda}_0}]_{ij} t^j \\ &= (-1)^{n-1} \mathbf{t}^T HL_{\boldsymbol{\lambda}_0} \mathbf{t}. \end{aligned}$$

Da der Vektor  $\mathbf{t} = \boldsymbol{\epsilon}(\nabla g_1, \dots, \nabla g_{n-1})$  senkrecht auf alle Gradienten steht, muss dieser im Tangentialraum liegen. Mit Satz 2 folgt jetzt die Behauptung.  $\square$

---

<sup>1</sup>Dies folgt aus der Cramerschen Regel zur Entwicklung der Determinante nach einer Spalte oder Zeile. Siehe auch *Zusammenhang Volumen - Oberfläche*