Mittelwertsatz

Theorem (Mittelwertsatz)

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

Sei f ausreichend differenzierbar. Dann wissen wir:
Lokal (d.h. für $\delta \sim 0$) gilt: $f(x + \delta) \sim f(x) + \delta f'(x)$.

Global ($\delta \gg 0$) und exakt gilt:

$$f(x + \delta) = f(x) + \delta f'(x') \text{ für ein } x' < \delta.$$

(Verwende den Mittelwertsatz für $a = x$, $b = x + \delta$, $x = x'$.)

(Später werden wir etwas ähnliches bei Taylor-Reihen sehen.)

Anwendung des Mittelwertsatzes: Monotonie

Theorem

Sei $f : (a, b) \to \mathbb{R}$ differenzierbar.

- *f ist monoton wachsend gdw $f'(x) \geq 0$ für alle $x \in (a, b)$.*
- *Wenn sogar $f'(x) > 0$, dann ist f streng monoton wachsend.*
- *Analog für fallend.*

Beweis: Für $x < y$ in (a, b) gibt es $z \in (x, y)$ mit $f'(z) = \frac{f(y)-f(x)}{y-x}$.
Angenommen f ist nicht streng monoton wachsend, dann gibt es $x < y$ mit $f(x) \geq f(y)$; daher ist $f'(z) \leq 0$.
Analog: Wenn f nicht monoton wachsend ist, dann $f(x) > f(y)$ und $f'(z) < 0$.
Wenn aber $f'(z) < 0$, dann gibt es Umgebung in der $f(x) > f(y)$ für all $x < z < y$. $\square$

Die Umkehrung gilt nicht für streng monoton.
Bsp: x^3 ist streng monoton, aber $\frac{dx^3}{dx}(0) = 0$.

Anwendung des Mittelwertsatzes: De l'Hopital

Theorem (Ohne Beweis)

Wir setzen voraus: $x_0 \in \mathbb{R} \cup \{\infty, -\infty\}$; $\lim_{x \to x_0} \frac{f'(x)}{g'(x)}$ *existiert; und* $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x)$ *ist entweder* 0 *oder* $+\infty$ *oder* $-\infty$.
Dann gilt:

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

Die Voraussetzung $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) \in \{0, \infty, -\infty\}$ ist nötig!
Bsp: Für $x_0 = 1$, $f(x) = x$, $g(x) = x^2$ ist $\lim_{x \to x_0} \frac{f(x)}{g(x)} = 1$, aber $\lim_{x \to x_0} \frac{f'(x)}{g'(x)} = \frac{1}{2}$.

"Beweis" für $f(x_0) = g(x_0) = 0$ und f, g stetig diffbar bei x_0:
$\frac{f(x_0+\delta)}{g(x_0+\delta)} \sim \frac{f(x_0)+\delta f'(x_0)}{g(x_0)+\delta g'(x_0)} = \frac{f'(x_0)}{g'(x_0)}$.

Bsp: $\lim_{x \to 0} \frac{\sin(x)}{x} = \lim_{x \to 0} \frac{\cos(x)}{1} = 1$.

Exkurs: Differentation und echte Daten

- Symbolisches Differenzieren ist leicht.
- $f'(x)$ für $x \sim x_0$ gut approximierbar durch

$$(*) \quad \frac{f(x) - f(x_0)}{x - x_0}$$

- $f(x)$ in der Praxis oft Messdaten: diskret, verrauscht.
- Direkt $(*)$ zu verwenden wäre sinnlos, konvergiert zu keinem sinnvollen Wert wenn $x \to x_0$.
- Wie üblich: Auf Grundlage eines Modells (oder "universell", zB mit splines) Funktion fitten; dann diese Funktion ableiten.

Mehrere Variablen: Partielle Differentiale

- Sei $g : \mathbb{R}^n \to \mathbb{R}$ und $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ und $1 \leq i \leq n$.
- Wir können den Wert für alle x_j mit $j \neq i$ festhalten (d.h., als Konstanten behandeln), und erhalten eine Funktion $h : \mathbb{R} \to \mathbb{R}$ mit $h(y) = g(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)$.
- Wir bezeichnen $h'(x_i)$ als $\frac{\partial g}{\partial x_i}(\vec{x})$.
- Beispiel: Sei $f(x_1, x_2) = \sin(x_1 x_2) + (x_2)^2$ und $\vec{x} = (0, 1)$.
  $\frac{\partial f}{\partial x_1} = x_2 \cos(x_1 x_2), \quad \frac{\partial f}{\partial x_1}\vec{x} = 1 \cdot \cos(0 \cdot 1) = 1$
  $\frac{\partial f}{\partial x_2} = x_1 \cos(x_1 x_2) + 2x_2, \quad \frac{\partial f}{\partial x_2}\vec{x} = 0 \cdot \cos(0 \cdot 1) + 2 \cdot 1 = 2.$

Mehrere Variablen: Die Ableitung

Betrachten wir $f : \mathbb{R}^n \to \mathbb{R}^m$, $f(\vec{x}) = (f_1(\vec{x}) \dots, f_m(\vec{x}))$.

- Eine lineare Funktion $\mathbb{R}^n \to \mathbb{R}^m$ entspricht einer $m \times n$ Matrix A:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} \qquad z_i = \sum_{j=1}^n a_{ij} x_j$$

- Die Ableitung von f bei $\vec{x}$ ist eine affine Approximation $g(\vec{y}) = f(\vec{x}) + A \cdot (\vec{y} - \vec{x})$, wobei A so eine Matrix ist.
- Damit die Ableitung existiert, muss die Ableitung "besser" sein entsprechend der Definition:

$$(\forall \varepsilon > 0)\, (\exists \delta > 0)\, d(\vec{y}, \vec{x}) < \delta \to d(f(\vec{y}), g(\vec{y})) < \varepsilon \cdot d(\vec{y}, \vec{x})$$

- (Ohne Beweis!) Wenn die Ableitung von f existiert, dann gilt : $a_{ij} = \frac{\partial f_i}{\partial x_j}(\vec{x})$.

Mehrere Variablen: Beispiel

- Dasselbe Beispiel wie zuvor: $f(x_1, x_2) = \sin(x_1 x_2) + (x_2)^2$ und $\vec{x} = (0, 1)$, d.h.: $\frac{\partial f}{\partial x_1}\vec{x} = 1$; $\frac{\partial f}{\partial x_2} = 2$ und $f(\vec{x}) = 1$.
  Für den Punkt $\vec{x}$ ist die affine Approximation $g(\vec{y})$ also:

$$g(\vec{y}) = f(\vec{x}) + A(\vec{y} - \vec{x}) = 1 + \begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} y_1 - 0 \\ y_2 - 1 \end{pmatrix} = y_1 + 2y_2 - 1$$

- Hier (d.h. im Fall $\mathbb{R}^2 \to \mathbb{R}$) entspricht die Ableitung einer "Tangentialebene".
  (Vgl: affine Approximation im 1-dim entspricht Gerade/Tangente.)
  Dabei gibt es viele Aspekte die im Eindimensionalen nicht auftreten.
  Es gibt z.B. (für nicht-konstante Ebenen) eine Richtung in die Ebene/Funktion am schnellsten steigt ("Gradient") etc.

Mehrere Variablen: Bemerkung

Wenn $f : \mathbb{R}^n \to \mathbb{R}^m$ an $\vec{x}$ differenzierbar ist, dann können wir also schreiben:

$$f(\vec{y}) \sim g(\vec{y}) = f(\vec{x}) + A(\vec{x}) \cdot (\vec{y} - \vec{x})$$

mit einer linearen Funktion/Matrix $A(\vec{x})$.
(Wir schreiben jetzt $A(\vec{x})$ um zu verdeutlichen dass die Matrix A von $\vec{x}$ abhängt, genauso wie $f'(x)$ i.A. von x abhängt.)

Sei A die Menge der $\vec{x} \in \mathbb{R}^n$ an denen f differenzierbar ist. Die Ableitung "als Funktion" Df mit Definitionsbereich A bildet nun also ein $\vec{x} \in A \subseteq \mathbb{R}^n$ auf eine lineare Abbildung $A(\vec{x})$ ab, d.h. auf ein Element von $\mathbb{R}^{m \times n}$!

Das deutet an dass bereits die zweite Ableitung im mehrdimensionalen komplizierter ausfallen wird als im Eindimensionalen (Krümmungstensor etc).

Höhere Ableitungen: Notation

- Sei $f : A_0 \to \mathbb{R}$, und $A_1 \subseteq A_0$ die Menge der x an denen f differenzierbar ist.
  Wir bekommen dann $f' : A_1 \to \mathbb{R}$.
  Sei $A_2 \subseteq A_1$ die Menge der x an denen f' differenzierbar ist.
  Das definiert $(f')' : A_2 \to \mathbb{R}$.
- Statt $(f')'$ schreibt man f'' oder $\frac{d^2 f}{dx^2}$, genannt "zweite Ableitung", und für $x \in A_2$ sagt man "f ist zweifach differenzierbar bei x."
- Analog: Statt $(f'')'$ schreibt man f''', genannt "dritte Ableitung", die k-te Ableitung bezeichnet man mit $f^{(k)}$. Wir setzen auch $f^{(0)} := f$.
- Wenn f (bei x) k-mal differenzierbar ist für alle $k \in \mathbb{N}$, dann heißt f unendlich oft differenzierbar.

Zweite Ableitung als Krümmung

Wenn f differenzierbar, dann ist $f'(x)$ die Steigung der Tangente in x.
Wenn f' differenzierbar, dann ist $f''(x)$ die Änderung der Steigung, d.h., die Krümmung der Kurve, im Punkt x.

Definition

Eine Funktion $f : (a, b) \to \mathbb{R}$ heißt konkav, wenn für alle $x < y < z$ in (a, b) der Wert $f(z)$ über der Strecke zwischen $(x, f(x))$ und $(y, f(x))$ liegt, d.h., wenn $f(y) \geq f(x) + (y - x)\frac{f(z)-f(x)}{z-x}$.
Wenn sogar $>$ gilt statt $\geq$ heißt f strikt konkav.
Analog für [strikt] konvex, wenn $\leq$ [bzw $<$] gilt.

Theorem (Ohne Beweis)

Sei $f : [a, b] \to \mathbb{R}$ differenzierbar. Dann ist f [strikt] konkav genau dann wenn f' [streng] monoton fallend ist. Analog für konvex und steigend.

Insbesondere: Wenn $f'' > 0$ auf $[a, b]$, dann ist f konvex, etc.

Kurvendiskussion

Ableitung und zweite Ableitung erlaubt es in vielen Fällen, diverse Eigenschaften einer Funktion $f : [a, b] \to \mathbb{R}$ zu bestimmen, z.B.:

- $f'(x) = 0$ ist notwendige Bedingung für "x lokales Extremum,"
- $f' \geq 0$ gdw f monoton steigend,
- $f' > 0$ ist hinreichende Bedingung für "f streng monoton steigend,"
- $f'' > 0$ ist hinreichende Bedingung für "f konvex", etc

Wir erwähnen ein weiteres klassisches hinreichendes Kriterium:

- Wenn $f'(x) = 0$ und $f''(x) < 0$, dann ist x ein lokales Maximum. (Und für $f''(x) > 0$ ein lokales Minimum).

Zweite Ableitung als quadratische Approximationen

- f ist stetig heißt:
  f ist durch konstante Funktion (=Polynom von Grad 0) "gut" approximierbar. (Gut entspricht $\dots < \varepsilon$.)
- f ist differenzierbar heißt:
  f ist durch affine Funktion (=Polynom von Grad 1) "besser" approximierbar. (Besser entspricht $\dots < \varepsilon \cdot |x - x_0|$.)
- Wir wollen nun f durch ein Polynom g von Grad 2, d.h. quadratisch, "noch besser" approximieren, entsprechend $\dots < \varepsilon \cdot |x - x_0|^2$.
  $f(x) \sim g(x) := a(x - x_0)^2 + b(x - x_0) + c$
- Es stellt sich heraus das dafür kein neues Konzept nötig ist, sondern dass wir dafür die zweite Ableitung verwenden können.
- Was ist die natürliche Wahl der Koeffizienten a, b, c? Wir wollen:
  (Wie in konstanter Approx.) $g(x_0) = f(x_0) \Rightarrow c = f(x_0)$
  (Wie in affiner Approx.) $g'(x_0) = f'(x_0) \Rightarrow b = f'(x_0)$
  (Neu) $g''(x_0) = f''(x_0)$. Achtung: $g''(x_0) = 2a$, d.h. $\Rightarrow a = \frac{f''(x_0)}{2}$.

Ableitung von Polynomen

- Noch ein Beispiel: Grad 3.

$$\begin{array}{rlcl}
f(x) = & a\cdot x^3 + b\cdot x^2 + c\cdot x + d & \Rightarrow f(0) = & d = 0!d\\
f'(x) = & a\cdot 3\cdot x^2 + b\cdot 2\cdot x + c & \Rightarrow f'(0) = & c = 1!c\\
f''(x) = & a\cdot 3\cdot 2\cdot x + b\cdot 2 & \Rightarrow f''(0) = & 2b = 2!b\\
f'''(x) = & a\cdot 3\cdot 2 & \Rightarrow f'''(0) = & 2\cdot 3a = 3!a
\end{array}$$

- Allgemein: $f(x) = a_n x^n + a_{n-1}x^n + \dots a_1 x + a_0$.
  Dann ist $f^{(k)}(0) = 1\cdot 2\cdots k\cdot a_k = k!a_k$ (für alle $0 \le k \le n$).
  (Insbes.: $f^{(0)}(0) = f(0) = a_0 = 0!a_0$, Daher gilt: $f^{(1)}(0) = a_1 = 1!a_1$)
  und $f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k$.
- Dementsprechend ist die natürliche n-te Approximation zu einer n-mal differentierbaren Funktion $f(x)$: $g(x) := \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k$.

Taylor-Reihen

Wie gut ist die Approximation? Schreibe $f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k + R_n(x)$.
Die Approximation ist gut wenn $|R_n(x)|$ klein ist.

Theorem (Satz von Taylor, ohne Beweis)

$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k + R_n(x)$, wobei $R_n(x) = \frac{f^{(n+1)}(\zeta)}{(n+1)!} x^{n+1}$ für ein $\zeta \in [0, x]$

Den Fall $n = 0$ kennen wir schon aus dem Mittelwertsatz:
$f(x) = f(0) + x f'(\zeta)$ für ein $\zeta \in [0, x]$.

Definition

Die Taylor-Reihe von f ist $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$.

Ein Beispiel

- Betrachte $f(x) = \sin(x)$.
- Es gilt $|f^{(k)}(x)| \leq 1$ für alle k, x. Daher ist $|R_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!}$
- $n!$ wächst schneller als m^n für jedes fixe $m \in \mathbb{N}$:
  $$\begin{array}{rl|l|l} m^n = & m \cdot m \cdots m & \cdot \quad m \quad \cdots \quad m & \cdot \quad m \quad \cdots m \\ n! = & 1 \cdot 2 \cdots m & \cdot (m+1) \cdots 2m & \cdot (2m+1) \cdots n \end{array}$$
  D.h. $\frac{n!}{m^n} > \frac{1}{m^m} \cdot 2^{n-2m} = 2^{n-2m-m \ln m}$. (Besser: $> (\frac{n}{me})^n$)
- D.h. für jedes x gilt $\lim_{n\to\infty} |R_n(x)| = 0$.
- Für jedes x konvergiert also $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ gegen $\sin(x)$.
- Die Ableitungen von sin sind $\sin, \cos, -\sin, -\cos, \sin, \dots$; für $\sin^{(k)}(0)$ bekommen wir $0, 1, 0, -1, \dots$. D.h.
  $\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} \pm \cdots$

Potenzreihen

Definition

Eine Potenzreihe ist eine unendliche Reihe der Form $g(x) := \sum_{n=0}^{\infty} a_n x^n$.

Genauer: Für jedes fixe x ergibt diese Darstellung eine unendliche Reihe, die konvergieren kann oder auch nicht.
Wenn sie konvergiert, dann ist $g(x)$ definiert, sonst nicht.
$g(0)$ existiert immer, und $g(0) = a_0$.

Theorem (Ohne Beweis)

Es gibt ein $0 \leq R \leq \infty$ (Konvergenzradius) mit:
$|x| < R \Rightarrow g(x)$ *existiert (Reihe konvergiert sogar absolut)*
$|x| > R \Rightarrow g(x)$ *existiert nicht*
(Für $|x| = R$ sagen wir nichts.)

Beispiele

- $a_n = n!$. Dann ist $g(x)$ für kein $x \neq 0$ definiert.
  (Weil ja $n! \gg x^n$ für jedes x, d.h., $n!x^n$ ist keine Nullfolge, daher kann die Reihe nicht konvergieren.)
  Dementsprechend ist $R = 0$.
- $a_n = 1$. Dann haben wir die geometrische Reihe $\sum_{n=0}^{\infty} x^n$, die konvergiert genau dann wenn $|x| < 1$ (und liefert $\frac{1}{1-x}$).
  Dementsprechend ist $R = 1$.
- $a_n = \frac{\sin^{(n)}(0)}{n!}$. Wie wir vorher aus dem Satz von Taylor gesehen haben konvergiert $g(x)$ für alle x, und zwar gegen $\sin(x)$.
  Dementsprechend ist $R = \infty$.

Entwicklung um andere Mittelpunkte

Bis jetzt haben wir Potenzreihen und Taylor-Reihen "um 0" betrachtet. Wir können dasselbe "um x_0" tun.

- Eine Potenzreihe um x_0 hat die Form $\sum_{n=0}^{\infty} a_n(x-x_0)^n$.
  Es gilt: Es gibt $0 \leq R < \infty$ s.d. absolut konvergent für $|x-x_0| < R$ und divergent für $|x-x_0| > R$.
- Die Tylor Reihe von f um x_0 ist $\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$

Taylor-Reihen

- Wenn f bei 0 unendlich oft differenzierbar ist, dann definiert f die Potenzreihe $g(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$, genannt Taylor Reihe.
- Fragen:
  - Was ist der Konvergenzradius R von g?
  - Gilt $g(x) = f(x)$ für alle $|x - x_0| < R$?
- Antwort, für alle anständigen Funktionen f:
  - R ist so groß wie nur irgendwie möglich.
  - Ja, $g(x) = f(x)$
- Was heißt "anständig"? Das beantwortet erst die komplexe Analysis. In der Praxis sind alle unendlich oft differenzierbaren Funktionen die Ihnen je begegnen werden anständig.

Eine unanständige Funktion

(Kein Prüfungsstoff)
Eine unanständige (bei 0) Funktion wäre z.B.

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{wenn } x \neq 0 \\ 0 & \text{sonst} \end{cases}.$$

Wir bekommen $f^{(n)}(0) = 0$ für alle n, daher ist $g(x) = 0$ und $R = \infty$.
Aber $g(x) \neq f(x)$ für alle $x \neq 0$.

Warum kann das passieren? Im reellen sieht es so aus als ob wir $f(x)$ bei 0 stetig (und sogar diffbar) definieren können, im komplexen geht das aber nicht.

Um $x_0 \neq 0$ funktioniert die Taylorreihe bestens, mit $R = |x_0|$.