GMA

Grundlagen
Mathematik und Analysis

Mehrdimensionale Funktionen 1

Christian Cenker
Gabriele Uchida

Data Analytics and Computing

Mehrdimensionale Funktionen

Definition

Mehrdimensionale Funktion

$$f : \mathbb{R}^n \to \mathbb{R}^m \qquad f(x) = y$$

$$f(x) = f(x_1, \ldots, x_n) = \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = y$$

Koordinatenfunktionen

$$f_j : \mathbb{R}^n \to \mathbb{R}, \quad f_j(x_1, \ldots, x_n) = y_j \in \mathbb{R}, \; j = 1, \ldots, m$$

Ziel: (A) Stetigkeit
(B) Differenzierbarkeit
(C) Integrierbarkeit

Wir werden von den Eigenschaften reeller Funktionen ausgehen und in Analogie dazu die Eigenschaften für Komponentenfunktionen (mehrdimensionale reelle Funktionen, d. h., das Ergebnis ist in $\mathbb{R}$) betrachten.

Mehrdimensionale Funktionen

Graph

Definition

Der Graph einer mehrdimensionalen Funktion $f: \mathbb{R}^n \to \mathbb{R}^m$ ist die Menge

$$G_f = \{(x, f(x)) \mid x \in \mathbb{R}^n,\ f(x) \in \mathbb{R}^m\} \subseteq \mathbb{R}^n \times \mathbb{R}^m$$

Eine Niveaumenge einer mehrdimensionalen Funktion $f: \mathbb{R}^n \to \mathbb{R}^m$ ist eine Menge

$$N_c(f) = \{x \in \mathbb{R}^n \mid f(x) = c \in \mathbb{R}^m\}$$

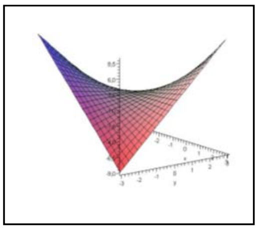
Mehrdimensionale Funktionen

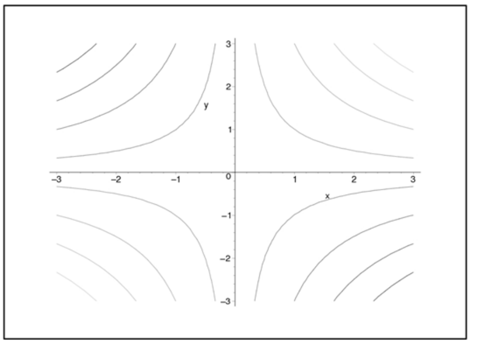
Niveaumengen

$$f: \mathbb{R}^2 \to \mathbb{R}, \quad f(x_1, x_2) = x_1 x_2$$

$$f(x_1, x_2) = x_1 x_2 = c \quad \Rightarrow \quad x_1 \neq 0: \ x_2 = \frac{c}{x_1} \quad \text{(Hyperbeln)}$$

$$c = 0 \quad \Rightarrow \quad x_1 = 0 \text{ oder } x_2 = 0$$





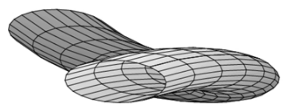
Kontourplot

Mehrdimensionale Funktionen

Niveaumengen

$f : \mathbb{R}^2 \to \mathbb{R}^2$

$$f(x_1, x_2) = \big(\cos(x_1) - 2\cos(0.4\, x_2), \sin(x_1) - 2\sin(0.4\, x_2) \big)$$



Graph/Kontourplot

Mehrdimensionale Funktionen

Stetigkeit

Definition

$f : \mathbb{R} \to \mathbb{R}$ ist stetig in x_0, wenn $\lim\limits_{x \to x_0} f(x) = f(x_0)$

$$\Leftrightarrow \quad \forall \varepsilon > 0 \;\exists \delta > 0 : |f(x) - f(x_0)| < \varepsilon \;\; \forall |x - x_0| < \delta$$

f ist stetig auf $\mathbb{R}$, wenn f in allen Punkten $x_0 \in \mathbb{R}$ stetig ist.

Die Stetigkeit wird über den Begriff des Grenzwertes bzw.
über einen Abstands- bzw. Längenbegriff eingeführt.

Im mehrdimensionalen Fall wird dies genauso geschehen.
Dazu müssen wir aber zuerst die Struktur des $\mathbb{R}^n$ betrachten (→ MBT).

Mehrdimensionale Funktionen

Der euklidische Vektorraum $\mathbb{R}^n$

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| x_1, \ldots, x_n \in \mathbb{R} \right\}$$

Vektoraddition $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$

Skalarmultiplikation $\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}, \lambda \in \mathbb{R}$

Skalarprodukt $\langle x, y \rangle = x^t y = x \cdot y = x_1 y_1 + \cdots + x_n y_n$

Länge (Norm) $\|x\| = \sqrt{x^t x} = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \cdots + x_n^2}$

Mehrdimensionale Funktionen

Der euklidische Vektorraum $\mathbb{R}^n$

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| x_1, \ldots, x_n \in \mathbb{R} \right\}$$

Länge (Norm) $\|x\| = \sqrt{x^t x} = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \cdots + x_n^2}$

Definition

Eine Abbildung $\| \cdot \| : \mathbb{R}^n \to \mathbb{R}$ heißt *Norm* genau dann, wenn folgende drei Bedingungen erfüllt sind:

(1) $\|x\| = 0 \Leftrightarrow x = 0, x \in \mathbb{R}^n, 0 \in \mathbb{R}^n$.

(2) $\|\lambda x\| = |\lambda| \, \|x\|, x \in \mathbb{R}^n, \lambda \in \mathbb{R}$.

(3) $\|x + y\| \leq \|x\| + \|y\|, x, y \in \mathbb{R}^n$ (Dreiecksungleichung)

Metrischer Raum $(\mathbb{R}^n, \| \cdot \|)$

Mehrdimensionale Funktionen

Folgen im metrischen Vektorraum $\mathbb{R}^n$

Definition

Eine Folge $(a_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}^n$ heißt *konvergent* gegen ihren Grenzwert oder Limes $a \in \mathbb{R}^n$ ($\lim_{n\to\infty} a_n = a$), wenn gilt:

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} : \; \|a_n - a\| < \varepsilon \quad \forall n \geq N \, .$$

Wir sprechen davon, dass ab einem gewissen N alle Folgenglieder in einer sogenannten ε-Umgebung von a liegen.

Definition

ε-Umgebung bzw. ε-Kugel

$$U_\varepsilon(x_0) = B_\varepsilon(x_0) = \{x \in \mathbb{R}^n \mid \|x - x_0\| < \varepsilon\}$$

Mehrdimensionale Funktionen

Stetige Funktionen $\mathbb{R}^n$

Definition

Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}^m$ heißt stetig in $x_0 \in \mathbb{R}^n$, wenn gilt:

$$\lim_{x\to x_0} f(x) = f(x_0)$$

$$\Leftrightarrow \quad \forall \varepsilon > 0 \; \exists \delta > 0 : \; \|f(x) - f(x_0)\| < \varepsilon \quad \forall \|x - x_0\| < \delta$$

f heißt stetig auf $A \subseteq \mathbb{R}^n$, wenn f stetig in jedem $x_0 \in A$ ist,

Mehrdimensionale Funktionen

Stetige Funktionen im $\mathbb{R}^n$

Bemerkung

f ist stetig in x_0, wenn

$$\lim_{k\to\infty} f(x_k) = f(x_0) \quad \forall (x_k) \subseteq \mathbb{R}^n \text{ mit } \lim_{k\to\infty} x_k = x_0\,,$$

dies gilt für alle denkbaren Folgen (x_k) mit Grenzwert x_0

Theorem

Eine Funktion $f: \mathbb{R}^n \to \mathbb{R}^m$ ist stetig auf $A \subseteq \mathbb{R}^n$,
wenn jede ihrer Komponentenfunktionen auf A stetig ist.

Mehrdimensionale Funktionen

Topologisches aus dem $\mathbb{R}^n$

Bevor wir den nächsten Schritt in Richtung Differenzierbarkeit tun,
müssen wir uns noch kurz mit der Topologie des $\mathbb{R}^n$ befassen,
d. h. mit Mengen im $\mathbb{R}^n$.

ε-Umgebung bzw. ε-Kugel

$$U_\varepsilon(x_0) = B_\varepsilon(x_0) = \{x \in \mathbb{R}^n \mid \|x - x_0\| < \varepsilon\}$$

Mehrdimensionale Funktionen

Spezielle Mengen im $\mathbb{R}^n$

Offene ε-Umgebung: $U_\varepsilon(a) = \{x \in \mathbb{R}^n \mid \|x - a\| < \varepsilon\}$

Abgeschlossene ε-Umgebung: $\bar{U}_\varepsilon(a) = \{x \in \mathbb{R}^n \mid \|x - a\| \leq \varepsilon\}$

Punktierte ε-Umgebung: $\dot{U}_\varepsilon(a) = U_\varepsilon(a) \setminus \{a\}$

Umgebung: Eine Menge M heißt Umgebung von $a \in \mathbb{R}^n$, wenn sie eine offene ε-Umgebung von a enthält, d. h., $U_\varepsilon(a) \subseteq M$.

Innerer Punkt: a heißt *innerer Punkt* von M, wenn eine Umgebung von a vollständig in M liegt, d. h., $\exists \varepsilon > 0 : U_\varepsilon(a) \subseteq M$.

Offene Menge: Eine Menge heißt *offen*, wenn sie nur aus inneren Punkten besteht.

Inneres: Das Innere M° einer Menge M ist die Menge ihrer inneren Punkte.

Mehrdimensionale Funktionen

Spezielle Mengen im $\mathbb{R}^n$

Isolierter Punkt: a heißt isolierter Punkt von M, wenn es mindestens eine punktierte Umgebung von a gibt, die nicht in M liegt, d. h., $\exists \dot{U}_\varepsilon(a) : \dot{U}_\varepsilon(a) \cap M = \emptyset$.

Randpunkt: a heißt *Randpunkt* der Menge M, falls in jeder Umgebung von a mindestens ein Punkt liegt, der in M liegt und mindestens ein Punkt, der nicht in M liegt: $(U_\varepsilon(a) \cap M \neq \emptyset) \wedge (U_\varepsilon(a) \cap M^c \neq \emptyset),\ \forall U_\varepsilon(a)$.

Rand ∂M: Der Rand ∂M einer Menge M ist die Menge all ihrer Randpunkte.

Abgeschlossene Menge: Eine Menge heißt abgeschlossen, wenn sie ihren Rand enthält.

Abgeschlossene Hülle: Die abgeschlossene Hülle $\bar{M}$ ist die Menge vereinigt mit ihrem Rand, d. h., $\bar{M} = M \cup \partial M$.

Beschränkt: Eine Menge M heißt beschränkt, wenn es ein $R > 0$ gibt, sodass $\|x\| < R$ für alle $x \in M$, oder anders ausgedrückt, $M \subseteq U_R(0)$.

Kompakt: Eine Menge M heißt kompakt, wenn sie beschränkt und abgeschlossen ist.

Mehrdimensionale Funktionen

Folgen und Mengen im $\mathbb{R}^n$

Theorem

Eine Menge M ist abgeschlossen, wenn mit jeder konvergenten Folge $(a_n) \subseteq M$ auch ihr Grenzwert in M liegt.

Theorem

M abgeschlossen $\Leftrightarrow M^c = \mathbb{R}^n \setminus M$ offen.

Definition

Eine Folge (a_k) heißt *Teilfolge* von (a_n), wenn sie durch Weglassen von Folgengliedern von (a_n) entsteht.

Theorem

Eine Menge M ist genau dann kompakt, wenn jede Folge in M eine in M konvergente Teilfolge besitzt.

Theorem

Konvergente Folgen sind beschränkt.

Bemerkung

Numerische Näherungslösungen iterativer Algorithmen sind nichts anderes als Glieder von Folgen im $\mathbb{R}^n$

Differenzialrechnung im $\mathbb{R}^n$

Definitionen

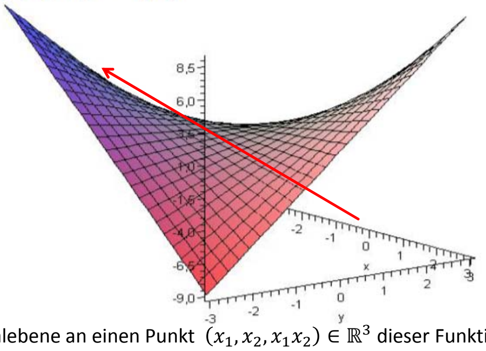
Differenzialquotient

$$f : \mathbb{R} \to \mathbb{R} \qquad f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \frac{d}{dx} f = \frac{df}{dx}$$

Wie sieht das im mehrdimensionalen Fall aus?

Differenzialrechnung im $\mathbb{R}^n$

$f : \mathbb{R}^2 \to \mathbb{R},\ f(x_1, x_2) = x_1 x_2$

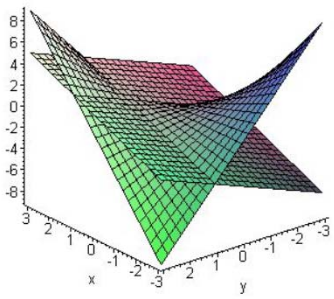


Die Tangentialebene an einen Punkt $(x_1, x_2, x_1x_2) \in \mathbb{R}^3$ dieser Funktion ist durch zwei linear unabhängige Richtungsvektoren gegeben.

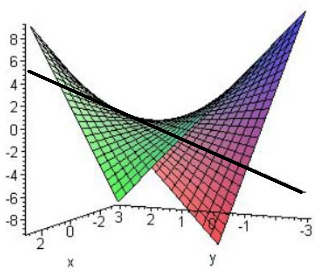
Wir bewegen uns entlang einer Gitternetzline parallel zur x_1-Achse und erhalten einen Richtungsvektor über die Tangente an die Gitternetzlinie.
Bei konstantem x_2 betrachten wir die Ableitung nach x_1.

Differenzialrechnung im $\mathbb{R}^n$

$f : \mathbb{R}^2 \to \mathbb{R},\ f(x_1, x_2) = x_1 x_2$



Tangentialebene



Projizierende Tangentialebene

Differenzialrechnung im $\mathbb{R}^n$

Partielle Ableitung

Definition

Die partielle Ableitung für reelle Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ ist definiert als

$$\frac{\partial}{\partial x_j} f(x_1, \ldots, x_n)$$
$$= \lim_{h \to 0} \frac{f(x_1, \ldots, x_{j-1}, x_j + h, x_{j+1}, \ldots, x_n) - f(x_1, \ldots, x_j, \ldots, x_n)}{h}$$

Bemerkung

Ableitung in Richtung x_j

x_j als Variable

alle anderen Argumente werden als konstant angesehen

Definition

Eine Funktion heißt *partiell differenzierbar* nach x_j im Punkt $x_0 \in \mathbb{R}^n$, wenn der obige Lines dort existiert.

Eine Funktion heißt *partiell differenzierbar* nach x_j auf $A \subseteq \mathbb{R}^n$, wenn sie in allen Punkten $x_0 \in A$ partiell differenzierbar ist.

Differenzialrechnung im $\mathbb{R}^n$

Partielle Ableitung, Notationen

$$\frac{\partial}{\partial x_j} f(x_1, \ldots, x_n)$$
$$= \lim_{h \to 0} \frac{f(x_1, \ldots, x_{j-1}, x_j + h, x_{j+1}, \ldots, x_n) - f(x_1, \ldots, x_j, \ldots, x_n)}{h}$$

$$\frac{\partial}{\partial x_j} f(x_1, \ldots, x_n) = \frac{\partial}{\partial x_j} f(x) = D_j f(x) = f_{.j}(x) = \frac{\partial f}{\partial x_j}|_x$$

Beispiel

$$f(x,y) = x \cdot y \qquad \frac{\partial f}{\partial x} = f_{.x}(x,y) = y \qquad \frac{\partial f}{\partial y} = f_{.y}(x,y) = x$$

Differenzialrechnung im $\mathbb{R}^n$

Gradient, erste Ableitung

Definition

Wenn für eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ alle partiellen Ableitungen existieren, so ist die erste Ableitung von f, der sogenannte *Gradient* $\nabla f(x)$, definiert als

$$\nabla f(x) = \operatorname{grad} f(x) = (\frac{\partial}{\partial x_1} f(x_1, \ldots, x_n), \ldots, \frac{\partial}{\partial x_n} f(x_1, \ldots, x_n))$$

Bemerkungen

Wir verwenden hier die Notation des Gradienten als Zeilenvektor,
da er dann direkt in Approximationsformeln etc. verwendet werden kann.
Manchmal wird der Gradient auch als Spaltenvektor definiert,
Insbesondere wenn durch ihn eine Richtung angezeigt werden soll.
Beides ist möglich und zulässig.

Der im Gradienten verwendete Nabla-Operator ∇ ist definiert als $\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$

Differenzialrechnung im $\mathbb{R}^n$

Höhere partielle Ableitungen

$$\frac{\partial}{\partial x_j}(\frac{\partial}{\partial x_k} f(x_1, \ldots, x_n)) = \frac{\partial^2}{\partial x_k \, \partial x_j} f(x_1, \ldots, x_n) = f_{.jk} = f_{.kj}$$

Beispiel

$$f(x, y) = x \cdot y \qquad \frac{\partial f}{\partial x} = f_{.x}(x, y) = y \qquad \frac{\partial f}{\partial y} = f_{.y}(x, y) = x$$

$$\frac{\partial^2 f}{\partial x \partial y} = 1 \qquad \frac{\partial^2 f}{\partial y \partial x} = 1$$

$$\frac{\partial^2 f}{\partial x^2} = 0 \qquad \frac{\partial^2 f}{\partial y^2} = 0$$

Differenzialrechnung im $\mathbb{R}^n$

Hessematrix, zweite Ableitung

Definition

Ist $f : \mathbb{R}^n \to \mathbb{R}$ jeweils (mindestens) zweimal partiell differenzierbar nach allen Koordinatenrichtungen, so ist die zweite Ableitung von f die Hessematrix

$$\nabla^2 f(x) = H_f(x) = \begin{pmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(x_1, \ldots, x_n) & \cdots & \frac{\partial^2}{\partial x_1 \partial x_n} f(x_1, \ldots, x_n) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_1} f(x_1, \ldots, x_n) & \cdots & \frac{\partial^2}{\partial x_n \partial x_n} f(x_1, \ldots, x_n) \end{pmatrix}$$

Theorem

Ist $f : \mathbb{R}^n \to \mathbb{R}$ zweimal partiell differenzierbar, so gilt

$$\frac{\partial^2}{\partial x_i \partial x_j} f(x_1, \ldots, x_n) = \frac{\partial^2}{\partial x_j \partial x_i} f(x_1, \ldots, x_n)$$

Die Hessematrix ist eine symmetrische $n \times n$-Matrix,

Differenzialrechnung im $\mathbb{R}^n$

Total differenzierbar

Definition

Eine Funktion $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ heißt in einem inneren Punkt $x_0 \in A$ (total) *differenzierbar*, wenn sie in x_0 partiell differenzierbar ist, alle partiellen Ableitungen in einer Umgebung von x_0 existieren und diese Ableitungen dort stetig sind.

Definition

Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ differenzierbar. Die lineare Abbildung

$$df(x_0) \doteq \nabla f(x_0) \cdot (x - x_0) = \frac{\partial}{\partial x_1} f(x_0) dx_1 + \cdots + \frac{\partial}{\partial x_n} f(x_0) dx_n ,$$

mit $dx_i \doteq (x_i - x_i^{(0)})$, heißt *totales Differential* der Funktion f in x_0. Das totale oder vollständige Differential wird oft auch mit $\frac{\partial f}{\partial x}$ bezeichnet.

Differenzialrechnung im $\mathbb{R}^n$

Total differenzierbar

Beispiel

$$f(x_1,x_2,x_3) = x_1x_2^2 + 2x_2\exp(x_3) + \exp(x_3^2)$$

$$\frac{\partial f}{\partial x_1} \qquad \frac{\partial f}{\partial x_2} \qquad \frac{\partial f}{\partial x_3}$$

$$\nabla f(x) = (x_2^2,\; 2x_1x_2 + 2\exp(x_3),\; 2x_2\exp(x_3) + 2x_3\exp(x_3^2))$$

$$\nabla^2 f(x) = \begin{pmatrix} 0 & 2x_2 & 0 \\ 2x_2 & 2x_1 & 2\exp(x_3) \\ 0 & 2\exp(x_3) & 2x_2\exp(x_3) + 2\exp(x_3^2) + 4x_3^2\exp(x_3^2) \end{pmatrix}$$

Differenzialrechnung im $\mathbb{R}^n$

Beispiel

$$f(x_1,x_2) = \begin{cases} \dfrac{x_1x_2}{\sqrt{x_1^2+x_2^2}} & (x_1,x_2) \neq (0,0) \\ 0 & (x_1,x_2) = (0,0) \end{cases}$$

f ist auf $\mathbb{R}^2 \setminus \{(0,0)\}$ differenzierbar

$$\frac{\partial}{\partial x_1} f(x_1,x_2) = \frac{x_2}{\sqrt{x_1^2+x_2^2}} - \frac{x_1^2x_2}{(x_1^2+x_2^2)^{\frac{3}{2}}}$$

$$\frac{\partial}{\partial x_1} f(0,x_2) = \frac{x_2}{x_2} = 1$$

$$\frac{\partial}{\partial x_1} f(x_1,0) = 0$$

Problem: $\lim\limits_{x_2\to 0} \frac{\partial}{\partial x_1} f(0,x_2) \neq \lim\limits_{x_1\to 0} \frac{\partial}{\partial x_1} f(x_1,0)$

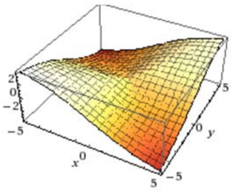
die partielle Ableitung $\frac{\partial}{\partial x_1} f$ ist in $(0,0)$ nicht stetig

daher ist f dort nicht differenzierbar

Differenzialrechnung im $\mathbb{R}^n$

Beispiel (Fortsetzung)

$$f(x_1,x_2) = \begin{cases} \dfrac{x_1x_2}{\sqrt{x_1^2+x_2^2}} & (x_1,x_2) \neq (0,0) \\ 0 & (x_1,x_2) = (0,0) \end{cases}$$

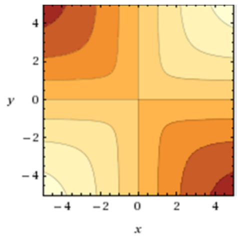


f ist symmetrisch bez. der Vertauschung von x_1 und x_2

f ist symmetrisch bez. der Vertauschung von x_1 und $-x_2$

$f(-x_1,x_2) = -f(x_1,x_2)$

$f(x_1,-x_2) = -f(x_1,x_2)$



Niveaumengen sind Hyperbeln

$$\frac{x_1x_2}{\sqrt{x_1^2+x_2^2}} = c \;\Leftrightarrow\; x_1x_2 = c\sqrt{x_1^2+x_2^2}$$

$$\Rightarrow\; x_1^2x_2^2 = c^2(x_1^2+x_2^2)$$

$$\Leftrightarrow\; x_2^2(x_1^2-c^2) = c^2x_1^2$$

$$\Rightarrow\; x_2 = \pm\frac{cx_1}{\sqrt{x_1^2-c^2}}$$

Differenzialrechnung im $\mathbb{R}^n$

Partielle Ableitungen

Theorem

Sei $A \subseteq \mathbb{R}^n$ eine offene Menge und $f : A \to \mathbb{R}^m$ zweimal stetig partiell differenzierbar. Dann gilt für alle $x_0 \in A$ und alle $i, j \in \{1, \dots, n\}$:

$$\frac{\partial^2}{\partial x_i \partial x_j} f(x_0) = \frac{\partial^2}{\partial x_j \partial x_i} f(x_0)$$

Theorem

Sei $A \subseteq \mathbb{R}^n$ eine offene Menge und $f : A \to \mathbb{R}^m$ k-mal stetig partiell differenzierbar. Dann gilt für jede Permutation π der Zahlen $1, \dots, k$:

$$\frac{\partial^k}{\partial x_{i_1} \partial x_{i_k}} f = \frac{\partial^k}{\partial x_{i_{\pi(1)}} \partial x_{i_{\pi(k)}}} f$$

Beispiel

$$\frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_3} = \frac{\partial^3 f}{\partial x_2 \partial x_3 \partial x_1} = \frac{\partial^3 f}{\partial x_3 \partial x_1 \partial x_2}$$

Differenzialrechnung im $\mathbb{R}^n$

Ableitungen für Funktionen $f: \mathbb{R}^n \to \mathbb{R}^m$

Bisher: Reelle Funktionen $f: \mathbb{R}^n \to \mathbb{R}$
Jetzt: Funktionen $f: \mathbb{R}^n \to \mathbb{R}^m$

$$f : \mathbb{R}^n \to \mathbb{R}^m \qquad f(x) = y$$

$$f(x) = f(x_1, \ldots, x_n) = \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = y$$

Jede Koordinatenfunktion f_j ist eine reelle Funktion $f_j: \mathbb{R}^n \to \mathbb{R}$

$$f_j : \mathbb{R}^n \to \mathbb{R}, \quad f_j(x_1, \ldots, x_n) = y_j \in \mathbb{R}, \; j = 1, \ldots, m$$

$\Longrightarrow$ Setze Ableitungen aus jenen der Koordinatenfunktionen zusammen.

Differenzialrechnung im $\mathbb{R}^n$

Jacobimatrix, Funktionalmatrix

Definition

Ist $f: \mathbb{R}^n \to \mathbb{R}^m$, $f(x) = \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix}$

Und sind alle Koordinatenfunktionen f_j partiell differenzierbar

$$f_j : \mathbb{R}^n \to \mathbb{R}, \quad f_j(x_1, \ldots, x_n) = y_j \in \mathbb{R}, \; j = 1, \ldots, m$$

so ist die erste Ableitung von f die Jacobimatrix (Funktionalmatrix)

$$\nabla f(x) = J_f(x) = \nabla \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} = \begin{pmatrix} \nabla f_1(x_1, \ldots, x_n) \\ \vdots \\ \nabla f_m(x_1, \ldots, x_n) \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial}{\partial x_1} f_1(x_1, \ldots, x_n) & \cdots & \frac{\partial}{\partial x_n} f_1(x_1, \ldots, x_n) \\ \vdots & \vdots & \vdots \\ \frac{\partial}{\partial x_1} f_m(x_1, \ldots, x_n) & \cdots & \frac{\partial}{\partial x_n} f_m(x_1, \ldots, x_n) \end{pmatrix}$$

Differenzialrechnung im $\mathbb{R}^n$

Jacobimatrix, Funktionalmatrix

$$\nabla f(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} f_1(x_1,\ldots,x_n) & \cdots & \frac{\partial}{\partial x_n} f_1(x_1,\ldots,x_n) \\ \vdots & \vdots & \vdots \\ \frac{\partial}{\partial x_1} f_m(x_1,\ldots,x_n) & \cdots & \frac{\partial}{\partial x_n} f_m(x_1,\ldots,x_n) \end{pmatrix}$$

Da wir m Komponentenfunktionen jeweils n-mal partiell ableiten, ist die Jacobimatrix eine $m \times n$-Matrix.

Ist die Jacobimatrix quadratisch, dann heißt ihre Determinante *Funktionaldeterminante*.

Bemerkung

Zweite Ableitungen von Funktionen $f: \mathbb{R}^n \to \mathbb{R}^m$ werden in geschlossener Form im Allgemeinen nicht mehr betrachtet, da wir jede partielle Ableitung wiederum partiell in n Richtungen ableiten können.
Das ergäbe einen „$m \times n \times n$-Würfel“, der nicht so leicht handhabbar ist.
Für Taylorreihen spielen höhere Ableitungen jedoch eine Rolle.

Differenzialrechnung im $\mathbb{R}^n$

Jacobimatrix, Funktionalmatrix

Beispiel

$$f(x_1,x_2,x_3) = \begin{pmatrix} 2x_1x_2^2 \\ x_2\exp(x_3^2) + 2x_1 \end{pmatrix}$$

$$\nabla f(x) = \begin{pmatrix} 2x_2^2 & 4x_1x_2 & 0 \\ 2 & \exp(x_3^2) & 2x_2x_3\exp(x_3^2) \end{pmatrix}$$

Beispiel

$$f(x) = \sin(x)\cos(x) \qquad f'(x) = \cos^2(x) - \sin^2(x)$$

$$f(x,y) = \sin(x)\cos(y) \qquad \nabla f(x,y) = (\cos(x)\cos(y), -\sin(x)\sin(y))$$

$$f(x,y) = \begin{pmatrix} \sin(x) \\ \cos(y) \end{pmatrix} \qquad \nabla f(x,y) = J_f(x,y) = \begin{pmatrix} \cos(x) & 0 \\ 0 & -\sin(y) \end{pmatrix}$$

Differenzialrechnung im $\mathbb{R}^n$

Differentiationsregeln

Wie im Eindimensionalen gelten auch hier weiterhin die bekannten Differentiationsregeln.

Sie können bewiesen werden, indem man die Funktionen und Ableitungen komponentenweise betrachtet.

Voraussetzung ist, dass die beteiligten Funktionen differenzierbar sind.

$\nabla(f+g)(x) = \nabla f(x) + \nabla g(x)$, $f, g : \mathbb{R}^n \to \mathbb{R}^m$ $(J_{f+g} = J_f + J_g)$

$\nabla(\lambda f)(x) = \lambda \nabla f(x)$, $\lambda \in \mathbb{R}$ $(J_{\lambda f} = \lambda J_f)$

Ist f in einer Umgebung $U_\varepsilon(x_0)$ konstant, so ist $\nabla f(x_0) = 0$

Ist f eine lineare Abbildung $(f(x) = A_f x)$, so ist $\nabla f(x) = A_f$

$\nabla(\frac{1}{f(x)}) = -\frac{\nabla f(x)}{f^2(x)}$, falls $f : \mathbb{R}^n \to \mathbb{R}$

Differenzialrechnung im $\mathbb{R}^n$

Differentiationsregeln

Theorem (Verallgemeinerte Kettenregel)

Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ und $g : \mathbb{R}^m \to \mathbb{R}^k$, f differenzierbar in x und g differenzierbar in $f(x) = y$, so ist $g \circ f : \mathbb{R}^n \to \mathbb{R}^k$ und $(g \circ f)(x) = g(f(x))$ differenzierbar in x und es gilt:

$$\nabla(g \circ f)(x) = \nabla g(f(x)) \cdot \nabla f(x)$$

Differenzialrechnung im $\mathbb{R}^n$

Verallgemeinerte Kettenregel

$$\nabla(g \circ f)(x) = \nabla g(f(x)) \cdot \nabla f(x)$$

Beispiel

$$f(x_1, x_2) = \begin{pmatrix} x_2 - x_1 \\ \exp(x_2) \end{pmatrix} \qquad g(x_1, x_2) = \begin{pmatrix} x_1 + x_2 \\ x_2^2 \\ x_1 \end{pmatrix}$$

$$\nabla f(x_1, x_2) = \begin{pmatrix} -1 & 1 \\ 0 & \exp(x_2) \end{pmatrix} \qquad \nabla g(x_1, x_2) = \begin{pmatrix} 1 & 1 \\ 0 & 2x_2 \\ 1 & 0 \end{pmatrix}$$

$$(g \circ f)(x_1, x_2) = \begin{pmatrix} x_2 - x_1 + \exp(x_2) \\ \exp(2x_2) \\ x_2 - x_1 \end{pmatrix}$$

$$\nabla(g \circ f)(x_1, x_2) = \nabla g(f(x_1, x_2)) \cdot \nabla f(x_1, x_2)$$

$$\nabla(g \circ f) = \begin{pmatrix} 1 & 1 \\ 0 & 2\exp(x_2) \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 & 1 \\ 0 & \exp(x_2) \end{pmatrix} = \begin{pmatrix} -1 & 1 + \exp(x_2) \\ 0 & 2\exp(2x_2) \\ -1 & 1 \end{pmatrix}$$

Differenzialrechnung im $\mathbb{R}^n$

Richtungsableitung

Sei $f : \mathbb{R}^2 \to \mathbb{R}$. Wir wollen die Ableitung von f im Punkt x_0 in Richtung eines Vektors $a \in \mathbb{R}^2$ betrachten, wobei a normiert sein soll, d. h. $\|a\| = 1$ (die Länge von a soll sich nicht auf den Wert der Ableitung auswirken).

Sei $h(t) = x_0 + t \cdot a$, dann ist $\nabla h(t) = a$, $t \in \mathbb{R}$.

Für $x = x_0 + ta = h(t)$ gilt:

$$f(x) = f(x_0 + ta) = f(h(t)) = (f \circ h)(t)$$

also $\quad \nabla(f \circ h)(t) = \nabla f(x_0 + ta) \cdot \nabla h(t)$

Also ist in x_0, d. h. $t = 0$: $\quad \nabla(f \circ h)(0) = \nabla f(x_0) \cdot a = \langle \nabla f(x_0), a \rangle$

Die Richtungsableitung ist also nichts anderes als das innere Produkt (Skalarprodukt) des Gradienten von f mit dem Richtungsvektor a.

Differenzialrechnung im $\mathbb{R}^n$

Richtungsableitung

Definition

Die Richtungsableitung einer Funktion $f : \mathbb{R}^n \to \mathbb{R}^m$ in Richtung $a \in \mathbb{R}^n$, mit $\|a\| = 1$, in einem Punkt x_0 ist definiert durch:

$$D_a f(x_0) = \nabla f(x_0) \cdot a = \langle \nabla f(x_0), a \rangle$$

Bemerkung

Für die Richtungsableitung gilt

$$D_a f(x_0) = \lim_{t \to 0} \frac{f(x_0 + ta) - f(x_0)}{t}$$

und

$$D_{e_k} f(x_0) = \frac{\partial}{\partial x_k} f(x_0)$$

wobei e_k den k-ten Einheitsvektor (die k-te Einheitsrichtung) bezeichnet.

Differenzialrechnung im $\mathbb{R}^n$

Richtungsableitung

Für welchen Vektor a ist die Richtungsableitung maximal?

Da diese das innere Produkt von a mit dem Gradienten von f ist,
ist sie proportional zum Cosinus des eingeschlossenen Winkels.

Da der Cosinus von 0° am größten ist, muss a parallel zum Gradienten liegen, also

$$a = \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$$

Der Gradient zeigt in Richtung des größten Anstiegs.

Bemerkung:
Wenn wir eine Funktion maximieren wollen und irgendwo anfangen,
werden wir in Richtung des Gradienten weitergehen (Gradientenverfahren).
Ist der Gradient gleich 0l, so sind wir in einem stationären Punkt angekommen,
d. h. wir kommen nicht weiter.
Wir könnten uns damit in einem Maximum befinden.

Differenzialrechnung im $\mathbb{R}^n$

Mittelwertsätze, Taylorreihen

Wie im eindimensionalen Fall, wo für eine differenzierbare Funktion $f : \mathbb{R} \to \mathbb{R}$ in einer Umgebung von x_0 gilt, dass $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$ (Definition des Differentialquotienten, Tangentengleichung), gibt es auch im mehrdimensionalen Fall solche Approximationen.

Theorem

Sei $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ eine reelle, auf der offenen Menge A differenzierbare Funktion und enthalte A die Strecke S von x nach x_0. Dann gibt es einen Punkt $\xi \in S$ auf dieser Strecke, für den gilt

$$f(x) = f(x_0) + \nabla f(\xi)(x - x_0)$$

bzw. $$f(x) = f(x_0) + \nabla f(x_0)(x - x_0) + O(\|x - x_0\|^2)\,.$$

Ist f zweimal stetig differenzierbar, so gilt

$$f(x) = f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^t \nabla^2 f(\xi)(x - x_0)$$

bzw.

$$f(x) = f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^t \nabla^2 f(x_0)(x - x_0) + O(\|x - x_0\|^3)$$

Differenzialrechnung im $\mathbb{R}^n$

Landau´sche O-Symbole

o (Klein-o) : $f(x) = o(g(x))$ mit $x \to a$, wenn $\lim_{x \to a} \frac{|f(x)|}{|g(x)|} = 0$

O (Groß-O) : $f(x) = O(g(x))$ mit $x \to a$, wenn $\exists C > 0 : \ |f(x)| \leq C|g(x)|$

Da im obigen Fall die Funktion $g(x) = \|x - x_0\|^2$ bzw. $g(x) = \|x - x_0\|^3$ mit $x \to x_0$ schnell gegen 0 geht, heißt das, dass der Fehler $O(g(x))$ mindestens genauso schnell gegen 0 geht: $\|x - x_0\| = h \ll 1$, es gilt nun für dieses h: h^3 geht schneller gegen 0 als h^2, wenn h gegen 0 geht.

Differenzialrechnung im $\mathbb{R}^n$

Mittelwertsätze, Taylorreihen

Taylorreihe erster Ordnung, lineare Approximation

Sei $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ eine reelle, auf der offenen Menge A differenzierbare Funktion. Dann

$$f(x) = f(x_0) + \nabla f(x_0)(x - x_0) + O(\|x - x_0\|^2)\,.$$

Taylorreihe zweiter Ordnung, quadratische Approximation

Ist f zweimal stetig differenzierbar, so gilt

$$f(x) = f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^t \nabla^2 f(x_0)(x - x_0) + O(\|x - x_0\|^3)$$

Differenzialrechnung im $\mathbb{R}^n$

Mittelwertsätze, Taylorreihen

Beispiel

Sei $f(x_1, x_2) = \begin{pmatrix} x_2 - x_1 \\ \exp(x_2) \end{pmatrix}$

dann ist $\nabla f(x_1, x_2) = \begin{pmatrix} -1 & 1 \\ 0 & \exp(x_2) \end{pmatrix}$

und in der Nähe des Ursprungs $x_0 = (0, 0)$ gilt

$$f(x_1, x_2) \approx f(x_0) + \nabla f(x_0)(x - x_0) =$$

$$= \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 - x_1 \\ 1 + x_2 \end{pmatrix}.$$

Differenzialrechnung im $\mathbb{R}^n$

Tangentialabbildung

Definition

Sei $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ in einem inneren Punkt von A differenzierbar.

Die Funktion $\quad g(x) = f(x_0) + \nabla f(x_0)(x - x_0)$

ist eine affin lineare Abbildung und heißt *Tangentialabbildung* von f.

Beispiel

Für $f : \mathbb{R}^2 \to \mathbb{R}$ und $x_0 = (x_1^{(0)}, x_2^{(0)}) \in \mathbb{R}^2$, sowie $x = (x_1, x_2)$ ist der Graph von

$$\begin{aligned} g(x_1, x_2) &= f(x_0) + \nabla f(x_0)(x - x_0) \\ &= f(x_1^{(0)}, x_2^{(0)}) + \tfrac{\partial}{\partial x_1} f(x_1^{(0)}, x_2^{(0)})(x_1 - x_1^{(0)}) + \tfrac{\partial}{\partial x_2} f(x_1^{(0)}, x_2^{(0)})(x_2 - x_2^{(0)}) \end{aligned}$$

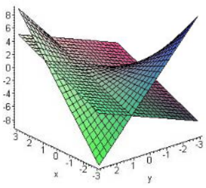
die *Tangentialebene* an f im Punkt $x_0 = (x_1^{(0)}, x_2^{(0)})^t$ (Ortsvektor).

Differenzialrechnung im $\mathbb{R}^n$

Tangentialabbildung

Beispiel

An die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = x_1 x_2$ soll im Funktionswert des Punktes $x_0 = (x_1^{(0)}, x_2^{(0)}) = (1, 1)$ eine Tangentialebene gelegt werden.



$$\begin{aligned} g(x_1, x_2) &= f(x_0) + \nabla f(x_0)(x - x_0) \\ &= x_1^{(0)} \cdot x_2^{(0)} + (x_2^{(0)}, x_1^{(0)}) \cdot \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \end{pmatrix} \right) \\ &= 1 \cdot 1 + (1, 1) \cdot \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \\ &= 1 + (1, 1) \cdot \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix} \\ &= 1 + (x_1 - 1) + (x_2 - 1) \\ &= x_1 + x_2 - 1 \end{aligned}$$

Differenzialrechnung im $\mathbb{R}^n$

Das totale Differential

Für eine differenzierbare Funktion $f : \mathbb{R}^n \to \mathbb{R}^m$ liefert die Tangentenabbildung

$$f_t(x) \doteq f(x_0) + \nabla f(x_0)(x - x_0)$$

in der Umgebung von x_0 eine gute Näherung der Funktion f, also

$$f_t(x) - f(x_0) = \nabla f(x_0)(x - x_0) \approx f(x) - f(x_0).$$

Definition

Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ differenzierbar. Die lineare Abbildung

$$df(x_0) \doteq \nabla f(x_0) \cdot (x - x_0) = \frac{\partial}{\partial x_1} f(x_0) dx_1 + \cdots + \frac{\partial}{\partial x_n} f(x_0) dx_n$$

mit $dx_i \doteq (x_i - x_i^{(0)})$, heißt *totales Differential* der Funktion f in x_0. Das totale oder vollständige Differential wird oft auch mit $\frac{\partial f}{\partial x}$ bezeichnet.

Das totale Differential gibt also die ungefähre Abweichung des Funktionswertes $f(x)$ an, wenn x von x_0 wenig abweicht.

Differenzialrechnung im $\mathbb{R}^n$

Das totale Differential

Beispiel (Fehlerfortpflanzungsgesetz)

Gegeben sei eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ von n Messgrößen $x_1, \ldots, x_n$. Die Messungenauigkeit betrage $\pm\Delta x_1, \ldots, \pm\Delta x_n$, wobei $\Delta x_i > 0$ für alle $i = 1, \ldots, n$.

Dann gilt für den Fehler $\Delta f(x)$ von $f(x_1, \ldots, x_n)$:

$$\Delta f(x_1, \ldots, x_n) < \left| \frac{\partial}{\partial x_1} f(x) \right| \Delta x_1 + \cdots + \left| \frac{\partial}{\partial x_n} f(x) \right| \Delta x_n$$

Definition

Die Größe

$$\Delta f_{max}(x_1, \ldots, x_n) := \left| \frac{\partial}{\partial x_1} f(x) \right| \Delta x_1 + \cdots + \left| \frac{\partial}{\partial x_n} f(x) \right| \Delta x_n$$

heißt der *maximale Fehler* von f.

Der Quotient $\frac{df_{max}(x)}{df(x)}$ heißt der *relative maximale Fehler* von f.

Differenzialrechnung im $\mathbb{R}^n$

Das totale Differential

Beispiel (Fehlerfortpflanzungsgesetz)

Aus der Physik wissen wir: $s = v \cdot t$, d.h., *Weg=Geschwindigkeit × Zeit*, mit den SI-Einheiten: $[m] = [m/s] \cdot [s]$.

Ein Marathonläufer soll über eine Strecke von $s = 10\,\text{km}$ eine Zeit von $t = 30\,\text{min}$ erzielen. D. h., die mittlere Geschwindigkeit ist $v = \frac{s}{t} = 20\,\text{km/h}$.

Der Messfehler für die Strecke s sei $\Delta s = \pm 1\,\text{m}$, jener der Zeit $\Delta t = \pm 1\,\text{sec}$. Wie groß ist dann Δv maximal?

$$\begin{aligned}
\Delta v(s,t) &= \left|\frac{\partial}{\partial s} v(s,t)\right| \Delta s + \left|\frac{\partial}{\partial t} v(s,t)\right| \Delta t \\
&= \left|\frac{1}{t}\right| \Delta s + \left|-\frac{s}{t^2}\right| \Delta t \\
&= \frac{0.001}{0.5}\frac{\text{km}}{\text{h}} + \frac{10}{0.25}\frac{\text{km}}{\text{h}^2}\frac{1}{3600}\text{h} \quad = 0.013\,\text{km/h}
\end{aligned}$$

$$\Rightarrow v = (20 \pm 0.013)\,\text{km/h}$$

Differenzialrechnung im $\mathbb{R}^n$

Die Jacobimatrix der Umkehrabbildung

Wir erinnern uns, dass im eindimensionalen Fall für die Umkehrabbildung f^{-1} einer differenzierbaren Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) = y$ gilt (falls sie existiert):

$$(f^{-1}(y))' = \frac{1}{f'(f^{-1}(y))}.$$

Das ist ableitbar aus $f \circ f^{-1} = id$ und der Kettenregel bei der Ableitung dieser Gleichung.

Ähnliches gilt auch für differenzierbare Funktionen $f : \mathbb{R}^n \to \mathbb{R}^n$, nur dass es sich bei der ersten Ableitung um eine Matrix handelt.

$$\nabla f^{-1}(y) = (\nabla f(f^{-1}(y)))^{-1}$$

Differenzialrechnung im $\mathbb{R}^n$

Die Jacobimatrix der Umkehrabbildung

Theorem

Sei

$$f : \mathbb{R}^n \to \mathbb{R}^n \qquad f(x) = \begin{pmatrix} f_1(x_1,\dots,x_n) \\ \vdots \\ f_n(x_1,\dots,x_n) \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

differenzierbar und

$$h : \mathbb{R}^n \to \mathbb{R}^n \qquad h(y) = \begin{pmatrix} h_1(y_1,\dots,y_n) \\ \vdots \\ h_n(y_1,\dots,y_n) \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

mit $h = f^{-1}$, also $h \circ f = f \circ h = id$

so gilt $\qquad \nabla h(y_1,\dots,y_n) = [\nabla f(h(y_1,\dots,y_n))]^{-1}$

Beweis. Es gilt $f \circ h = id$, daher $\nabla(f \circ h) = \nabla id = I$, die $n \times n$ Einheitsmatrix. Wegen der allgemeinen Kettenregel gilt nun

$$\nabla(f \circ h)(x) = \nabla f(h(x)) \cdot \nabla h(x) = I$$

und daher $\qquad \nabla h(x) = [\nabla f(h(x))]^{-1}$

Differenzialrechnung im $\mathbb{R}^n$

Die Jacobimatrix der Umkehrabbildung

Beispiel

$$f(x_1,x_2) = \begin{pmatrix} x_2 - x_1 \\ \exp(x_2) \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\nabla f(x_1,x_2) = \begin{pmatrix} -1 & 1 \\ 0 & \exp(x_2) \end{pmatrix}$$

Die Umkehrabbildung zu f ist $\quad h(y_1,y_2) = \begin{pmatrix} \log y_2 - y_1 \\ \log y_2 \end{pmatrix}$

$$\nabla h(y_1,y_2) = [\nabla f(h(y_1,y_2))]^{-1}$$

$$= \begin{pmatrix} -1 & 1 \\ 0 & y_2 \end{pmatrix}^{-1} = -\frac{1}{y_2}\begin{pmatrix} y_2 & -1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 & \frac{1}{y_2} \\ 0 & \frac{1}{y_2} \end{pmatrix}$$

Differenzialrechnung im $\mathbb{R}^n$

Die Jacobimatrix der Umkehrabbildung

Theorem (Satz von der Umkehrabbildung)

Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ stetig differenzierbar und $J_f(x_0)$ regulär, dann gibt es eine Umgebung $U_\varepsilon(x_0)$ und $V_\delta(f(x_0))$ und eine Umkehrfunktion $f^{-1} : V_\delta \to U_\varepsilon$ von f auf U_ε, die stetig differenzierbar ist.

Es ist

$$J_{f^{-1}}(f(x_0)) = (J_f(x_0))^{-1}$$

Differenzialrechnung im $\mathbb{R}^n$

Die Jacobimatrix der Umkehrabbildung

Beispiel

Die Funktion $f(x_1, x_2) = \begin{pmatrix} e^{x_1} \cos x_2 \\ e^{x_1} \sin x_2 \end{pmatrix}$ ist nicht global invertierbar.

Es ist $f(0,0) = f(0, 2\pi) = (1,0)^t$.

Es ist jedoch $J_f(x_1, x_2) = \begin{pmatrix} e^{x_1} \cos x_2 & -e^{x_1} \sin x_2 \\ e^{x_1} \sin x_2 & e^{x_1} \cos x_2 \end{pmatrix}$

und die **Funktionaldeterminante** $\det J_f = e^{2x_1} > 0$ und damit J_f regulär, d. h. invertierbar, daher ist f überall lokal invertierbar und es gilt:

$$x_0 = (0, \frac{\pi}{2}) \qquad J_f(0, \frac{\pi}{2}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad f(0, \frac{\pi}{2}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$J_{f^{-1}}(f(0, \frac{\pi}{2})) = J_{f^{-1}}(0,1) = J_f(0, \frac{\pi}{2})^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Differenzialrechnung im $\mathbb{R}^n$

Die Jacobimatrix der Umkehrabbildung

Beispiel

Die komplexe Funktion $f : \mathbb{C} \to \mathbb{C}$, $z \mapsto w = z^2$ kann als mehrdimensionale Funktion $g : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (u, v)$ aufgefasst werden, mit $z = x + iy$ und $w = u + iv$, also $u = x^2 - y^2$ und $v = 2xy$.

Es gilt

$$J_f(z) = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix} \quad \Rightarrow \quad \det J_f(z) = 4(x^2 + y^2) = 4|z|^2$$

d. h., für alle $z \neq 0$ gibt es eine Umgebung, die bijektiv auf z^2 abgebildet wird! Daher gibt es dort auch eine (lokale) Umkehrfunktion.