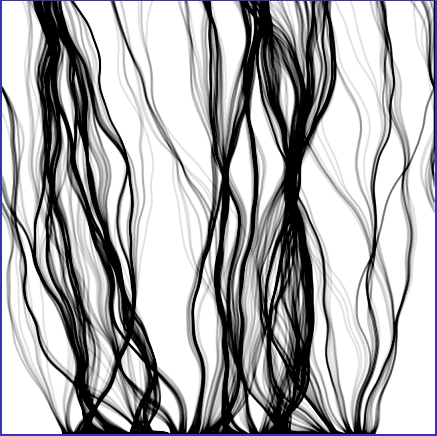
Angewandte Mathematik



Dr. rer. nat. Johannes Riesterer

Für fast alle (hinreichend große) Mathlib

Für einen Filter I bedeutet die Bedingung $\forall^f x\ p(x)$ dass die Menge der Elemente, für die $p(x)$ gilt, ein Element des Filters I ist, also $\{X \mid p(x)\} \in I$.

Definition

Big O Notation Mathlib

Für einen Filter l und Funktionen f, g definieren wir

$$f = \mathcal{O}[l]g \leftrightarrow \exists c > 0, \forall^{\mathrm{f}} x \in l, \|f(x)\| \leq c \cdot \|g(x)\|$$

Erklärung

Die Aussage besagt, dass für fast alle x in der Menge l, die Norm von $f(x)$ durch ein Vielfaches der Norm von $g(x)$ beschränkt ist. Das Vielfache wird durch die Konstante c dargestellt.

Definition

Klein-o Notation Mathlib

Für einen Filter l und Funktionen f, g definieren wir

$$f = o[l]g \leftrightarrow \forall c > 0, \forall^{\mathrm{f}} x \in l, \|f(x)\| \leq c \cdot \|g(x)\| \text{ für } x \geq N$$

Erklärung

Die Aussage besagt, dass für jede positive Konstante c und für fast alle x in der Menge l die Norm von $f(x)$ kleiner oder gleich $c \cdot \|g(x)\|$ ist. Dies beschreibt, dass $f(x)$ asymptotisch schneller gegen 0 geht als $g(x)$. In anderen Worten, der Ausdruck $\frac{\|f(x)\|}{\|g(x)\|}$ geht gegen 0 entlang l, wobei mögliche Probleme durch Division durch Null durch diese Definition vermieden werden.

Beweis: Klein-o impliziert Groß-O Mathlib

Wenn $f = o[l]g$, dann ist $f = O[l]g$.

Ableitungsdefinitionen

Klassische Definition in einer Dimension:

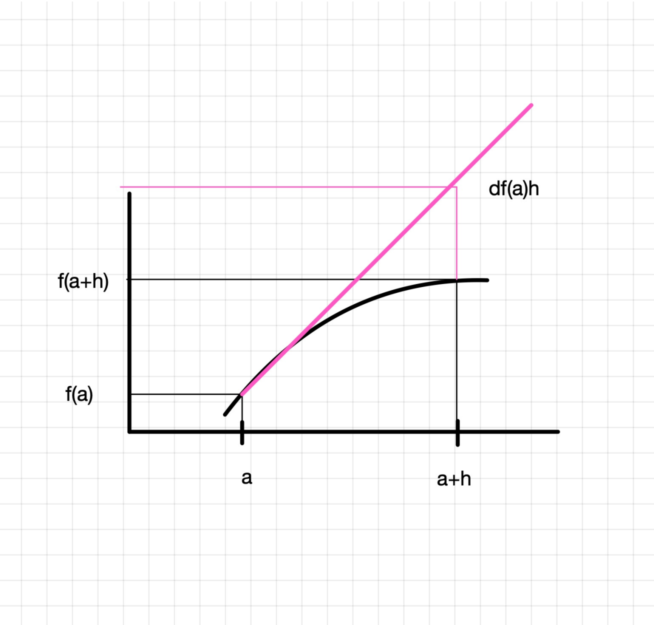
$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

Definition mit o-Kalkül:

$$f(a+h) = f(a) + f'(a)h + o(h)$$

Äquivalenz:

1. **Von $o(h)$ zur klassischen Definition:**
   $\frac{f(a+h)-f(a)}{h} = f'(a) + \frac{o(h)}{h}$
   Mit $\lim_{h \to 0} \frac{o(h)}{h} = 0$ folgt die klassische Definition.
2. **Von der klassischen Definition zur $o(h)$-Form:**
   Nehmen wir $f'(a) = \lim_{h \to 0} \frac{f(a+h)-f(a)}{h}$, dann:
   $f(a+h) = f(a) + f'(a)h + r(h)$, wobei $r(h)$ der Restterm ist.
   Da $\lim_{h \to 0} \frac{r(h)}{h} = 0$, gilt $r(h) = o(h)$.



df(a)h
f(a+h)
f(a)
a
a+h

Ableitung in mehreren Dimensionen

Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}^m$ hat die (lineare) Abbildung df als Ableitung, falls

$$f(a+h) = f(a) + df(a)h + o(h), \tag{1}$$

Die Ableitung wird auch mit $f' := df$ bezeichnet.

Eindeutigkeit

Die Ableitung df ist eindeutig bestimmt.

Beweis

Ist df' eine weiter Abbildung mit Eigenschaft (1), so gilt für jeden Basisvektor e_i

$$\lim_{t\to 0} \frac{f(a+te_i) - f(a) - df(a)te_i}{||te_i||} = 0 \tag{2}$$

$$\lim_{t\to 0} \frac{f(a+te_i) - f(a) - df'(a)te_i}{||te_i||} = 0 \tag{3}$$

$$\Rightarrow (df(a) - df'(a))(e_i) = lim_{t\to 0} \frac{(df'(a) - df(a))(te_i)}{||te_i||} = 0 \tag{4}$$

Definition der Ableitung entlang eines Filters

Ableitung Lean

Eine Funktion $f : V \to W$ zwischen normierten Räumen hat die lineare Abbildung f' als Ableitung entlang des Filters L, falls

$$f(x') = f(x) + f'(x' - x) + o(x' - x),$$

Lineare Abbildungen

Definition

Eine Abbildung $T : V \to W$ zwischen zwei Vektorräumen V und W über einem Körper K heißt **linear**, wenn für alle $v_1, v_2 \in V$ und alle $\alpha, \beta \in K$ gilt:

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2).$$

Eigenschaften

- Lineare Abbildungen erhalten die Vektorraumstruktur: Sie respektieren die Addition und die Skalarmultiplikation.
- Jede lineare Abbildung ist durch ihr Verhalten auf einer Basis des Vektorraums eindeutig bestimmt.
- Die Ableitung einer linearen Abbildung ist die Abbildung selbst: Für eine lineare Abbildung T gilt $D(T) = T$.

Lineare Abbildungen

Beispiele

- Die Identitätsabbildung $\mathrm{id}_V : V \to V$, definiert durch $\mathrm{id}_V(v) = v$, ist linear.
- Projektionen und Rotationen in $\mathbb{R}^n$ sind lineare Abbildungen.
- Matrizen wirken als lineare Abbildungen auf Vektorräumen.

Vergleich: Lineare Abbildungen in Mathematik und Lean (Mathlib)

Definition in Lean (Mathlib)

In Lean4 (Mathlib) wird eine lineare Abbildung T zwischen zwei normierten Vektorräumen V und W über $\mathbb{R}$ als eine stetige lineare Abbildung (`continuous_linear_map`) definiert:

$$T : V \to L[\mathbb{R}]W$$

Die lineare Struktur wird durch zwei Eigenschaften beschrieben:

- `map_add` : $T(v_1 + v_2) = Tv_1 + Tv_2$
- `map_smul` : $T(c \cdot v) = c \cdot Tv$

Vergleich

Im endlichdimensionalen Fall entspricht die Definition in Lean der üblichen Definition einer linearen Abbildung. Im unendlichdimensionalen Fall wird zusätzlich die Stetigkeit gefordert, da diese nicht automatisch gegeben ist.

Beispiel

$A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$

$$f(x) := A \cdot x + b \tag{5}$$

$$df(a) := A \tag{6}$$

Beweis

$$\lim_{h \to 0} \frac{A(x+h) - A \cdot x - A \cdot h}{||h||} = \tag{7}$$

$$\lim_{h \to 0} \frac{A \cdot x + A \cdot h - A \cdot x - A \cdot h}{||h||} = 0 \tag{8}$$

Vergleich: Lineare Abbildungen in Mathematik und Lean (Mathlib)

Partielle Ableitung

In Lean4 (Mathlib) wird eine lineare Abbildung T zwischen zwei normierten Vektorräumen V und W über $\mathbb{R}$ als eine stetige lineare Abbildung (`continuous_linear_map`) definiert:

$$T : V \to L[\mathbb{R}]W$$

Die lineare Struktur wird durch zwei Eigenschaften beschrieben:

- `map_add` : $\quad T(v_1 + v_2) = Tv_1 + Tv_2$
- `map_smul` : $\quad T(c \cdot v) = c \cdot Tv$

Vergleich: Lineare Abbildungen in Mathematik und Lean (Mathlib)

Ableitung Berechnen

Wie kann man die Ableitung einer Funktion berechnen?

Vergleich: Lineare Abbildungen in Mathematik und Lean (Mathlib)

Partielle Ableitung

Für eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ definert man die partielle Ableitung

$$D_i f(x) = \lim_{h \to 0} \frac{f(x + h \cdot e_i) - f(x)}{h}$$

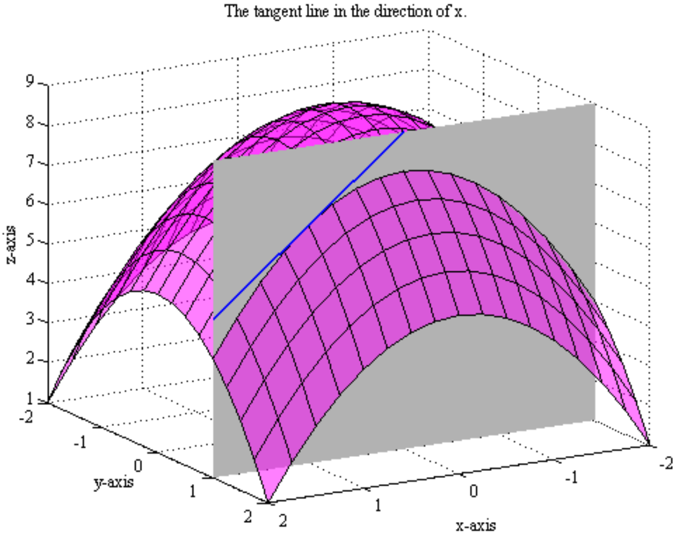
wobei e_i der i-te Einheitsvektor ist.
$\frac{\partial f(x)}{\partial x_i} := D_i f(x)$.

Richtungsableitung

Allgemeiner definiert man für $f : \mathbb{R}^n \to \mathbb{R}$ und $v \in \mathbb{R}^n$

$$D_v f(x) = \lim_{h \to 0} \frac{f(x + h \cdot v) - f(x)}{h}$$

die Richtungsableitung von f an der Stelle x in Richtung v.


The tangent line in the direction of x.
z-axis
9
8
7
6
5
4
3
2
1
-2
-1
0
y-axis
1
2
2
1
0
-1
-2
x-axis

Gradient

Der Vektor

$$\nabla f(a) := \begin{pmatrix} \frac{\partial f(a)}{\partial x_1} \\ \vdots \\ \frac{\partial f(a)}{\partial x_n} \end{pmatrix}$$

wird als Gradient bezeichnet.

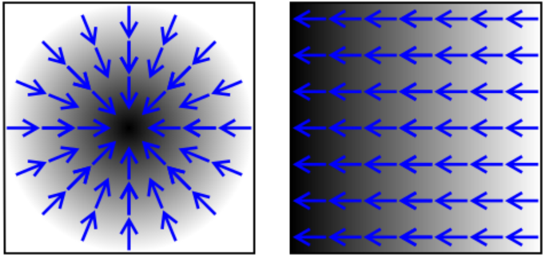


Figure: Quelle: Wikipedia:

Beispiel 1: Quadratische Funktion in 2D

$$f(x, y) = x^2 + y^2$$

Gradient:

$$\nabla f(x, y) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

Beispiel 2: Quadratische Funktion in 3D

$$f(x, y, z) = x^2 + y^2 + z^2$$

Gradient:

$$\nabla f(x, y, z) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}$$

Gradienten-Beispiele: Komplexere Funktionen

Beispiel 3: Exponentialfunktion

$$f(x, y) = e^{x^2+y^2}$$

Gradient:

$$\nabla f(x, y) = \begin{pmatrix} \frac{\partial}{\partial x} e^{x^2+y^2} \\ \frac{\partial}{\partial y} e^{x^2+y^2} \end{pmatrix} = \begin{pmatrix} 2xe^{x^2+y^2} \\ 2ye^{x^2+y^2} \end{pmatrix}$$

Beispiel 4: Logarithmische Funktion

$$f(x, y) = \ln(x^2 + y^2)$$

Gradient:

$$\nabla f(x, y) = \begin{pmatrix} \frac{\partial}{\partial x} \ln(x^2 + y^2) \\ \frac{\partial}{\partial y} \ln(x^2 + y^2) \end{pmatrix} = \begin{pmatrix} \frac{2x}{x^2+y^2} \\ \frac{2y}{x^2+y^2} \end{pmatrix}$$

Vergleich: Lineare Abbildungen in Mathematik und Lean (Mathlib)

Richtungsableitungv und Ableitung

$$D_v f(x) = df(x)v$$

und damit insbesondere $df(a) \cdot h = \langle \nabla f(a), h \rangle$.

Beweisg

Folgt aus der Eindeutigkeit der Ableitung.

Gradient

Sei $f : U \to \mathbb{R}$ differenzierbare Funktion, $a \in U$ und $v := \operatorname{argmax}_{||h||=1}\{\partial_h f(a)\}$. Dann gilt

$$||\nabla f(a)||v = \nabla f(a) .$$

Gradient

Der Gradient zeigt in die Richtung des steilsten Anstiegs.

Beweis

Für beliebiges h gilt

$$\partial_h f(a) = df(a)h = \langle \nabla f(a), h \rangle = ||\nabla f(a)|| \cdot ||h|| \cdot \cos(\varphi)$$

wobei φ den Innenwinkel zwischen $\nabla f(a)$ und h bezeichnet. Für $||h|| = 1$ wird somit $\partial_h f(a)$ maximal, wenn $\varphi = 0$ und somit$h = \frac{\nabla f(a)}{||\nabla f(a)||}$ ist.

Extrema

Sei $f : X \subset \mathbb{R}^n \to \mathbb{R}$ eine relle Funktion. Ein Punkt $a \in X$ heißt lokales Maximum bzw. Minimum, falls eine Umgebung U von a existiert, so dass $f(x) \leq f(a)$ bzw. $f(x) \geq f(a)$ für alle $x \in U$ gilt. Liegt einer der beiden Fälle vor, so spricht man von einem lokalen Extremum. Gilt strikt $f(x) < f(a)$ bzw. $f(x) > f(a)$, so nennt man das Extremum isoliert. Ist $U = X$ so nennt man es auch globales Maximum bzw. Minimum.

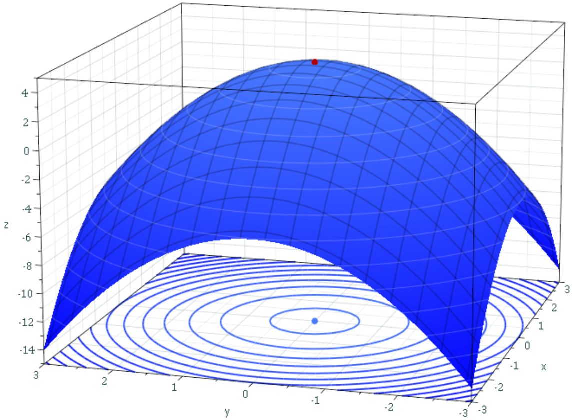


Figure: Quelle: Wikipedia:
https://en.wikipedia.org/wiki/File:MaximumParaboloid.png

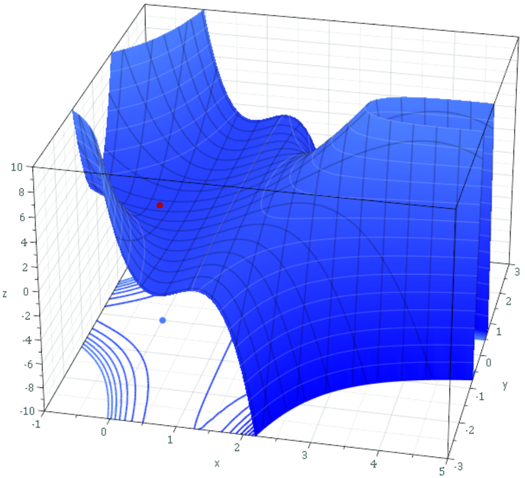


Figure: Quelle: Wikipedia:
https://en.wikipedia.org/wiki/File:MaximumCounterexample.png

Extrema

Ist $f : U \to \mathbb{R}$ differenzierbar und hat f in $a \in U$ ein lokales Extremum, so gilt

$$\frac{\partial}{\partial x_1} f(a) = \cdots = \frac{\partial}{\partial x_n} f(a) = 0 \ .$$

Sind die partiellen Ableitungen stetig, ist dies gleichbedeutend mit $df(a) = 0$.

Kritischer Punkt

Ein Punkt a mit $df(a) = 0$ wird kritischer Punkt genannt.

Beweis

Setze $F_k(t) := f(a + te_k)$. Da f ein Extremum in a hat, hat F_k in einer hinreichend kleinen Umgebung um 0 ein Extremum. Da F_k eine Funktion einer Veränderlichen ist, gilt $F'(0) = 0$. Da $\frac{\partial}{\partial x_k} f(a) = F_k'(0)$ folgt die Behauptung.