

8 Mehrdimensionale Differentialrechnung¹

8.1 Matrizen und quadratische Formen

8.1.1 Definition: Matrizen

$$A := \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & & & \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

ist eine $m \times n$ -Matrix mit $a_{\mu,\nu} \in \mathbb{R}$.

m ist die Zahl der Zeilen: $(a_{\mu,1} \ a_{\mu,2} \ \dots \ a_{\mu,n})$ ist eine $1 \times n$ -Matrix.

n ist die Zahl der Spalten: $\begin{pmatrix} a_{1,\nu} \\ a_{2,\nu} \\ \vdots \\ a_{m,\nu} \end{pmatrix}$ ist eine $m \times 1$ -Matrix.

8.1.2 Rechenregeln für Matrizen

1. Für $C := A + B$ gilt: $c_{\mu,\nu} := a_{\mu,\nu} + b_{\mu,\nu}$.
2. Für $C := \lambda A$ gilt: $c_{\mu,\nu} := \lambda a_{\mu,\nu}$.
3. Sei A eine $m \times n$ -Matrix und B eine $n \times p$ -Matrix. Dann ist $C := A \cdot B$ eine $m \times p$ -Matrix mit

$$c_{\mu,\varrho} = \sum_{\nu=1}^n a_{\mu,\nu} b_{\nu,\varrho}$$

für $\mu = 1, \dots, m$ und $\varrho = 1, \dots, p$.

8.1.3 Transponierte Matrizen

Die Transponierte von A (siehe oben) ist

$$A^T := \begin{pmatrix} a_{1,1} & \dots & a_{m,1} \\ a_{1,2} & & a_{m,2} \\ \vdots & & \vdots \\ a_{1,n} & \dots & a_{m,n} \end{pmatrix}.$$

Wenn $x \in \mathbb{R}^n$ ist, wird oft

$$x^T = (x_1, \dots, x_n)$$

¹Version 184 vom 13. Februar 2006

geschrieben. Dies ist als Spaltenvektor gemeint:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

8.1.4 Skalarprodukt

$$x^T y = (x_1, \dots, x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \dots + x_n y_n = x \cdot y$$

ist das Skalarprodukt im $\mathbb{R}^n$.

8.1.5 Matrizenorm

Sei A eine $m \times n$ -Matrix, $\|\cdot\|_n$ eine Norm im $\mathbb{R}^n$ und $\|\cdot\|_m$ eine Norm im $\mathbb{R}^m$. Dann ist

$$\|A\| := \sup\{\|A \cdot x\|_m : x \in \mathbb{R}^n, \|x\|_n = 1\}$$

die von $\|\cdot\|_n$ und $\|\cdot\|_m$ erzeugte Matrixnorm.

Beachte:

$$\|Ax\|_m \leq \|A\| \cdot \|x\|_n \text{ für alle } x \in \mathbb{R}^n.$$

8.1.6 Beispiele

(1) Beidesmal die Euklidnorm: Zeige (als Aufgabe)

$$\|A\| \leq \sqrt{\sum_{\substack{\mu=1,\dots,m \\ \nu=1,\dots,n}} a_{\mu,\nu}^2}.$$

(2) Beidesmal Maximumnorm:

Für $y = Ax$ ist

$$y_\mu = \sum_{\nu=1}^n a_{\mu,\nu} x_\nu.$$

Es gilt dann

$$\begin{aligned} |y_\mu| &\leq \sum_{\nu=1}^n |a_{\mu,\nu}| \cdot \underbrace{|x_\nu|}_{\leq \|x\|_\infty} \\ &\leq \left(\sum_{\nu=1}^n |a_{\mu,\nu}| \right) \cdot \|x\|_\infty \end{aligned}$$

Also ist

$$\|y\|_\infty \leq \left(\max_{\mu=1}^m \sum_{\nu=1}^n |a_{\mu,\nu}| \right) \cdot \|x\|_\infty$$

Es ist damit

$$\|A\| \stackrel{(*)}{\leq} \max_{\mu=1}^m \sum_{\nu=1}^n |a_{\mu,\nu}|.$$

Zeige, daß für $(*)$ Gleichheit gilt:

Sei

$$\max_{\mu=1}^m \sum_{\nu=1}^n |a_{\mu,\nu}| = \sum_{\nu=1}^n |a_{\varrho,\nu}|$$

Setze

$$x_\nu = \begin{cases} 1 & \text{wenn } a_{\varrho,\nu} \geq 0 \\ -1 & \text{wenn } a_{\varrho,\nu} < 0 \end{cases}.$$

Dann ist

$$\left| \sum_{\nu=1}^n a_{\varrho,\nu} x_\nu \right| = \sum_{\nu=1}^n |a_{\varrho,\nu}| \cdot \underbrace{\|x\|_\infty}_{=1}$$

und damit

$$\|A\| \stackrel{\odot}{\geq} \sum_{\nu=1}^n |a_{\varrho,\nu}|.$$

Aus $\odot$ und $(*)$ folgt insgesamt die Gleichheit bei $(*)$.

8.1.7 Definition: Quadratische Formen

Sei A eine symmetrische $n \times n$ -Matrix ($A = A^T$, $a_{\mu,\nu} = a_{\nu,\mu}$). Dann heißt

$$Q_A: \mathbb{R}^n \rightarrow \mathbb{R}, \quad Q_A(x) := x^T A x = \sum_{\mu,\nu=1}^n a_{\mu,\nu} x_\mu x_\nu$$

quadratische Form.

8.1.8 Bemerkung

Wenn $A^T = A$ ist, sind alle Eigenwerte $\lambda_1, \dots, \lambda_n$ reell. Man kann die Eigenvektoren $x^{(1)}, \dots, x^{(n)}$ orthonormal wählen, d. h. es gilt

$$x^{(j)} \cdot x^{(k)} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}.$$

$x^{(1)}, \dots, x^{(n)}$ bilden eine Orthonormalbasis des $\mathbb{R}^n$.

Sei nun $x \in \mathbb{R}^n$. Dann gilt:

$$x = \sum_{j=1}^n \xi_j x^{(j)} \quad \text{und} \quad Ax = \sum_{j=1}^n \xi_j \lambda_j x^{(j)}.$$

Damit ist

$$x^T A x = x \cdot (Ax) = \sum_{j=1}^n \lambda_j \xi_j^2.$$

8.1.9 Definition: positiv/negativ (semi-)definit

A heißt

$$\left. \begin{array}{l} \text{positiv definit} \\ \text{negativ definit} \\ \text{positiv semidefinit} \\ \text{negativ semidefinit} \\ \text{indefinit} \end{array} \right\} \Longleftrightarrow x^T A x \begin{cases} > 0 & \text{für } x \neq 0 \\ < 0 & \text{für } x \neq 0 \\ \geq 0 & \text{für } x \neq 0 \\ \leq 0 & \text{für } x \neq 0 \\ \text{sonst} \end{cases}$$

$$\Longleftrightarrow \begin{cases} \text{alle } \lambda_j > 0 \\ \text{alle } \lambda_j < 0 \\ \text{alle } \lambda_j \geq 0 \\ \text{alle } \lambda_j \leq 0 \\ \text{mind. ein } \lambda_j > 0, \text{ ein } \lambda_j < 0 \end{cases}$$

8.1.10 Bemerkungen

A ist positiv definit („ $A > 0$ “)

$$\Longleftrightarrow \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} > 0 \text{ für } k = 1, 2, \dots, n,$$

bzw. negativ definit („ $A < 0$ “)

$$\Longleftrightarrow (-1)^k \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} > 0 \text{ für } k = 1, 2, \dots, n.$$

Es gilt $A > 0 \Longleftrightarrow (-A) < 0$.

Ohne Beweis.

Insbesondere ist die Matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ genau dann positiv definit, wenn $a > 0$ und $ac > b^2$, bzw. genau dann negativ definit, wenn $a < 0$ und $ac > b^2$ ist.

8.1.11 Beispiele

(1)

$$A = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \text{ ist positiv semidefinit, d. h. } A \geq 0.$$

$$x^T A x = \sum_{\mu, \nu=1}^n x_\mu x_\nu = \sum_{\mu=1}^n x_\mu \cdot \sum_{\nu=1}^n x_\nu = \left(\sum_{\nu=1}^n x_\nu \right)^2 \geq 0$$

$$\text{„} = 0 \text{“} \Longleftrightarrow \sum_{\nu=1}^n x_\nu = 0$$

(2)

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

$$\begin{aligned} Q_A(x) &= x_1^2 + 2x_2^2 + 3x_3^2 - 2x_1x_2 + 2x_2x_3 \\ &= (x_1 - x_2)^2 + x_2^2 + 3x_3^2 + 2x_2x_3 \\ &= (x_1 - x_2)^2 + (x_2 + x_3)^2 + 2x_3^2 \geq 0 \end{aligned}$$

$$\begin{aligned} „= 0“ &\Leftrightarrow x_3 = 0 \wedge x_2 + x_3 = 0 \wedge x_1 - x_2 = 0 \\ &\Leftrightarrow x = 0 \in \mathbb{R}^3 \Rightarrow A > 0. \end{aligned}$$

8.1.12 Definition: Landau-Symbole

Sei $E \subseteq \mathbb{R}^n$, ξ Häufungspunkt von E , $\varphi: E \rightarrow \mathbb{R}^m$ und $\psi: E \rightarrow \mathbb{R}$.

- (a) $\varphi(x) = O(\psi(x))$ für $x \rightarrow \xi$ bedeutet:
Es gibt ein $C > 0$ und ein $\delta > 0$ mit

$$\|\varphi(x)\| \leq C\psi(x) \text{ in } x \in E, 0 < \|x - \xi\| < \delta.$$

- (b) $\varphi(x) = o(\psi(x))$ für $x \rightarrow \xi$ bedeutet:
Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ mit

$$\|\varphi(x)\| \leq \varepsilon\psi(x) \text{ in } x \in E, 0 < \|x - \xi\| < \delta.$$

8.1.13 Beispiele

1. Sei $\varphi(x) = x^T Ax$.
Es gilt: $\phi(x) = O(\|x\|^2)$ für $x \rightarrow 0$.
Äquivalent ist: $|x^T Ax| \leq C\|x\|^2$ für $\|x\| < \delta$.

$$\begin{aligned} |x^T Ax|^2 &= |x \cdot (Ax)|^2 \\ &\stackrel{\text{CSU}}{\leq} \|x\|^2 \cdot \|Ax\|^2 \\ &\leq \|x\|^2 \cdot (C\|x\|^2) = C^2\|x\|^4 \end{aligned}$$

$$\Rightarrow \|x^T Ax\| \leq C\|x\|^2 \text{ für } x \in \mathbb{R}^n.$$

2. Sei $f: E \rightarrow \mathbb{R}^m$ stetig in $\xi \in E$.
 $\Leftrightarrow$ Zu $\varepsilon > 0$ ex. ein $\delta > 0$ mit

$$\begin{aligned} \|f(x) - f(\xi)\| &< \varepsilon \text{ für } \|x - \xi\| < \delta, x \in E \\ \Leftrightarrow f(x) - f(\xi) &=: \tilde{\varphi}(x) \text{ mit } \tilde{\varphi}(x) \rightarrow 0 \text{ für } x \rightarrow \xi \\ \Leftrightarrow f(x) - f(\xi) &= o(1) \text{ für } x \rightarrow \xi \\ \Leftrightarrow f(x) &= f(\xi) + o(1) \text{ für } x \rightarrow \xi \end{aligned}$$

8.1.14 Regeln für Landau-Symbole

(Die Regeln unter (a) und (b) gelten sowohl für O als auch für o .)

(a) Sei $\varphi_j(x) = O(\psi(x))$ für $x \rightarrow \xi$ und $j = 1, \dots, k$.

Dann ist

$$\sum_{j=1}^k \varphi_j(x) = O(\psi(x)) \text{ für } x \rightarrow \xi.$$

Anwendungsbeispiel:

Seien f_1 und f_2 stetig in ξ . Dann ist $f_1 + f_2$ stetig in ξ ,

denn: $f_j(x) = f_j(\xi) + o(1)$ für $j = 1, 2$.

$$f_1(x) + f_2(x) = f_1(\xi) + f_2(\xi) + o(1)$$

(b) Sei $\varphi_j(x) = O(\psi(x))$ für $x \rightarrow \xi$ und $j = 1, \dots, k$.

Dann gilt

$$\varphi(x) = \max_1^k \|\varphi_j(x)\| \Rightarrow \varphi(x) = O(\psi(x)) \text{ für } x \rightarrow \xi$$

(c) Gilt $\varphi(x) = o(\psi(x))$ für $x \rightarrow \xi$, dann ist auch $\varphi(x) = O(\psi(x))$ für $x \rightarrow \xi$. Dies gilt i. A. aber nicht umgekehrt.

8.1.15 Beispiel

Sei $f \in C^k(-r, r)$. Dann ist für $x \rightarrow 0$

$$f(x) = \sum_{j=0}^k \frac{f^{(j)}(0)}{j!} x^j + \begin{cases} o(|x|^k) & , \text{ falls } f \in C^k(-r, r) \\ O(|x|^{k+1}) & , \text{ falls } f \in C^{k+1}(-r, r) \end{cases}$$

8.2 Differenzierbare Funktionen

Für diesen Abschnitt ist $D \subseteq \mathbb{R}^n$ ein Gebiet.

Funktionen $D \rightarrow \mathbb{R}^m$ werden üblicherweise mit $f, g, \psi, \varphi, \dots$ bezeichnet.

Funktionen $D \rightarrow \mathbb{R}$ werden üblicherweise mit $u, v, w, \dots$ bezeichnet.

$\|\cdot\|$ ist die Euklidnorm.

8.2.1 Definition: Differenzierbarkeit

$f: D \rightarrow \mathbb{R}^m$ heißt *differenzierbar* in $x \in D$, wenn es eine $m \times n$ -Matrix A gibt mit

$$f(x+h) - f(x) - Ah = o(\|h\|) \text{ für } h \rightarrow 0 \ (h \in \mathbb{R}^n).$$

Äquivalent hierzu ist:

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - A \cdot h}{\|h\|} = 0.$$

Interpretation:

$$f(x+h) = \underbrace{f(x) + A \cdot h}_{\otimes} + \underbrace{o(\|h\|)}_{\odot}$$

$\otimes$: Affine Funktion von h .

$\odot$: Fehlerterm: $< \varepsilon \cdot \|h\|$ für $\|h\| < \delta$.

8.2.2 Bemerkungen

(a) Sei $n = m = 1$:

$$f(x+h) - f(x) = a \cdot h + o(|h|)$$

$$\iff \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = a =: f'(x)$$

(b) Man nennt A die Ableitung von f und schreibt $f'(x) = A$. Andere Schreibweise:

$$A = \frac{\partial f}{\partial x} \text{ ist die Jacobi-Matrix.}$$

8.2.3 Beispiele

(1) Sei $f(x) = c$ konstant mit $c \in \mathbb{R}^m$.

Dann ist $f'(x) = 0$ die $m \times n$ -Nullmatrix,
denn $f(x+h) - f(x) - 0 \cdot h = 0 = o(\|h\|)$.

(2) Sei $f(x) = Ax$ mit einer $m \times n$ -Matrix A .

$$f(x+h) - f(x) = A \cdot h = A \cdot h + o(\|h\|),$$

also ist $f'(x) = A$ für $x \in \mathbb{R}^n$.

(3) Sei $f(x) = x^T A x$ mit einer symmetrischen $n \times n$ -Matrix A .

Es ist $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Also ist $f'(x)$ eine $1 \times n$ -Matrix, ein Zeilenvektor.

$$\begin{aligned} f(x+h) - f(x) &= (x+h)^T A (x+h) - x^T A x \\ &= x^T A h + h^T A x + h^T A h \\ &= (x^T A)h + (x^T A^T)h + h^T A h \\ &= (x^T A)h + (x^T A)h + h^T A h \end{aligned}$$

Da $|h^T A h| = O(\|h\|^2) = o(\|h\|)$ für $h \rightarrow 0$ ist, ist

$$f'(x) = x^T A + x^T A = 2x^T A = 2(Ax)^T.$$

8.2.4 Satz

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : D \rightarrow \mathbb{R}^m$$

ist genau dann in x differenzierbar, wenn alle $f_\mu: D \rightarrow \mathbb{R}$ in x differenzierbar sind. Es gilt:

$$f'(x) = \begin{pmatrix} f'_1(x) \\ \vdots \\ f'_m(x) \end{pmatrix}.$$

Beweis: Sei $A = f'(x)$ und $a^\mu := f'_\mu(x)$.

„ $\Rightarrow$ “: $f'(x)$ existiere nach Voraussetzung. Dann ist

$$\begin{aligned} |f_\mu(x+h) - f_\mu(x) - a^\mu h| &\leq \|f(x+h) - f(x) - A \cdot h\| \\ &= o(\|h\|) \text{ für } h \rightarrow 0, \mu = 1, \dots, m \end{aligned}$$

Also ist $f'_\mu(x) = a^\mu$.

„ $\Leftarrow$ “: $f'_\mu(x) = a^\mu$ existieren nach Voraussetzung. Setze

$$A = \begin{pmatrix} a^1 \\ \vdots \\ a^m \end{pmatrix} \text{ ist } m \times n\text{-Matrix}$$

$$\begin{aligned} \|f(x+h) - f(x) - A \cdot h\| &\leq \sum_{\mu=1}^m \underbrace{|f_\mu(x+h) - f_\mu(x) - a^\mu \cdot h|}_{=o(\|h\|)} \\ &= o(\|h\|) \text{ für } h \rightarrow 0 \end{aligned}$$

Also ist $f'(x) = A$.

8.2.5 Satz: Die Ableitung ist eindeutig

Die Ableitung ist eindeutig bestimmt. Ist f differenzierbar in x , so ist f auch stetig in x .

Beweis:

Eindeutigkeit: Sei

$$\begin{aligned} f(x+h) &= f(x) + A \cdot h + o(\|h\|) \\ f(x+h) &= f(x) + B \cdot h + o(\|h\|) \end{aligned}$$

Subtraktion der beiden Gleichungen liefert:

$$\mathbb{R}^m \ni 0 = (A - B) \cdot h + o(\|h\|) \text{ für } h \rightarrow 0, h \in \mathbb{R}^n$$

Folgt daraus $A = B$?

Setze $h = t \cdot e$ mit $t \in \mathbb{R}$, $e \in \mathbb{R}^n$ und $\|e\| = 1$.

Es gilt: $h \rightarrow 0 \iff t \rightarrow 0$.

$$t(A - B)e = o(|t|) \text{ für } t \rightarrow 0$$

$$(A - B)e = \frac{1}{t} o(|t|) = o(1) \text{ für } t \rightarrow 0$$

Für $t \rightarrow 0$ gilt dann:

$$(A - B) \cdot e = 0 \text{ für alle } e \in \mathbb{R}^n \text{ mit } \|e\| = 1.$$

Also ist $A - B = 0$, d. h. $A = B$.

Stetigkeit:

$$\begin{aligned}
\|f(x+h) - f(x)\| &= \|A \cdot h + o(\|h\|)\| \\
&\leq \|A\| \cdot \|h\| + \varepsilon \|h\| \text{ für } \|h\| < \delta \text{ bei vorgegebenem } \varepsilon \\
&= \text{const} \cdot \|h\| \text{ für } \|h\| < \delta
\end{aligned}$$

Setze nun $y = x + h$. Dann ist

$$\|f(y) - f(x)\| \leq \text{const} \cdot \|x - y\| \text{ für } \|x - y\| < \delta.$$

Also ist f stetig.

8.2.6 Regeln für differenzierbare Funktionen

Sind $f, g: D \rightarrow \mathbb{R}^m$ differenzierbar in x , so auch $f + g$, $\lambda \cdot f$ ($\lambda \in \mathbb{R}$) und $f \cdot g: D \rightarrow \mathbb{R}$ (Skalarprodukt). Im Punkt x gilt:

$$\begin{aligned}
(f + g)' &= f' + g' \\
(\lambda f)' &= \lambda f' \\
(f \cdot g)' &= f^T g' + g^T f'
\end{aligned}$$

Vergleiche mit den eindimensionalen Regeln (4.1.3, Seite 70).

Beweis: Als Aufgabe.

8.2.7 Kettenregel für differenzierbare Funktionen

Zur Wiederholung zuerst die Kettenregel im 1-dimensionalen: Seien $f, g: \mathbb{R} \rightarrow \mathbb{R}$. Dann ist

$$\frac{d}{dx} f(g(x)) = f'(g(x)) \cdot g'(x).$$

Sei nun $g: D \subseteq \mathbb{R}^n \rightarrow G \subseteq \mathbb{R}^m$ differenzierbar in $x \in D$ und $f: G \rightarrow \mathbb{R}^p$ sei differenzierbar in $y = g(x)$. Dann ist $f \circ g: D \rightarrow \mathbb{R}^p$ differenzierbar in x mit

$$(f \circ g)'(x) = f'(y) \cdot g'(x) \text{ mit } y = g(x).$$

Beweis: Sei $\Phi = f \circ g$, $B = g'(x)$ und $A = f'(y)$. Dann ist

$$g(x+h) = g(x) + B \cdot h + o(\|h\|) \text{ für } h \rightarrow 0$$

und

$$f(y+k) = f(y) + A \cdot k + o(\|k\|) \text{ für } k \rightarrow 0.$$

Es ist $y = g(x)$, $y+k = g(x+h)$, also $k = g(x+h) - g(x)$. Damit ist

$$\begin{aligned}
\Phi(x+h) &= f(g(x+h)) = f(y+k) = f(y) + A \cdot k + o(\|k\|) \text{ für } k \rightarrow 0 \\
&= f(g(x)) + A \cdot (g(x+h) - g(x)) + o(\|g(x+h) - g(x)\|) \text{ für } h \rightarrow 0 \\
&= \Phi(x) + A \cdot (g(x+h) - g(x)) + o(\|g(x+h) - g(x)\|) \text{ für } h \rightarrow 0
\end{aligned}$$

Außerdem ist $\|k\| = O(\|h\|)$ für $h \rightarrow 0$. Dann gilt:

$$\begin{aligned}
\Phi(x+h) &= \Phi(x) + A \cdot (B \cdot h + o(\|h\|)) + \underbrace{o(O(\|h\|))}_{=o(\|h\|)} = \Phi(x) + A \cdot B \cdot h + o(\|h\|) + o(\|h\|) \\
&= \Phi(x) + A \cdot B \cdot h + o(\|h\|) = \Phi(x) + f'(y)g'(x)h + o(\|h\|)
\end{aligned}$$

Also ist $\Phi'(x) = f'(y)g'(x)$ mit $y = g(x)$.

8.2.8 Mittelwertsatz I

Sei $u: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ und S sei eine Strecke in D mit Anfangspunkt a und Endpunkt b . Ist u differenzierbar in D , so gibt es ein $\xi \in S \setminus \{a, b\}$ mit

$$u(b) - u(a) = u'(\xi)(b - a).$$

Beweis: Sei $\varphi(t) := u(a + t \cdot (b - a))$ für $0 \leq t \leq 1$. Nach dem 1-dimensionalen Mittelwertsatz (4.2.4, Seite 74) gibt es ein $\tau \in (0, 1)$ mit

$$\varphi(1) - \varphi(0) = \varphi'(\tau)(1 - 0).$$

Mit der Kettenregel gilt dann:

$$u(b) - u(a) = u'(\underbrace{a + \tau(b - a)}_{=: \xi}) \cdot (b - a).$$

8.2.9 Mittelwertsatz II

Ist $f: D \rightarrow \mathbb{R}^m$ differenzierbar und sind a, b und S wie im MWS I, so gilt:

$$f(b) - f(a) = J(\xi^1, \dots, \xi^m) \cdot (b - a).$$

Dabei ist

$$J(\xi^1, \dots, \xi^m) = \begin{pmatrix} f'_1(\xi^1) \\ f'_2(\xi^2) \\ \vdots \\ f'_m(\xi^m) \end{pmatrix} \text{ mit } \xi^1, \dots, \xi^m \in S \setminus \{a, b\}.$$

Beweis: Mittelwertsatz I anwenden auf f_μ :

$$f_\mu(b) - f_\mu(a) = f'_\mu(\xi^\mu)(b - a).$$

Dabei ist $\xi^\mu \in S \setminus \{a, b\}$ für $\mu = 1, \dots, m$.

Daraus folgt dann die Behauptung.

8.2.10 Definition: Integrierbarkeit von Matrizen

Seien $a_{\mu,\nu}: [0, 1] \rightarrow \mathbb{R}$ für $\mu = 1, \dots, m$ und $\nu = 1, \dots, n$ integrierbar und $A(t) = (a_{\mu,\nu}(t))_{\substack{\mu=1,\dots,m \\ \nu=1,\dots,n}}$.

Man setzt

$$\int_0^1 A(t) dt := \left(\int_0^1 a_{\mu,\nu}(t) dt \right)$$

und nennt A integrierbar. Das Ergebnis ist eine $m \times n$ -Matrix.

8.2.11 Mittelwertsatz III

Ist $f: D \rightarrow \mathbb{R}^m$ differenzierbar, sind a, b und S wie im MWS I und ist $f'(a + t \cdot (b - a))$ integrierbar über $[0, 1]$, dann gilt:

$$f(b) - f(a) = \underbrace{\left(\int_0^1 f'(a + t(b - a)) dt \right)}_{m \times n\text{-Matrix}} \cdot (b - a).$$

Beweis: Hier für $m = 1$. Sei $u: D \rightarrow \mathbb{R}$.

Setze $\varphi(t) = u(a + (b - a)t)$. Es ist $\varphi'(t) = u'(a + t(b - a)) \cdot (b - a)$.

Mit dem Hauptsatz gilt dann

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t) dt,$$

also

$$u(b) - u(a) = \int_0^1 u'(a + t(b - a)) \cdot (b - a) dt.$$

8.2.12 Satz

Ist $D \subseteq \mathbb{R}^n$ ein Gebiet und $f: D \rightarrow \mathbb{R}^m$ differenzierbar mit $f'(x) = 0$ in D , so ist f konstant.

Beweis: Zuerst als Erinnerung für $m = n = 1$:

$$f(y) - f(x) = \underbrace{f'(\xi)}_{=0} (y - x) = 0.$$

Seien nun $x, y \in D$ und sei s ein achsenparalleler Streckenzug in D von x nach y .

Die Ecken von s seien $x = a^1, a^2, \dots, a^k = y$. Es gilt:

$$f(a^j) - f(a^{j-1}) \stackrel{\text{MWS II}}{=} J(\xi^1, \dots, \xi^m) \cdot (a^j - a^{j-1}) = 0.$$

Daraus folgt:

$$f(y) - f(x) = f(a^k) - f(a^1) = \sum_{j=2}^k f(a^j) - f(a^{j-1}) = 0.$$

Also ist f konstant.

8.3 Partielle Ableitungen

8.3.1 Vorüberlegungen

Sei $u: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar in x .

Dann ist

$$u'(x) = (a_1, a_2, \dots, a_n) = a$$

und

$$u(x + h) = u(x) + a \cdot h + o(\|h\|) = u(x) + a_1 h_1 + \dots + a_n h_n + o(\|h\|)$$

mit $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$.

Was ist dann a_1 ?

Setze $h = (t, 0, \dots, 0)$. Dann gilt $h \rightarrow 0$ genau für $t \rightarrow 0$. Es ist

$$\begin{aligned} u(x_1 + t, x_2, \dots, x_n) - u(x_1, \dots, x_n) &= a_1 \cdot t + o(|t|) \\ \Rightarrow \frac{u(x_1 + t, x_2, \dots, x_n) - u(x_1, \dots, x_n)}{t} &= a_1 + \frac{o(|t|)}{t} \\ \Rightarrow a_1 &= \frac{u(x_1 + t, x_2, \dots, x_n) - u(x_1, \dots, x_n)}{t} + o(1) \text{ für } t \rightarrow 0 \\ \Rightarrow a_1 &= \lim_{t \rightarrow 0} \left[\frac{u(x_1 + t, x_2, \dots, x_n) - u(x_1, \dots, x_n)}{t} + o(1) \right] \end{aligned}$$

8.3.2 Definition: Partielle Ableitung, Gradient

Sei $u: D \rightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^n$ offen und $x \in D$. Dann heißt

$$\lim_{t \rightarrow 0} \frac{u(x_1, \dots, x_{\nu-1}, x_{\nu} + t, x_{\nu+1}, \dots, x_n) - u(x_1, \dots, x_n)}{t}$$

partielle Ableitung von u nach x_{ν} . Sie wird bezeichnet mit

$$D_{\nu}u(x) \text{ oder } \frac{\partial u}{\partial x_{\nu}}(x) \text{ oder } u_{x_{\nu}}(x).$$

Existieren alle partiellen Ableitungen, so heißt

$$\text{grad } u(x) := (u_{(x_1)}(x), \dots, u_{(x_n)}(x))$$

der *Gradient* von u .

Bemerkung: Der Gradientenvektor ist ein Zeilenvektor.

8.3.3 Satz

Ist u differenzierbar in x , so ist

$$u'(x) = \text{grad } u(x).$$

Beweis: Oben in 8.3.1.

8.3.4 Definition: Richtungsableitung

Sei $u: D \rightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^n$ offen, $x \in D$ und $e \in \mathbb{R}^n$ mit $\|e\| = 1$. Dann heißt

$$\frac{\partial u}{\partial e}(x) := \lim_{t \rightarrow 0} \frac{u(x + te) - u(x)}{t}$$

die *Richtungsableitung* von u in Richtung e .

Spezialfall: Für $e = e^{\nu}$ ist $\frac{\partial u}{\partial e^{\nu}} = \frac{\partial u}{\partial x_{\nu}}$.

8.3.5 Satz

Ist $u: D \rightarrow \mathbb{R}$ in x differenzierbar, so existieren alle Richtungsableitungen $\frac{\partial u}{\partial e}(x)$ und es ist

$$\frac{\partial u}{\partial e}(x) = e \cdot \text{grad } u(x).$$

Es gilt dann:

$$-\|\operatorname{grad} u(x)\| \leq \frac{\partial u}{\partial e}(x) \leq \|\operatorname{grad} u(x)\|.$$

Falls $\operatorname{grad} u(x) \neq 0$ ist, so ist Gleichheit nur für

$$e = \pm \frac{\operatorname{grad} u(x)}{\|\operatorname{grad} u(x)\|}.$$

Beweis: u ist differenzierbar. Damit ist

$$u(x+h) = u(x) + u'(x) \cdot h + o(\|h\|) \text{ für } h \rightarrow 0.$$

Setze $h = t \cdot e$ mit $e \in \mathbb{R}^n$, $\|e\| = 1$ und $t \in \mathbb{R}$. Damit ist

$$\begin{aligned} u(x+te) - u(x) &= tu'(x)e + o(|t|) \text{ für } t \rightarrow 0 \\ \Rightarrow \frac{u(x+te) - u(x)}{t} &= u'(x) \cdot e + o(1) \text{ für } t \rightarrow 0 \\ \Rightarrow \frac{\partial u}{\partial e}(x) &= u'(x) \cdot e = \operatorname{grad} u(x) \cdot e \end{aligned}$$

Beweis der Ungleichung:

$$\begin{aligned} \frac{\partial u}{\partial e} &= \operatorname{grad} u \cdot e \stackrel{\text{CSU}}{\leq} \|\operatorname{grad} u\| \cdot \underbrace{\|e\|}_{=1} \\ \frac{\partial u}{\partial e} &= \operatorname{grad} u \cdot e \stackrel{\text{CSU}}{\geq} -\|\operatorname{grad} u\|. \end{aligned}$$

Falls $\operatorname{grad} u \neq 0$ ist, gilt:

$$\begin{aligned} „ = “ &\iff \operatorname{grad} u \text{ und } e \text{ sind linear abhängig} \\ &\iff e = \lambda \operatorname{grad} u \text{ mit } \lambda \in \mathbb{R}. \end{aligned}$$

Für positives λ liegt rechts, für negatives λ liegt links Gleichheit vor. Es ist also $\lambda = \pm \frac{1}{\|\operatorname{grad} u\|}$.

8.3.6 Satz

Ist $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ differenzierbar, so ist

$$f'(x) = \begin{pmatrix} \operatorname{grad} f_1(x) \\ \operatorname{grad} f_2(x) \\ \vdots \\ \operatorname{grad} f_m(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Beweis: Ohne.

8.3.7 Beispiele

1. Sei $u(x, y) = e^x \cos(y - x)$. Dann ist

$$u_x = e^x \cos(y - x) + e^x (-\sin(y - x)) \cdot (-1) = e^x (\cos(y - x) + \sin(y - x))$$

und

$$u_y = -e^x \sin(y - x).$$

Frage: Ist u nun differenzierbar?

2. Sei

$$u(x, y) = \begin{cases} \frac{x^2}{y} & \text{für } |y| > x^2 \\ 0 & \text{sonst} \end{cases}.$$

u' und u sind unstetig in $(0, 0)$.

Alle Richtungsableitungen $\frac{\partial u}{\partial e}$ existieren in $(0, 0)$. Es ist

$$\frac{\partial u}{\partial e}(0, 0) = \begin{cases} 0 & \text{für } e = (\pm 1, 0) \\ \frac{a^2}{b} & \text{für } e = (a, b), a^2 + b^2 = 1, b \neq 0 \end{cases},$$

denn für $e = (\pm 1, 0)$ ist

$$\frac{u(\pm t, 0) - u(0, 0)}{t} = \frac{0 - 0}{t} \rightarrow 0 \text{ für } t \rightarrow 0$$

und für $e = (a, b)$ mit $b \neq 0$ ist

$$\begin{aligned} \frac{u(at, bt) - u(0, 0)}{t} &= \frac{\frac{a^2 t^2}{bt} - 0}{t} \text{ für kleines } |t| \\ &= \frac{a^2}{b} \rightarrow \frac{a^2}{b} \text{ für } t \rightarrow 0. \end{aligned}$$

Zeige nun, daß u in $(0, 0)$ nicht stetig ist:

Es ist $u(x, 0) = 0$ und $u(x, 2x^2) = \frac{x^2}{2x^2} = \frac{1}{2}$.

Also ist u nicht stetig und damit nicht differenzierbar in $(0, 0)$.

8.3.8 Satz

Existiert $\text{grad } u$ von $u: D \rightarrow \mathbb{R}$ und sind $u_{x_1}, \dots, u_{x_n}$ im Punkt x stetig, so ist u in x differenzierbar.

Beweis: Sei $h = (h_1, \dots, h_n)$ und $h^\nu = (h_1, \dots, h_\nu, 0, \dots, 0)$.

Insbesondere sind dann:

$$h^0 = (0, \dots, 0)$$

$$h^1 = (h_1, 0, \dots, 0)$$

$$h^2 = (h_1, h_2, 0, \dots, 0).$$

Es gilt dann:

$$\begin{aligned} u(x+h) - u(x) - \text{grad } u(x) \cdot h &= \sum_{\nu=1}^n \underbrace{u(x+h^\nu) - u(x+h^{\nu-1})}_{h_\nu u_{x_\nu}(x+\xi_\nu)} - u_{x_\nu}(x) h_\nu \\ &= \sum_{\nu=1}^n h_\nu [u_{x_\nu}(x+\xi_\nu) - u_{x_\nu}(x)]. \end{aligned}$$

Sei $\varepsilon > 0$. Dazu existiert (Stetigkeit des Gradienten) ein $\delta > 0$ mit

$$\|\text{grad } u(x) - \text{grad } u(y)\| < \varepsilon \text{ für } \|x - y\| < \delta.$$

Setze nun $y = x + h$ mit $\|h\| < \delta$. Dann ist

$$\begin{aligned} |u(x+h) - u(x) - \text{grad } u(x) \cdot h| &\stackrel{\text{CSU}}{\leq} \|h\| \cdot \sqrt{\sum_{\nu=1}^n (u_{x_\nu}(x+\xi_\nu) - u_{x_\nu}(x))^2} \\ &< \varepsilon \cdot \|h\| \text{ für } \|h\| < \delta. \end{aligned}$$

Also ist $u'(x) = \text{grad } u(x)$.

8.3.9 Bemerkungen

Sei $u: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar im Punkt x . Dann gilt

$$u(x+h) = u(x) + u'(x) \cdot h + o(\|h\|).$$

Für $n = 1$ ist $z = u(x) + u'(x) \cdot t$ die Gleichung der Tangenten an den Graphen an der Stelle x .

Für $n \geq 2$ ist

$$z = u(a) + u'(a) \cdot (x - a)$$

die Gleichung einer Hyperebene im $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$.

$$T_a := \{(x, z) \in \mathbb{R}^n \times \mathbb{R} : z = u(a) + u'(a) \cdot (x - a)\}$$

ist eine Tangentialhyperebene. Für diese Ebene gilt

$$(z - u(a)) - u'(a) \cdot (x - a) = 0.$$

Setze nun

$$\nu = \begin{pmatrix} -u_{x_1}(a) \\ \vdots \\ -u_{x_n}(a) \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1}.$$

Für $(x, z) \in T_a$ gilt nun

$$\nu \cdot \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \\ z - u(a) \end{pmatrix} = 0, \quad \nu \text{ ist orthogonal zu } \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \\ z - u(a) \end{pmatrix}.$$

Also ist ν Normalenvektor zur Tangentialhyperebene.

8.4 Höhere Ableitungen

8.4.1 Definition: Stetig differenzierbar

Sei $u: D \rightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}^n$ offen, Gebiet. u heißt *stetig differenzierbar*, wenn $u_{x_1}, \dots, u_{x_n}$ in D stetig sind: $u \in C^1(D)$. $f: D \rightarrow \mathbb{R}^m$ heißt stetig differenzierbar, wenn $f_1, \dots, f_m \in C^1(D)$ sind: $f \in C^1(D, \mathbb{R}^m)$.

8.4.2 Definition: Mehrfach stetig differenzierbar

$u: D \rightarrow \mathbb{R}$ heißt k -mal stetig differenzierbar, geschrieben: $u \in C^k(D)$, wenn alle Ableitungen $D_{j_1} \cdot D_{j_2} \dots D_{j_k} u$ für $j_1, \dots, j_k \in \{1, \dots, n\}$ in D stetige Funktionen sind.

8.4.3 Beispiele

- Es ist z. B. $D_1 D_2 u = (u_{x_2})_{x_1} =: u_{x_2 x_1}$
und $D_3 D_7 D_4 u = ((u_{x_4})_{x_7})_{x_3} =: u_{x_4 x_7 x_3}$.

- Sei $u(x, y) = x^2 - e^{xy}$. Dann ist

$$\begin{aligned} u_x &= 2x - ye^{xy} & u_y &= -xe^{xy} \\ u_{xx} &= 2 - y^2 e^{xy} & u_{xy} &= -e^{xy} - xye^{xy} \\ u_{yy} &= -x^2 e^{xy} & u_{yx} &= -e^{xy} - xye^{xy}. \end{aligned}$$

8.4.4 Satz von H. A. Schwarz (1. Version)

Sei $D \subseteq \mathbb{R}^2$ offen, $u: D \rightarrow \mathbb{R}$, $u \in C^1(D)$ und u_{xy} existiere in D und ist stetig in $(a, b) \in D$. Dann existiert auch $u_{yx}(a, b) = u_{xy}(a, b)$.

Beweis: Für $(a, b) = (0, 0)$:

Es gilt

$$\begin{aligned} R(x, y) &:= (u(x, y) - u(x, 0)) - (u(0, y) - u(0, 0)) \\ &= x \cdot [u_x(\xi, y) - u_x(\xi, 0)] \text{ mit } \xi \text{ aus dem MWS zwischen } 0 \text{ und } x \\ &= x \cdot y \cdot u_{xy}(\xi, \eta) \text{ mit } \eta \text{ zwischen } 0 \text{ und } y \text{ (MWS)} \end{aligned}$$

Daraus folgt:

$$\begin{aligned} u_{xy}(\xi, \eta) &= \frac{R(x, y)}{xy} = \frac{u(x, y) - u(x, 0) - u(0, y) + u(0, 0)}{xy} \\ &= \frac{1}{x} \cdot \frac{(u(x, y) - u(x, 0)) - (u(0, y) - u(0, 0))}{y} \end{aligned}$$

Für $y \rightarrow 0$ ist damit

$$\rightarrow \frac{1}{x} (u_y(x, 0) - u_y(0, 0))$$

und danach für $x \rightarrow 0$:

$$\rightarrow u_{yx}(0, 0)$$

Dabei geht die linke Seite wegen der Stetigkeit von $u_{x,y}$ gegen $u_{xy}(0, 0)$.

8.4.5 Satz von H. A. Schwarz (2. Version)

Ist $u \in C^k(D)$, so ist

$$D_{j_1} D_{j_2} \dots D_{j_k} u \text{ für } j_1, \dots, j_k \in \{1, \dots, n\}$$

unabhängig von der Reihenfolge der Differentiation.

Beweis: Als Aufgabe.

8.4.6 Definition und Regeln: Multiindex

$p \in \mathbb{N}_0^n$ heißt Multiindex. Es gilt:

$$\begin{aligned} p &= (p_1, p_2, \dots, p_n) \text{ mit } p_\nu \in \mathbb{N}_0 \\ |p| &:= p_1 + p_2 + \dots + p_n \text{ ist die Länge von } p \\ p! &:= p_1! \cdot p_2! \cdot \dots \cdot p_n! \end{aligned}$$

Für $x \in \mathbb{R}^n$ gilt:

$$x^p := x_1^{p_1} x_2^{p_2} \dots x_n^{p_n}.$$

8.4.7 Definition: Polynom, Taylorpolynom

$$\sum_{|p| \leq k} a_p x^p \text{ mit festem } k \text{ und } a_p = a_{p_1, \dots, p_n}$$

ist ein Polynom in $x \in \mathbb{R}^n$.

Z. B. ist für $n = 2$

$$a_{00} + a_{10}x_1 + a_{01}x_2 + a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2.$$

Sei nun $u \in C^k(D)$ und $p \in \mathbb{N}_0^n$ mit der Länge $|p| = k$.

Dann bedeutet

$$\frac{\partial^k u}{\partial x^p} = \frac{\partial^k u}{\partial x_1^{p_1} \partial x_2^{p_2} \dots \partial x_n^{p_n}}$$

differenziere p_1 -mal nach x_1 ,

differenziere p_2 -mal nach x_2 ,

$\vdots$

differenziere p_n -mal nach x_n .

Sei nun $u \in C^k(D)$ und $a \in D$. Dann ist

$$T_k(x; a) := \sum_{|p| \leq k} c_p (x - a)^p$$

mit

$$c_p := \frac{1}{p!} \cdot \frac{\partial^{|p|} u}{\partial x^p}(a)$$

das k -te Taylorpolynom von u an der Stelle a .

8.4.8 Satz von Taylor

Ist $u \in C^k(D)$ und ist $a \in D$, so gilt:

$$|u(x) - T_k(x; a)| = o(\|x - a\|^k)$$

für $x \rightarrow a$.

Beweis: Wähle $r > 0$ so, daß $K(a, r) \subseteq D$ ist. Sei dann $x \in K(a, r)$.

Setze

$$\varphi(t) := u(a + t(x - a)) \text{ für } 0 \leq t \leq 1.$$

Es ist $\varphi \in C^k([0, 1])$. Damit ist nach dem eindimensionalen Satz von Taylor (4.4.5, Seite 87):

$$\begin{aligned} \varphi(1) &= \sum_{\ell=0}^{k-1} \frac{\varphi^{(\ell)}(0)}{\ell!} 1^\ell + \frac{\varphi^{(k)}(\tau)}{k!} 1^k \text{ mit } 0 < \tau < 1 \\ &= \sum_{\ell=0}^k \frac{\varphi^{(\ell)}(0)}{\ell!} + \frac{\varphi^{(k)}(\tau) - \varphi^{(k)}(0)}{k!}. \end{aligned}$$

Mit Induktion (wird weiter unten gezeigt) gilt dann:

$$\frac{\varphi^{(\ell)}}{\ell!}(t) = \sum_{|p|=\ell} \frac{1}{p!} \cdot \frac{\partial^\ell u}{\partial x^p}(a + t(x - a))(x - a)^p. \quad (8.1)$$

Wenn (8.1) gilt, dann folgt:

$$\begin{aligned} u(x) &= \sum_{\ell=0}^k \sum_{|p|=\ell} \frac{1}{p!} \cdot \frac{\partial^\ell u}{\partial x^p}(a)(x-a)^p \\ &\quad + \left[\sum_{|p|=k} \frac{1}{p!} \left(\frac{\partial^k u}{\partial x^p}(a + \tau(x-a)) - \frac{\partial^k u}{\partial x^p}(a) \right) (x-a)^p \right] \\ &= T_k(x; a) + R_k(x; a). \end{aligned}$$

Es gilt:

$$|R_k(x; a)| \leq \sum_{|p|=k} \frac{1}{p!} \cdot \left| \frac{\partial^k u}{\partial x^p}(a + \tau(x-a)) - \frac{\partial^k u}{\partial x^p}(a) \right| \cdot \|x-a\|^k,$$

weil

$$|h^p| = |h_1^{p_1} \dots h_n^{p_n}| \leq \|h\|^{p_1+p_2+\dots+p_n} = \|h\|^{|p|}.$$

Sei nun $\varepsilon > 0$. Dazu existiert (wegen der Stetigkeit) ein $\delta > 0$ mit

$$\left| \frac{\partial^k u}{\partial x^p}(a+h) - \frac{\partial^k u}{\partial x^p}(a) \right| < \varepsilon \text{ für } \|h\| < \delta.$$

Für $\|x-a\| < \delta$ gilt dann:

$$|R_k(x; a)| \leq \varepsilon \|x-a\|^l \cdot (\# \text{Anzahl der Summanden}).$$

Beweise nun die Gleichung (8.1): Hier für $a = 0$.

$\ell = 1$: Es ist

$$\varphi(t) \stackrel{!}{=} u(tx) \checkmark$$

$\ell \mapsto \ell + 1 \leq k$:

$$\begin{aligned} \frac{\varphi^{(\ell+1)}}{(\ell+1)!}(t) &= \frac{1}{\ell+1} \left(\frac{\varphi^{(\ell)}(t)}{\ell!} \right)' \\ &= \frac{1}{\ell+1} \sum_{|p|=\ell} \frac{1}{p!} \sum_{\nu=1}^n \frac{\partial}{\partial x_\nu} \cdot \frac{\partial^\ell u}{\partial x^p}(tx) \cdot x_\nu x^p \\ &= \frac{1}{\ell+1} \sum_{\nu=1}^n \sum_{|p|=\ell} \frac{1}{p!} \cdot \frac{\partial^{\ell+1} u}{\partial x^{p+e^\nu}}(tx) x^{p+e^\nu} \end{aligned}$$

mit e^ν , dem ν -ten Einheitsvektor. Es gilt:

$$|p+e^\nu| = |p| + 1 = \ell + 1.$$

Sei nun q ein Multiindex der Länge $\ell + 1$. Dann ist

$$M_q := \{p: |p| = \ell \text{ und } q = p + e^\nu \text{ für ein } \nu\}.$$

Mit

$$\sum_{p \in M_q} \frac{1}{p!} = \sum_{q_\nu > 0} \frac{p_\nu + 1}{q!} = \sum_{q_\nu > 0} \frac{q_\nu}{q!} = \frac{1}{q!}(\ell + 1)$$

gilt dann:

$$\begin{aligned}
 \frac{\varphi^{(\ell+1)}}{(\ell+1)!}(t) &= \frac{1}{\ell+1} \sum_{|q|=\ell+1} \sum_{p \in M_q} \frac{1}{p!} \cdot \frac{\partial^{\ell+1} u}{\partial x^q}(tx) x^q \\
 &= \frac{1}{\ell+1} \sum_{|q|=\ell+1} \frac{1}{q!} \cdot \frac{\partial^{\ell+1} u}{\partial x^q}(tx) x^q (\ell+1) \\
 &= \sum_{|q|=\ell+1} \frac{1}{q!} \cdot \frac{\partial^{\ell+1} u}{\partial x^q}(tx) x^q.
 \end{aligned}$$

8.4.9 Spezialfall für den Satz von Taylor

Sei $u \in C^2(D)$ und $a \in D$. Dann ist

$$\begin{aligned}
 u(x) &= u(a) + \sum_{\nu=1}^n u_{x_\nu}(a)(x_\nu - a_\nu) \\
 &\quad + \frac{1}{2} \sum_{\mu, \nu=1}^n u_{x_\nu x_\mu}(a)(x_\nu - a_\nu)(x_\mu - a_\mu) \\
 &\quad + o(\|x - a\|^2).
 \end{aligned}$$

8.4.10 Definition: Hesse-Matrix

Für $u \in C^2(D)$ heißt

$$H_u(a) := \begin{pmatrix} u_{x_1 x_1} & u_{x_1 x_2} & \cdots & u_{x_1 x_n} \\ \vdots & & & \vdots \\ u_{x_n x_1} & u_{x_n x_2} & \cdots & u_{x_n x_n} \end{pmatrix} (a)$$

Hesse-Matrix. H_u ist symmetrisch, da $u_{x_\mu x_\nu} = u_{x_\nu x_\mu}$.

Setze man $Q_u(h) := h^T H_u(a) h$, dann gilt mit Taylor:

$$u(a+h) = u(a) + u'(a) \cdot h + \frac{1}{2} Q_u(h) + o(\|h\|^2).$$

8.4.11 Definition: Extremum

Sei $u: D \rightarrow \mathbb{R}$. $a \in D \subseteq \mathbb{R}^n$ heißt Stelle eines Maximums bzw. Minimums (zusammen: Extremums) der Funktion u , wenn

$$u(x) \leq u(a) \text{ bzw. } u(x) \geq u(a) \text{ für alle } \|x - a\| < \delta.$$

8.4.12 Satz

Sei $u: D \rightarrow \mathbb{R}$ und $a \in D$.

- (1) Ist a Stelle eines Extremums und existiert $u'(a)$, so ist $u'(a) = 0$.
- (2) Ist sogar $u \in C^2(D)$, so gilt im Falle eines Maximums bzw. Minimums, daß $H_u(a)$ negativ bzw. positiv semidefinit ist.

Beweis:

(1) $\varphi(t) = u(a_1 + t, a_2, \dots, a_n)$ hat Extremum für $t = 0$, d. h. $\varphi'(0) = 0 = u_{x_1}(a)$, also ist $u'(a) = 0$.

(2) Gilt:

$$\begin{array}{ll} \text{Minimum:} & \varphi''(0) \leq 0 \quad u_{x_1 x_1}(a) \leq 0 \\ \text{Maximum:} & \varphi''(0) \geq 0 \quad u_{x_1 x_1}(a) \geq 0 \end{array} ?$$

Z. B. für das Minimum: Sei $\|h\| \leq \delta$:

$$u(a) \leq u(a+h) = u(a) + \underbrace{u'(a) \cdot h}_{=0} + \frac{1}{2} Q_u(h) + o(\|h\|^2)$$

$$\Rightarrow \frac{1}{2} Q_u(h) + o(\|h\|^2) \geq 0 \text{ für } \|h\| < \delta.$$

Wähle nun $e \in \mathbb{R}^n$ und setze $h = te$ mit kleinem $t \in \mathbb{R}$. Dann ist

$$0 \leq \frac{1}{2} Q_u(te) + o(\|te\|^2) = \underbrace{t^2}_{>0} \underbrace{\left(\frac{1}{2} Q_u(e) + o(1) \right)}_{\geq 0} \text{ für } t \rightarrow 0$$

Für $t \rightarrow 0$ ist damit dann $Q_u(e) \geq 0$ für $e \in \mathbb{R}^n$, d. h. Q_u ist positiv semidefinit.

8.4.13 Satz

Sei $u \in C^2(D)$, $a \in D$, $u'(a) = 0$ und $H_u(a)$ positiv bzw. negativ definit. Dann ist a Stelle eines Minimums bzw. Maximums.

Ist H_u indefinit, so liegt in a weder ein Maximum noch ein Minimum.

Beweis: Sei Q_u positiv definit. Dann ist

$$\begin{aligned} u(a+h) - u(a) &= \underbrace{\frac{1}{2} Q_u(h)}_{\substack{\geq \alpha \|h\|^2 \\ \text{mit } \alpha > 0}} + \underbrace{o(\|h\|^2)}_{\substack{|\cdot| < \frac{\alpha}{2} \|h\|^2 \\ \text{für } \|h\| < \delta}} \\ &\geq \left(\alpha - \frac{\alpha}{2} \right) \|h\|^2 > 0 \text{ für } 0 < \|h\| < \delta \end{aligned}$$

Also ist in a ein Minimum.

8.4.14 Beispiele

(1) Sei $u(x, y) = x^2 - y^2$. Dann ist $u_x = 2x$ und $u_y = -2y$. $u' = 0$ gilt genau für $(x, y) = (0, 0)$.

Es ist $u_{xx} = 2$, $u_{xy} = 0 = u_{yx}$ und $u_{yy} = -2$.

Also ist $H_u = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$. H_u ist indefinit.

Also gibt es keine Extrema.

(2) Sei $u(x, y) = xy(1 - x - y)$.

Dann ist $u_x = y(1 - x - y) + xy(-1) = y(1 - 2x - y)$ und $u_y = x(1 - 2y - x)$.

$u_x = 0, u_y = 0 \iff (x, y) \in \{(0, 0), (1, 0), (0, 1), (\frac{1}{3}, \frac{1}{3})\}$.

Es ist $u_{xx} = -2y$, $u_{yy} = -2x$

und $u_{xy} = u_{yx} = -2x - 2y$.

$$H_u(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ ist indefinit.}$$

$$H_u(1,0) = \begin{pmatrix} 0 & -1 \\ -1 & -2 \end{pmatrix} \text{ ist indefinit.}$$

$$H_u(0,1) = \begin{pmatrix} -2 & -1 \\ -1 & 0 \end{pmatrix} \text{ ist indefinit.}$$

$$H_u\left(\frac{1}{3}, \frac{1}{3}\right) = \begin{pmatrix} -\frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{2}{3} \end{pmatrix} \text{ ist negativ definit, also liegt ein Maximum vor.}$$

- (3) Sei $u(x, y, z) = x^2 + y^2 - xy - x + z^2 - z$.
Es sind dann

$$u_x = 2x - y - 1$$

$$u_y = 2y - x$$

$$u_z = 2z - 1 = 0$$

Es ist $(u_x, u_y, u_z) = (0, 0, 0)$ für $(x, y, z) = \left(\frac{2}{3}, \frac{1}{3}, \frac{1}{2}\right)$

und $H_u = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ ist positiv definit,

also liegt in $\left(\frac{2}{3}, \frac{1}{3}, \frac{1}{2}\right)$ ein Minimum vor.

8.5 Implizite Funktionen und Umkehrfunktionen

8.5.1 Beispiel

Die Gleichung $x^y = y^x$ hat für $x, y > 0$ z. B. folgende Lösungen:

$$y = x$$

$$x = 4, y = 2$$

$$y = 4, x = 2$$

Gibt es weitere Lösungen?

8.5.2 Schreibweisen

Im Folgenden ist $\mathbb{R}^{n+m} := \mathbb{R}^n \times \mathbb{R}^m$. Für $\begin{pmatrix} x \\ y \end{pmatrix}$ oder (x, y) ist $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^m$.

8.5.3 Satz über implizit definierte Funktionen

Es sei $U = K(x^0, r) \subseteq \mathbb{R}^n$ und $V = K(y^0, \varrho) \subseteq \mathbb{R}^m$.

Außerdem sei $f: U \times V \rightarrow \mathbb{R}^m$ mit $f(x^0, y^0) = 0$ gegeben.

f sei stetig,

$$f_y := \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{pmatrix}$$

sei stetig und $\det f_y(x^0, y^0) \neq 0$.

Dann gibt es Kugeln $U_0 = K(x^0, r_0)$ und $V_0 = K(y^0, \varrho_0)$, sowie eine stetige Funktion $g: U_0 \rightarrow V_0$ mit

$$\{(x, y): x \in U_0, y \in V_0, f(x, y) = 0\} = \{(x, g(x)): x \in U_0\}.$$

Ist f_x stetig, so ist $g \in C^1(U_0, \mathbb{R}^m)$ und in U_0 ist

$$g'(x) = -(f_y(x, g(x)))^{-1} \cdot f_x(x, g(x)).$$

Beweis: Sei OBdA $x_0 = 0 \in \mathbb{R}^n$ und $y^0 = 0 \in \mathbb{R}^m$ (sonst betrachte $f(x^0+x, y^0+y) = \tilde{f}(x, y) = 0$). Setze

$$A := (f_y(0, 0))^{-1}.$$

Dann ist

$$f(x, y) = 0 \iff y = y - Af(x, y) =: F(x, y).$$

Beachte: Es ist $F(0, 0) = 0 \in \mathbb{R}^m$ und $F_y(0, 0) = 0$ ist $(m \times n)$ -Nullmatrix.

Wähle nun U_0 und V_0 so, daß $\|F(x, 0)\| < \varrho_0/2$ für $x \in U_0$ ist (dies ist wegen der Stetigkeit von F möglich), und daß $\|F_y(x, y)\| < \frac{1}{2}$ ist für $(x, y) \in U_0, V_0$.

Eindeutigkeit von g :

Zeige: Zu $x \in U_0$ gibt es höchstens ein $y \in V_0$ mit $f(x, y) = 0$.

Annahme: Es gibt $y, z \in V_0$ mit $f(x, y) = f(x, z) = 0$. Es gilt dann:

$$y = F(x, y)$$

$$z = F(x, z)$$

Nach Subtraktion gilt dann:

$$\begin{aligned} y - z &= F(x, y) - F(x, z) \\ &\stackrel{\text{MWS III}}{=} \left[\int_0^1 F_y(x, ty + (1-t)z) dt \right] \cdot (y - z). \end{aligned}$$

Daraus folgt dann:

$$\begin{aligned} \|y - z\| &\leq \left\| \int_0^1 \underbrace{F_y(x, ty + (1-t)z)}_{\|\cdot\| < \frac{1}{2}} dt \right\| \cdot \|y - z\| \\ &\leq \frac{1}{2} \|y - z\|, \end{aligned}$$

also ist $y = z$.

Existenz von g :

Setze

$$M = \{u: u \text{ ist stetig in } U_0, u: U_0 \rightarrow V_0, u(0) = 0\}.$$

Definiere folgende Metrik auf M :

$$d(u, v) := \sup \{\|u(x) - v(x)\| : x \in U_0\}.$$

Anmerkung zur Definition dieser Norm:

Sei $B = \{u \in C(U_0, \mathbb{R}^m) : u(0) = 0\}$.

Auf B ist folgende Norm definiert: $\|u\|_\infty := \sup \{\|u(x)\| : x \in U_0\}$.

B ist ein Banachraum.

M ist abgeschlossene Teilmenge von B mit $d(u, v) = \|u - v\|_\infty$.

Also ist M vollständig.

Definiere nun eine Kontraktion T auf M :

$$T : \begin{cases} u \in M & \mapsto Tu \\ (Tu)(x) & := F(x, u(x)) \end{cases}$$

Zeige nun:

1. $T: M \rightarrow M$, T ist Selbstabbildung
2. $d(Tu, Tv) \leq \frac{1}{2}d(u, v)$, T ist eine Kontraktion.

Dann gibt es nach dem Banachschen Fixpunktsatz (7.5.5, Seite 176) (genau) ein $u \in M$, stetig mit $u = Tu$, d. h.

$$\begin{aligned} u(x) &= F(x, u(x)) = u(x) - Af(x, u(x)) \\ \iff f(x, u(x)) &= 0 \text{ in } U_0. \end{aligned}$$

Setze dann $g = u$.

Beweise nun, daß T eine Selbstabbildung (1) und Kontraktion(2) ist:

- (1) $\bar{u}(x) := (Tu)(x) = F(x, u(x))$ ist wohldefiniert, da $\|u(x)\| < \varrho_0$ ist, und da F in $U_0 \times V_0$ definiert ist.
 Es gilt: $\bar{u}(0) = F(0, 0) = 0$.
 $\bar{u}$ ist stetig.
 Ist $\|\bar{u}\| < \varrho_0$, d. h. gilt $\bar{u}: U_0 \rightarrow V_0$?

$$\begin{aligned} \|\bar{u}(x)\| &\leq \|F(x, 0)\| + \|F(x, u(x)) - F(x, 0)\| \\ &< \frac{\varrho_0}{2} + \frac{1}{2}\|u(x) - 0\| < \varrho_0. \end{aligned}$$

- (2) Beweis der Kontraktion:

Seien $u, v \in M$, $\bar{u} = Tu$ und $\bar{v} = Tv$. Dann ist

$$\begin{aligned} \|\bar{u}(x) - \bar{v}(x)\| &= \|F(x, u(x)) - F(x, v(x))\| \\ &\leq \frac{1}{2}\|u(x) - v(x)\| \leq \frac{1}{2}d(u, v) \end{aligned}$$

für alle $x \in U_0$.

Also ist $d(\bar{u}, \bar{v}) \leq \frac{1}{2}d(u, v)$.

Differenzierbarkeit von g (zunächst in $x = 0$). (Es ist $F \in C^1(U_0 \times V_0, \mathbb{R}^m)$)

Es gilt die Approximation

$$F(x, y) = F_x(0, 0)x + o(\|(x, y)\|) \text{ für } (x, y) \rightarrow (0, 0).$$

Beachte dabei, daß $F_y(0, 0) = 0$ ist.

Insbesondere gilt:

$$g(x) = F(x, g(x)) = F_x(0, 0)x + o(\|(x, g(x))\|),$$

also ist $g(x) = O(\|x\|)$ für $x \rightarrow 0$.

Es ist dann

$$\begin{aligned} g(x) - F_x(0, 0)x &= o(\|(x, g(x))\|) = o(\|x\|) \text{ für } x \rightarrow 0 \\ g(x) - g(0) - F_x(0, 0)x &= o(\|x\|), \end{aligned}$$

d. h.

$$g'(0) = +F_x(0, 0) = -Af_x(0, 0) = (f_y(0, g(0)))^{-1}f_x(0, g(0)).$$

Also gilt die Behauptung für $x = 0$.

Differenzierbarkeit von g für $\xi \in U_0$: Es ist zu zeigen:

$$g'(\xi) = -(f_y(\xi, g(\xi)))^{-1}f_x(\xi, g(\xi)) \text{ für } \xi \in U_0.$$

Die Rolle, die (x^0, y^0) gespielt hat, kann jeder Punkt $(\xi, g(\xi))$ übernehmen.

8.5.4 Beispiel

Für $n = 2$ und $m = 1$. Die Gleichung

$$\sin(x + z) + \sin(y + z) = z$$

ist in einer Umgebung von $(x, y) = (0, 0)$ in der Form $z = g(x, y)$ aufzulösen, wobei $g(0, 0) = 0$ ist.

Überprüfe die Voraussetzungen: Es ist $f(x, y, z) = \sin(x + z) + \sin(y + z) - z$.

$$f(0, 0, 0) = 0$$

$$\text{Außerdem ist } f_z(x, y, z) = \cos(x + z) + \cos(y + z) - 1$$

$$f_z(0, 0, 0) = 1 + 1 - 1 = 1 \neq 0.$$

Also existiert eine Funktion $g: U_0 \rightarrow V_0$ mit $g(0, 0) = 0$ und $f(x, y, g(x, y)) = 0$ für alle $(x, y) \in U_0$.

Dabei ist $U_0 = K((0, 0), r_0) \subseteq \mathbb{R}^2$ und $V_0 = (-\varrho, \varrho)$.

Berechne nun $g_x(0, 0)$ und $g_y(0, 0)$:

Es ist $f(x, y, g(x, y)) = 0$ in U_0 . Differenziere nach x und y .

Dann ist

$$f_x + f_z \cdot g_x = 0$$

und

$$f_y + f_z \cdot g_y = 0.$$

In $(x, y) = (0, 0)$, $z = 0$ gilt dann:

$$1 \cdot g_x(0, 0) = -f_x(0, 0, 0) = -\cos(0 + 0) = -1$$

$$1 \cdot g_y(0, 0) = -f_y(0, 0, 0) = -\cos(0 + 0) = -1.$$

8.5.5 Satz über die Umkehrfunktion

Sei $f: D \rightarrow \mathbb{R}^n$, $D \subseteq \mathbb{R}^n$ Gebiet, $f \in C^1(D, \mathbb{R}^n)$.

Außerdem sei $\xi \in D$ und $\det f'(\xi) \neq 0$. Dann gibt es in einer Kugel $V_0 = K(\eta, r)$ eine Umkehrfunktion g von f mit $g(\eta) = \xi$.

g ist dadurch eindeutig bestimmt und es gilt:

$$g'(y) = (f'(x))^{-1} \Big|_{x=g(y)}$$

Beweis: Setze $F: D \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $F(x, y) := f(x) - y$.

$F \in C^1(D \times \mathbb{R}^n, \mathbb{R}^n)$ und $F(\xi, \eta) = 0$.

$F_x(x, y) = f'(x)$, also ist $\det F_x(\xi, \eta) = \det f'(\xi) \neq 0$.

Nach dem Satz über implizite Funktionen existiert V_0 sowie $g \in C^1(V_0, \mathbb{R}^n)$ mit

$$g(\eta) = \xi$$

und

$$F(g(y), y) = 0 \text{ für alle } y \in V_0.$$

Es gilt:

$$\begin{aligned} f(g(y)) - y &= 0, \quad g = f^{-1} \text{ in } V_0 \\ f(g(y)) &= y \\ f'(g(y)) \cdot g'(y) &= I = \text{diag}(1, \dots, 1). \end{aligned}$$

Daraus folgt die Behauptung für $g'(y)$.

8.5.6 Definition: Diffeomorphismus

Eine bijektive C^1 -Funktion $f: D \rightarrow G$ mit $f^{-1} \in C^1(G, \mathbb{R}^n)$ heißt *Diffeomorphismus*.

8.5.7 Satz über die Gebietstreue

Ist $f \in C^1(D, \mathbb{R}^n)$ mit einem Gebiet $D \subseteq \mathbb{R}^n$ und regulärem $f'(x)$ in D , so ist $f(D)$ wieder ein Gebiet. ($f'(x)$ ist regulär, wenn die Determinante ungleich Null ist.)

Beweis: $G = f(D)$ ist zusammenhängend, weil D zusammenhängend und weil f stetig ist.

Sei nun $\eta \in G = f(D)$ und $\eta = f(\xi)$.

Da $\det f'(\xi) \neq 0$ ist, existiert eine Kugel $V_0 = K(\eta, r_0)$ sowie $f^{-1}: V_0 \rightarrow D$.

Es gilt: $V_0 \subseteq f(D)$, η ist innerer Punkt, beliebig.

Also ist $f(D) = G$ offen.

8.6 Extrema mit Nebenbedingungen

8.6.1 Allgemeines Problem

Sei $D \subseteq \mathbb{R}^n$ ein Gebiet, $f: D \rightarrow \mathbb{R}$ und $g: D \rightarrow \mathbb{R}^m$ mit $m < n$.

Gesucht sind Maxima bzw. Minima von f auf

$$M = \{x: g(x) = 0\}.$$

Extremum von f unter der Nebenbedingung $g = 0$.

8.6.2 Konkrete Probleme

(a) A sei eine symmetrische $n \times n$ -Matrix.

Maximiere bzw. Minimiere $x^T A x$ auf der Menge $\{x: \|x\| = 1\}$.

Hier: $f(x) = x^T A x$, $D = \mathbb{R}^n$, $g(x) = \|x\|^2 - 1$.

(b) Maximiere $x_1 x_2 \dots x_n$ unter der NB $x_1 + \dots + x_n = n$, alle $x_\nu > 0$.

Hier: $f(x) = x_1 \dots x_n$, $g(x) = x_1 + \dots + x_n - n$, $D = \{x: x_1 > 0, \dots, x_n > 0\}$.

8.6.3 Lagrangesche Multiplikationsregel

$f \in C^1(D)$ habe in $\xi \in D$ ein Extremum unter der Nebenbedingung $g(x) = 0$, wobei $g \in C^1(D, \mathbb{R}^m)$ und $\text{rg } g'(\xi) = m$, $m < n$.

Dann gibt es ein $\lambda^0 \in \mathbb{R}^m$ mit: Die Ableitung von

$$H(x, \lambda) := f(x) - \lambda \cdot g(x)$$

$H: D \times \mathbb{R}^m \rightarrow \mathbb{R}$ verschwindet in (ξ, λ^0) .

λ^0 heißt Lagrangemultiplikator.

Beweis: Sei $\text{rg } g'(\xi) = m$. Dann enthält die $m \times n$ -Matrix eine reguläre $m \times m$ -Matrix. Nach Ummumerierung sei OBdA $\left(\frac{\partial g_\mu}{\partial x_\nu}\right)_{\mu, \nu=1, \dots, m}$ regulär.

Schreibe $x = (y, z)$ und $\xi = (\eta, \zeta)$. Dabei ist $y = (x_1, \dots, x_m) \in \mathbb{R}^m$, $z = (x_{m+1}, \dots, x_n) \in \mathbb{R}^{n-m}$, $\eta \in \mathbb{R}^m$ und $\zeta \in \mathbb{R}^{n-m}$.

Es ist $g(y, z) = 0$, $g(\eta, \zeta) = 0$ und $g_y(\eta, \zeta)$ regulär.

Nach dem Satz über implizite Funktionen gibt es eine Kugel $U = (\zeta, r) \subseteq \mathbb{R}^{n-m}$ und $h \in C^1(U, \mathbb{R}^m)$ mit

$$g(h(z), z) = 0 \text{ in } U$$

$$g(y, z) \neq 0 \text{ für } z \in U, y \neq h(z)$$

$\Phi(z) = f(h(z), z)$, definiert in U , hat Extremum in $z = \zeta \in U$, also

$$\Phi'(\zeta) = 0.$$

Es ist $\Phi'(z) = f_y(h(z), z) \cdot h'(z) + f_z(h(z), z)$ und $g(h(z), z) = 0$ in U .

$$\Rightarrow \underbrace{g_y(h(z), z)}_{m \times n\text{-Matrix}} \cdot \underbrace{h'(z)}_{m \times (n-m)} + g_z(h(z), z) = 0$$

$g_y(h(\zeta), \zeta) = g_y(\eta, \zeta)$ ist regulär.

Mit Einsetzen von (2) in (1) ergibt sich

$$\begin{aligned} h'(\zeta) &= (g_y(\eta, \zeta))^{-1} \cdot g_z(\eta, \zeta) \\ -f_y(\eta, \zeta) \cdot (g_y(\eta, \zeta))^{-1} g_z(\eta, \zeta) + f_z(\eta, \zeta) &= 0 \\ -f_y(\xi) \cdot (g_y(\xi))^{-1} g_z(\xi) + f_z(\xi) &= 0 \end{aligned}$$

Setze $(\lambda^0)^T := -f_y(\xi) \cdot (g_y(\xi))^{-1} \in \mathbb{R}^m$. Daraus folgt dann:

$$\begin{aligned} f_z + \lambda^0 g_z &= 0 \text{ in } \xi \\ f_y + \lambda^0 g_y &= 0 \text{ in } \xi \end{aligned}$$

8.6.4 Beispiel

Maximiere bzw. Minimiere $x^T A x$ unter der NB $\|x\|^2 = 1$.

(1) $S^{n-1} = \{\|x\|^2 = 1\}$ ist kompakt.

Also existieren Maximum und Minimum von $x^T A x$ auf S^{n-1} .

(2) Setze

$$g(x) = \|x\|^2 - 1 = \left(\sum_{\nu=1}^n x_\nu^2 \right) - 1$$

Es ist $m = 1 < n$,

$$g'(x) = (2x_1, 2x_2, \dots, 2x_n) \neq 0 \text{ auf } S^{n-1}$$

und $\text{rg } g'(x) = 1$.

(3) Es ist

$$H(x, \lambda) = \sum_{\nu, \mu=1}^n a_{\nu\mu} x_\mu x_\nu - \lambda \cdot \left(\left(\sum_{\nu=1}^n x_\nu^2 \right) - 1 \right)$$

Dann ist

$$H_{x_\varrho} = \sum_{\nu=1}^n a_{\varrho\nu} x_\nu + \sum_{\mu=1}^n a_{\mu\varrho} x_\mu - 2\lambda x_\varrho = 2 \left(\left(\sum_{\nu=1}^n a_{\varrho\nu} x_\nu \right) - \lambda x_\varrho \right) \stackrel{!}{=} 0 \text{ für } \varrho = 1, \dots, n$$

$$H_\lambda = \sum_{\nu=1}^n x_\nu^2 - 1 \stackrel{!}{=} 0$$

$$\iff Ax - \lambda x = 0 \text{ und } \|x\|^2 = 1$$

Minimum und Maximum von $x^T Ax$ werden *nur* in Eigenvektoren auf S^{n-1} angenommen. Dabei ist λ der zugehörige Eigenwert.

$$\min_{S^{n-1}}, \max_{S^{n-1}} x^T Ax \in \{x^T Ax : x \in S^{n-1}, x \text{ Eigenvektor}\}$$

$\max, \min x^T Ax$ = größter bzw. kleinster Eigenwert.

8.6.5 Praktisches Vorgehen

Gesucht: Extremum von $f(x)$ unter der NB $g(x) = 0$.

1. Ansatz: $H(x, \lambda) := f(x) + \lambda g(x)$.
2. Ableitung berechnen und gleich 0 setzen:

$$H_x = f' + \lambda g' \stackrel{!}{=} 0 \qquad H_\lambda = g \stackrel{!}{=} 0$$

Dies sind $n + m$ Gleichungen mit $n + m$ Unbekannten.

3. Auflösen nach (x, λ) . λ nicht unbedingt explizit erforderlich: Lösungen (ξ, λ) .
4. Ist $\text{rg } g'(\xi) = m$? Hat f unter NB überhaupt Extrema? Wenn ja, suche unter den Werten $f(\xi)$.

8.6.6 Beispiele

1. Maximiere $x_1 x_2 \dots x_n$ unter der NB $x_1 + x_2 + \dots + x_n = n$ ($x_1, \dots, x_n > 0$).
 Es ist $f(x) = x_1 \dots x_n$, $g(x) = x_1 + \dots + x_n - n$ und $D = \{x: x_\nu > 0 \text{ für } \nu = 1, \dots, n\}$.
 Ansatz:

$$\begin{aligned} H(x, \lambda) &= x_1 \dots x_n + \lambda(x_1 + \dots + x_n - n) \\ H_{x_j} &= \frac{x_1 \dots x_n}{x_j} + \lambda \stackrel{!}{=} 0 \text{ für alle } j \\ H_\lambda &= x_1 + \dots + x_n - n = 0 \end{aligned}$$

Also ist $x_1 = x_2 = \dots = x_n = 1$, da $x_1 + \dots + x_n = n$ ist und alle x_ν gleich sein müssen.
 Es ist $\xi = (1, \dots, 1)$. Ist

$$f(\xi) = 1 \dots 1 = 1 = \text{Maximum unter NB?}$$

Es ist $g'(x) = (1, \dots, 1) \neq 0$ und $\text{rg } g'(x) = 1 = m$.

Existenz des Maximums:

$$\{x: x_1 + \dots + x_n = n, x_1, \dots, x_n > 0\}$$

ist kompakt. Darauf hat f ein Maximum. Es wird *nicht* angenommen, wenn ein $x_\nu = 0$, also angenommen, wo alle $x_\nu > 0$ sind, also in D .

2. A sei eine $n \times n$ -Matrix mit

$$\sum_{\mu, \nu=1}^n a_{\mu, \nu}^2 = n.$$

Gesucht ist A , so daß $\det A$ maximal wird.

- (a) $\{(a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{n1}, \dots, a_{nn}): a_{\mu, \nu} \in \mathbb{R}\} = \mathbb{R}^{(n^2)}$
 (b) $\{a_{\mu, \nu}: \sum_{\mu, \nu=1}^n a_{\mu, \nu}^2 = n\}$ ist kompakt.
 (c) $A \mapsto \det A$ ist stetig. Somit existiert das Maximum.
 (d) Es gilt $\det A = \sum_{k=1}^n a_{jk} A_{jk}$ (LaPlace'scher Entwicklungssatz).
 (e) Setze $f(A) = \det A$ und $g(A) = \sum_{\mu, \nu=1}^n a_{\mu, \nu}^2 - n$.

Ansatz:

$$\begin{aligned} H(A, \lambda) &= \det A + \lambda \left(\sum_{\mu, \nu=1}^n a_{\mu, \nu}^2 - n \right) \\ &= \sum_{k=1}^n a_{jk} A_{jk} + \lambda \left(\sum_{\mu, \nu=1}^n a_{\mu, \nu}^2 - n \right) \end{aligned}$$

$$\frac{\partial H}{\partial a_{j\ell}} = \sum_{k=1}^n \frac{\partial a_{jk}}{\partial a_{j\ell}} A_{jk} + 2\lambda a_{j\ell} = A_{j\ell} + 2\lambda a_{j\ell} = 0$$

$$A_{j\ell} = -(2\lambda) a_{j\ell}$$

$$\det A = \sum_{\ell=1}^n a_{j\ell} A_{j\ell} = -2\lambda \sum_{\ell=1}^n a_{j\ell}^2$$

für $j = 1, \dots, n$.
Dann ist

$$n \cdot \det A = -2\lambda \cdot \sum_{\ell=1}^n \sum_{j=1}^n a_{j\ell}^2 = n(-2\lambda)$$

$$\det A = -2\lambda.$$

Es gilt:

$$\frac{1}{\det A} \begin{pmatrix} A_{11} & \dots & A_{n1} \\ \vdots & & \vdots \\ A_{1n} & \dots & A_{nn} \end{pmatrix} = A^{-1}$$

$$\frac{-2\lambda}{\det A} A^T = A^{-1} \Rightarrow A^{-1} = A^T$$

A ist Orthogonalmatrix, also ist $\det A = 1$.

8.6.7 Satz

Unter allen $n \times n$ -Matrizen mit $\sum_{\mu,\nu=1}^n a_{\mu,\nu}^2 = n$ haben genau die Orthogonalen maximale Determinanten $\det A = 1$.

Hadamardsche Determinantengleichung: Sei A eine $n \times n$ -Matrix:

$$(\det A)^2 \leq \prod_{j=1}^n \sum_{k=1}^n a_{jk}^2.$$

„=“ genau dann, wenn $\sum_{k=1}^n a_{jk}a_{\ell k} = 0$ ist für $j \neq \ell$.

Beweis: Ist o.k., wenn $\det A = 0$ ist.

Sei $\det A > 0$ und setze

$$S_j^2 := \sum_{k=1}^n a_{jk}^2.$$

Es ist $S_j > 0$. Setze

$$b_{jk} = \frac{a_{jk}}{x_j}.$$

Dann ist

$$\sum_{j,k=1}^n b_{jk}^2 = n \Rightarrow (\det B)^2 \leq 1$$

$$(\det A)^2 = s_1^2 \cdots s_n^2 (\det B)^2 \leq x_1^2 \cdots x_n^2.$$

Gleichheit ist dabei genau für orthogonales B . Rest als Aufgabe.

