Häufungspunkte von Mengen

Sei (X, d^X) metrisch, $A \subseteq X$.

Definition

 $x_0 \in X$ ist Häufungspunkt von A, wenn es eine Folge mit Elementen aus $A \setminus \{x_0\}$ gibt, die nach x_0 konvergiert.

Ein Punkt $x_0 \in A$ der kein Häufungspunkt ist heißt isoliert. Beispiele für $X = \mathbb{R}$:

- Für A = (1, 2) sind genau die Elemente von [1, 2] Häufungspunkte, es gibt keine isolierten Punkte von A.
- $A = \{0\}$ hat keine Häufungspunkte, und 0 ist isoliert. (Die Folge $(0,0,0,\ldots)$ hat Limes und Häufungspunkt 0, aber wir bräuchten ja eine Folge aus $A \setminus \{0\} = \emptyset$.)
- $A = \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$ hat Häufungspunkt 0 (wie die entsprechende Folge), alle Punkte von A sind isoliert.

< A → 102

Grenzwerte

Sei (X, d^X) und (Y, d^Y) metrisch, $A \subseteq X$, und $f : A \to Y$; und sei $x_0 \in X$ Häufungspunkt von A.

Definition

Wir schreiben $\lim_{x\to x_0} f(x) = c$, wenn gilt: Wann immer a_n eine Folge in $A \setminus \{x_0\}$ ist, mit $\lim_{n\to\infty} a_n = x_0$, dann $\lim_{n\to\infty} f(a_n) = c$.

Es gilt:

- Der Limes ist eindeutig (wenn er existiert). Beweis: Weil x_0 Häufungspunkt ist gibt es irgendeine Folge a_n in $A \setminus \{x_0\}$ mit $\lim(a_n) = x_0$. Wenn $\lim_{x \to x_0} f(x)$ existiert, gilt nach Definition $\lim_{x \to x_0} f(x) = \lim_{n \to \infty} f(a_n)$.
- Wenn x_0 zusätzlich in A ist, dann ist f stetig bei x_0 gdw $\lim_{x\to x_0} f(x) = f(x_0)$. Beweis: Übung.



Uneigentliche Limiten für reelle Funktionen

Für $f:A\to\mathbb{R}$ mit $A\subseteq\mathbb{R}$ definieren wir auch die folgenden "uneigentlichen Limiten":

Definition

- $\lim_{x\to x_0}(f(x))=c$ für $c=\pm\infty$ heißt: Wenn $\lim_{n\to\infty}(a_n)=x_0$ mit $a_n\in A$, dann $\lim_{n\to\infty}(f(a_n))=c$.
- Angenommen A ist nicht nach oben beschränkt. Dann schreiben wir (für $c \in \mathbb{R}$ oder $c = \pm \infty$) $\lim_{x \to \infty} f(x) = c$, wenn gilt: Wenn $\lim_{n \to \infty} (a_n) = \infty$ mit $a_n \in A$, dann $\lim_{n \to \infty} f(a_n) = c$.
- Analog für $\lim_{x\to-\infty}$, falls A nicht nach unten beschränkt ist.

Beispiele:

- $\lim_{x\to\infty} e^x = \infty$, $\lim_{x\to 0} \ln(x) = -\infty$, $\lim_{x\to 0} \frac{1}{x^2} = \infty$.
- Es gibt keinen (nicht einmal einen uneigentlichen) Limes $\lim_{x\to 0} \frac{1}{x}$.
- Aber für $g = \frac{1}{x} \upharpoonright \mathbb{R}^{>0}$ ist $\lim_{x \to 0} g(x) = \infty$ ("rechtsseitiger Limes"), analog $\lim_{x \to 0} (\frac{1}{x} \upharpoonright \mathbb{R}^{<0}) = -\infty$ ("linksseitiger").

Ein klassisches Beispiel

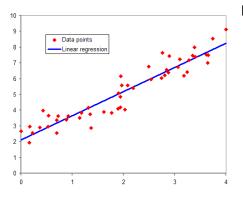
- Sei $f(x) = \frac{\sin(x)}{x}$. Der natürliche Definitionsbereich ist $\mathbb{R} \setminus \{0\}$.
- Man kann zeigen: $\lim_{x\to 0} f(x) = 1$. (Beweis später.) Gut zu wissen: Für kleine x gilt $\sin(x) \sim x$ (und, weniger wichtig, $\sim \tan(x)$).



• Wenn man also die Definition erweitert auf f(0) := 1, dann erhält man ein stetiges $f: \mathbb{R} \to \mathbb{R}$.



Messdaten



Ist diese Funktion "stetig"?

- Messdaten sind diskret (endlich viele Punkte).
 Sinnlos nach Stetigkeit zu fragen.
- Wir interpretieren die Daten in einem Modell.

Hier: lineare Regression
Viele (nicht alle) dieser Modelle
operieren mit stetigen
Funktionen.

 Die meisten Funktionen denen Sie je begegnen werden sind "stückweise stetig": f: A → R mit A eine Vereinigung I₁ ∪ · · · ∪ I_n von Intervallen, so dass f ↑ I_n stetig ist für jedes I_n.
 Bsp: Vorzeichenfunktion, Sägezahn- und Rechteckschwingung, . . .

Differenzialrechnung



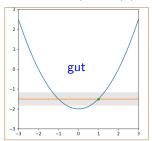
Ableitung als affine Approximation

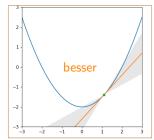
- Stetig heißt: f(x) ist nahe x_0 gut durch eine konstante Funktion g(x) = b approximierbar. (Dann muss $b = f(x_0)$ sein.) Schlampig: $f(x) \sim f(x_0)$.
- Gut heißt: $(\forall \varepsilon > 0) (\exists \delta > 0) (\forall x) d(x, x_0) < \delta \rightarrow d(f(x), g(x)) < \varepsilon$.
- Differenzierbar: f(x) ist nahe bei x_0 "noch besser" durch eine affine Funktion $g(x) = a(x x_0) + b$ approximierbar.

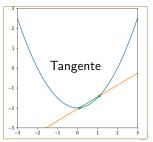
(Dabei wird a $f'(x_0)$ genannt, und b muss wieder $f(x_0)$ sein).

Schlampig: $f(x_0 + \Delta) \sim f(x_0) + f'(x_0) \cdot \Delta$.

• Besser: $(\forall \varepsilon > 0) (\exists \delta > 0) (\forall x) d(x, x_0) < \delta \rightarrow d(f(x), g(x)) \leq \varepsilon \cdot d(x, x_0)$.







Definition der Ableitung für reelle Funktionen

Äquivalent zur vorigen Definition ist die "Steigung der Tangente" (d.h.: Limes der Steigungen der Sekanten):

Um notationelle Unschönheiten zu vermeiden, sei $I \subseteq \mathbb{R}$ ein Intervall das nicht nur aus einem Punkt besteht, oder eine Vereinigung solcher Intervalle.

Definition

- $f: I \to \mathbb{R}$, $x_0 \in I$. Dann ist f in x_0 differenzierbar, wenn $\lim_{x \to x_0} \frac{f(x) f(x_0)}{x x_0}$ existiert. Den Limes nennt man $f'(x_0)$ oder $\frac{df}{dx}(x_0)$.
- Dadurch wird eine neue Funktion $f': A \to \mathbb{R}$ definiert, wobei $A \subseteq I$ die Menge der Punkte ist an denen f differenzierbar ist.
- $f: I \to \mathbb{R}$ heißt differenzierbar, wenn f für alle $x_0 \in I$ differenzierbar ist. In dem Fall ist $f': I \to \mathbb{R}$.



Äquivalenz

Differenzierbar ist äquivalent zu "besser affin approximierbar":

Theorem

• Angenommen $f'(x_0)$ existiert. Setze $g(x) = f(x_0) + (x - x_0)f'(x_0)$. Dann gilt:

(*)
$$(\forall \varepsilon > 0) (\exists \delta > 0) (\forall x \in I) |x - x_0| < \delta \rightarrow |g(x) - f(x)| \le |x - x_0| \cdot \varepsilon.$$

② Wenn es ein affines $g = a \cdot x + b$ gibt das (*) erfüllt, dann existiert $f'(x_0)$ und ist gleich a.

$$\lim_{y \to x_0} \frac{g(y) - f(y)}{y - x_0} = \lim_{y \to x_0} \frac{f(x_0) - f(y)}{y - x_0} + f'(x_0) = -f'(x_0) + f'(x_0) = 0$$

D.h., wenn
$$\lim_{n\to\infty} y_n = x_0$$
, dann $\lim_{n\to\infty} \frac{g(y_n) - f(y_n)}{y_n - x_0} = 0$.

Angenommen (*) gilt nicht. Dann gibt es $\varepsilon > 0$ so dass es für alle $\delta > 0$ ein Gegenbeispiel x_{δ} gibt. Sei $y_n = x_{\underline{1}}$.

D.h.
$$|y_n - x_0| < \frac{1}{n}$$
 und $|g(y_n) - f(y_n)| > |y_n - x_0| \cdot \varepsilon$. Widerspruch.

Beweis (Forts.)

(2) Wir setzen voraus: Es gibt ein affines $g = a \cdot x + b$ so dass $(\forall \varepsilon > 0) (\exists \delta > 0) (\forall x \in I) |x - x_0| < \delta \rightarrow |g(x) - f(x)| \le |x - x_0| \cdot \varepsilon$, und wollen zeigen dass $f'(x_0) = a$.

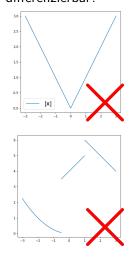
Es gilt dann:
$$|g(x_0) - f(x_0)| = 0$$
, d.h. $g(x_0) = ax_0 + b = f(x_0)$, d.h. $b = f(x_0) - ax_0$, d.h. $g(x) = a(x - x_0) + f(x_0)$.

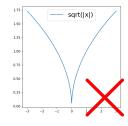
Und es gilt

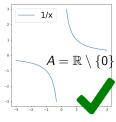
$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{f(x) - g(x) + a(x - x_0)}{x - x_0} = \lim_{x \to x_0} \frac{f(x) - g(x)}{x - x_0} + a = a.$$

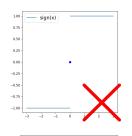
4 🗗 ト

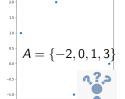
Beispiele: $f: A \to \mathbb{R}$ (mit "natürlichem" A.) Ist f (auf ganz A) differenzierbar?











Ableitungsregeln

Differenzieren ist einfach!

- $(f \pm g)'(x) = f'(x) \pm g'(x)$; und $(a \cdot f)'(x) = a \cdot f'(x)$ für $a \in \mathbb{R}$.
- Produktregel: $(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x)$. Warum? $(f \cdot g)(x+\delta) = f(x+\delta) \cdot g(x+\delta) \sim (f(x)+\delta f'(x)) \cdot (g(x)+\delta g'(x)) = f(x)g(x) + \delta(f'(x)g(x) + f(x)g'(x)) + \delta^2 f'(x)g'(x),$ und wir können $\delta^2 \ll \delta$ vernachlässigen.
- Kettenregel: $(g \circ f)'(x) = (g' \circ f)(x) \cdot f'(x)$. Warum? $(g \circ f)(x + \delta) = g(f(x + \delta)) \sim g(f(x) + \delta f'(x)) \sim g(f(x)) + \delta f'(x)g'(f(x))$.
- Quotientenregel: $(\frac{f}{g})'(x) = \frac{f'(x)g(x) g'(x)f(x)}{g(x)^2}$. Warum? $(f\frac{1}{g})'(x) = f'(x)\frac{1}{g(x)} + f(x)(\frac{1}{g})'(x) = f'(x)\frac{1}{g(x)} + f(x)(-1)\frac{1}{g^2(x)}g'(x)$.
- Wenn g Inverse von f ist, dann $g'(f(x)) = \frac{1}{f'(x)}$; bzw $g'(y) = \frac{1}{f'(g(y))}$. Warum? x = g(f(x)), d.h. $x' = 1 = (g \circ f)'(x) = (g' \circ f)(x) \cdot f'(x)$.

J. Kellner analinf

113

Ableitung von x^r

- $(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k = x^n + n x^{n-1} y + \frac{n(n-1)}{2} x^{n-2} y^2 + \dots + y^n$. Dabei ist $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, genannt Binomialkoeffizient.
- Damit sieht man: $(x + \delta)^n = x^n + nx^{n-1}\delta + \frac{n(n-1)}{2}x^{n-2}\delta^2 + \dots$, wobei wir δ^n vernachlässigen können für $n \ge 2$. D.h., $(x + \delta)^n \sim x^n + nx^{n-1} \cdot \delta$, d.h. $\frac{dx^n}{dx} = nx^{n-1}$
- Man zeit weiters (für n, m in \mathbb{N}) (Übung): $\frac{dx^{\frac{1}{n}}}{dx} = \frac{1}{n}x^{\frac{1}{n}-1}. \text{ Damit: } \frac{dx^{\frac{n}{m}}}{dx} = \frac{n}{m}x^{\frac{n}{m}-1}.$ Und: Wenn $\frac{dx^r}{dx} = rx^{r-1}$ dann $\frac{dx^{-r}}{dx} = -rx^{-r-1}.$ D.h. $\frac{dx^r}{dx} = rx^{r-1}$ für alle $r \in \mathbb{Q}$.
- Beachte die Definitionsbereiche. Bsp: $x^{\frac{1}{3}}$ überall definiert, Ableitung ist $x^{-\frac{2}{3}}$, nur auf $\mathbb{R} \setminus \{0\}$ definiert.
- Für allgemeine $r \in \mathbb{R}$ ist x^r durch rationale Approximationen definiert, damit zeigt man: $\frac{dx^r}{dx} = rx^{r-1} \text{ für } r \in \mathbb{R} \text{ und } x > 0.$

114

Mehr elementare Funktionen

- Für ein Polynom $f(x) = a_n x^n + \cdots + a_1 x + a_0$ ist $f'(x) = na_n x^{n-1} + (n-1)a_{n-1} x^{n-2} \cdots + a_1$ (für alle $x \in \mathbb{R}$).
- Wir definieren $e:=\lim_{n\to\infty}(1+\frac{1}{n})^n$. D.h. für große n ist $e\sim (1+\frac{1}{n})^n$, d.h. $e^{\frac{1}{n}}\sim 1+\frac{1}{n}$. In anderen Worten: $e^{\delta}\sim 1+\delta$ (für kleine δ).

Damit bekommen wir $e^{x+\delta}=e^x\cdot e^\delta\sim e^x(1+\delta)=e^x+e^x\cdot \delta$, d.h.: $(e^x)'=e^x$.

(Ein wesentlicher Grund warum e von so zentraler Bedeutung ist.)

- Daher: $(a^x)' = (e^{x \ln(a)})' = \ln(a)e^{x \ln(a)} = \ln(a) \cdot a^x$ (für a > 0).
- Daher auch: $\ln'(x) = \frac{1}{e^{\ln x}} = \frac{1}{x}$ (für x > 0).
- $\sin'(x) = \cos(x)$, und $\cos'(x) = -\sin(x)$ (mehr dazu später).



Bedeutung der Ableitung

• f' hat enorme konzeptuelle Bedeutung als Änderungsrate von f.

```
Beispiel: Sei f(t) die Position in Abhängigkeit von Zeit t. Dann ist v(t) := f'(t) die Geschwindigkeit. Analog ist v'(t) = f''(t) die Beschleunigung (Änderungsrate von v). Die meisten fundamentalen Gleichungen der Physik sind Differenzialgleichungen (zB Schrödinger- und Maxwell- und Einsteinsche Feldgleichungen)
```

• f' hat enorme Bedeutung als (erste, einfachste) Approximation: $f(x + \delta) \sim f(x) + \delta \cdot f'(x)$.

Viele komplizierte Probleme lassen sich in dieser "ersten Näherung" behandeln, mit immer noch nützlichen Resultaten.

 Ableitung ist in der Mathematik enorm nützlich: Klassisches Beispiel: Extremwerte.



Extremwerte

Definition

Sei $f: A \to \mathbb{R}$, $x_0 \in A$. x_0 ist lokales Maximum, wenn $(\exists \varepsilon > 0) (\forall x \in B_{\varepsilon}(x_0) \cap A) f(x) \leq f(x_0).$

Analog für Minimum.

Theorem

Wenn $f:[a,b]\to\mathbb{R}$ differenzierbar ist, und $x_0\in(a,b)$ lokales Maximum oder Minimum ist, dann $f'(x_0) = 0$.

Beweis (für Minimum): $f(x) \ge f(x_0)$, d.h. $\frac{f(x)-f(x_0)}{x-x_0}$ ist ≥ 0 für $x > x_0$ und ist ≤ 0 für $x < x_0$. Für jede Folge y_n die von rechts kommend gegen x_0 konvergiert gilt also $f'(x_0) = \lim_{n \to \infty} \frac{f(y_n) - f(x_0)}{y_n - y_0} \ge 0$; von links kommend bekommen wir $f'(x_0) < 0$.

< A →

Bemerkungen

- Das gilt nicht für x = a oder x = b: Sei f: [0,1] → ℝ mit f(x) = x.
 Dann ist 0 das (sogar globale) Minimum und 1 das Maximum, aber die Ableitung ist überall 1.
- Die Umkehrung gilt nicht: f'(x) = 0 impliziert nicht dass x lokales Minimum oder Maximum ist. Bsp: x^3 ist streng monoton (und hat daher gar keine lokalen Minima oder Maxima), $\frac{dx^3}{dx}(0) = 0$.
- Immer wieder nützlich, zB für Ungleichungen: Will man zeigen f(x) < g(x) für alle $x \ge 0$, d.h. h(x) := g(x) f(x) > 0, dann reicht es zu zeigen: h(0) > 0, und h(x) > 0 für unbeschränkt viele x, und h(x) > 0 für alle x mit h'(x) = 0.



Mittelwertsatz

Theorem

Sei $f:[a,b] \to \mathbb{R}$ differenzierbar. Dann gibt es ein $x \in (a,b)$ mit $f'(x) = \frac{f(b) - f(a)}{b - a}$.

Beweis: $F(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$. Es gilt F(a) = F(b) = f(a) und $F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}$. Wir suchen ein x mit F'(x) = 0, dann ist $f'(x) = \frac{f(b) - f(a)}{b - a}$ wie verlangt.

Weil F stetig ist, ist F''[a, b] = [c, d] mit $c \le d$ in \mathbb{R} .

Fall 1: c = d. Dann ist F konstant, und F'(x) = 0 für jedes $x \in [a, b]$.

Fall 2: c < d. Dann ist entweder c oder d ungleich f(a) = F(a) = F(b).

Angenommen $d \neq f(a)$. Dann ist das Maximum d = F(y) für ein $y \in (a, b)$, und 0 = F'(y) = 0. Analog für $c \neq f(a)$ und Minimum.

