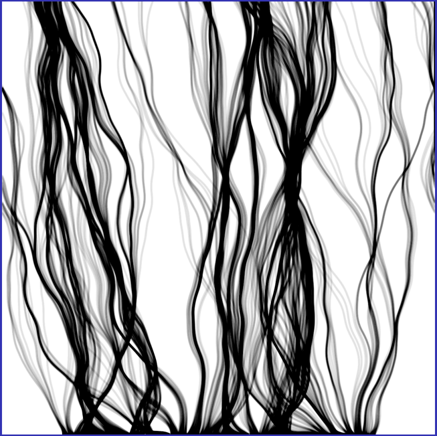
Angewandte Mathematik



Dr. rer. nat. Johannes Riesterer

Wiederholung Taylorreihe für 1-dimensionale Funktionen

Für eine $\mathcal{C}^n$-Funktion $f : \mathbb{R} \to \mathbb{R}$ und $a \in \mathbb{R}$ ist die Taylorreihe um a gegeben durch

$$T_N f(x; a) := \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

Taylorreihe Wiki

Wiederholung Taylorreihe für 1-dimensionale Funktionen

Für eine $\mathcal{C}^{n+1}$-Funktion $f : \mathbb{R} \to \mathbb{R}$ und $a \in \mathbb{R}$ gilt

$$f(x) = T_n f(x; a) + R_n f(x; a) = T_n f(x; a) + o(|x-a|^n), \quad x \to a$$

Taylorreihe Wiki

Wiederholung Taylorreihe für 1-dimensionale Funktionen

Wie kann man das auf höhere Dimensionen verallgemeinern und was für Eigenschaften brauchen wir dafür?

Satz von Schwarz

Wenn Für eine Funktion $f : U \subset \mathbb{R}^n \to \mathbb{R}$ die Ableitungen $\frac{\partial}{\partial x_i} f(a)$, $\frac{\partial}{\partial x_j} f(a)$ und $\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(a)$ existieren und letztere stetig ist, dann existiert auch $\frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f(a)$ und es gilt

$$\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(a) = \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f(a)$$

Bedeutung

Die Reihenfolge spielt beim wiederholten ableiten keine Rolle.

$\mathcal{C}^k$-Funktionen

Eine Funktion $f : U \subset \mathbb{R}^n \to \mathbb{R}$ für die alle partiellen Ableitungen

$$\frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_k}} f(a)$$

mit $i_1 + \cdots + i_k \leq k$ existieren und stetig sind heißt $\mathcal{C}^k$-Funktion oder k-mal stetig differenzierbar.

$\mathcal{C}^k$-Funktionen

Eine $\mathcal{C}^1$-Funktion ist also eine differenzierbare Funktion.

p-te Ableitung

Für eine Funktion $f : U \to \mathbb{R}$, $a \in U$ und Vektoren $v^1, \cdots, v^p \in \mathbb{R}^n$ heißt

$$d^p f(a)\big(v^1, \cdots, v^p\big) := \partial_{v^1} \ldots \partial_{v^p} f(a)$$

die p-te Richtungsableitung von f. Sie ist wegen dem Satz von Schwarz wohldefiniert.

$$d^p f(a)\big(v^1, \cdots, v^p\big) = \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} \frac{\partial}{\partial x_{i_1}} \ldots \frac{\partial}{\partial x_{i_p}} f(a) \cdot v^1_{i_1} \cdots v^p_{i_p} \; .$$

Für einen Vektor $z \in \mathbb{R}^n$ definieren wir $d^p f(a) z^p := d^p f(a) \underbrace{(z, \cdots, z)}_{p-mal}$.

Hessematrix

Für $p = 2$ und $u, v \in \mathbb{R}^n$ ist

$$d^2 f(a)(u, v) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f(a) v_i u_i$$

und mit

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

ist $d^2 f(a)(u, v) = u^T \cdot f''(a) \cdot v$. Die Matrix $f''(a)$ wird auch Hesse-Matrix genannt. Nach dem Satz von Schwarz ist sie symmetrisch.

Taylorapproximation

Sei $f : U \subset \mathbb{R}^n \to \mathbb{R}$ eine $\mathcal{C}^{p+1}$-Funktion und $x, a \in U$, so dass die Verbindung zwischen x und a in U liegt. Dann gilt

$$f(x) = f(a) + \sum_{k=1}^{p} \frac{1}{p!} d^k f(a)(x-a)^k + R_{p+1}(x; a)$$

mit dem Restglied $R_{p+1}(x; a) = o(||x-a||^n)$.

Beweis

Sei $F(t) := f(a + th)$ mit $t \in [0, 1]$. Wiederholte Anwendung der Kettenregel mit $\gamma(t) := a + th$ ergibt

$$F'(t) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(a + th) h_i$$

$$F''(t) = \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} f(a + th) h_i h_j$$

$$\vdots$$

$$F^p(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} \frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_p}} f(a + th) h_{i_1} \cdots h_{i_p} .$$

Beweis

Für $h := (x - a)$ ist $F(0) = f(a)$ und $F(1) = f(x)$ und mit der Taylorapproximation für Funktionen einer Veränderlichen

$$F(1) = F(0) + F'(0) + \frac{1}{2!}F''(0) + \cdots + \frac{1}{p!}F^p(0) + R_{p+1}$$

Extrema

Ist $f : U \to \mathbb{R}$ eine $\mathcal{C}^2$-Funktion und ist $f'(a) = 0$ für ein $a \in U$. Dann gilt:

- $f''(a) > 0 \Rightarrow f$ hat in a ein isoliertes lokales Minimum.
- $f''(a) < 0 \Rightarrow f$ hat in a ein isoliertes lokales Maximum.
- $f''(a) \gtrless 0 \Rightarrow f$ hat in a einen Sattelpunkt.

Positive/negative Definitheit

- $f''(a) > 0 \Leftrightarrow \forall h \in \mathbb{R}^n \; h^t f''(a) h > 0.$
- $f''(a) < 0 \Leftrightarrow \forall h \in \mathbb{R}^n \; h^t f''(a) h < 0.$
- $f''(a) \gtrless 0 \Leftrightarrow h^t f''(a) h \gtrless 0.$

Beweis

Sei $f'(a) = 0$ und $f''(a) > 0$. Mit der Taylorformel gilt für hinreichend kurze Vektoren $h \in \mathbb{R}^n$

$$f(a+h) = f(a) + \frac{1}{2} h^T f''(a) h + R(h)$$

mit $\lim_{h \to 0} \frac{R(h)}{||h||^2} = 0$. Für $||h|| \leq 1$ hat $h^T f''(a) h$ sein Maximum m auf dem Einheitskreis $\{h \in \mathbb{R}^n \mid ||h|| = 1\}$ da $f''(a) > 0$.

$$h^T f''(a) h = ||h|| \frac{1}{||h||} h^t f''(a) ||h|| \frac{1}{||h||} h \geq m ||h||^2 \, .$$

Beweis

Wir wählen ϵ so klein, dass $R(h) \leq \frac{m}{2}||h||^2$ gilt für $||h|| < \epsilon$ (was geht wegen Taylorformel). Damit erhalten wir

$$f(a+h) \geq f(a) + m||h||^2 \ .$$

und damit hat f ein lokales Minimum in a.
Der Fall $f''(a) < 0$ wird mit Betrachtung von $-f$ durch den vorigen Fall bewiesen.

Beweis

Es sei nun $f''(a) \gtrless 0$ und v mit $v^T f''(a)v > 0$ und w mit $w^T f''(a)w > 0$. Betrachten wir die Funktionen

$$F_v(t) := f(a + tv)$$
$$F_w(t) := f(a + tw)$$

dann ist

$$F_v'(t) = 0;\ F_v''(0) = v^T f''(a)v > 0$$
$$F_w'(t) = 0;\ F_w''(0) = w^T f''(a)w < 0$$

und somit hat F_v ein isoliertes lokales Maximum und F_w ein isoliertes lokales Minimum und damit f kein lokales Extremum in a.

- Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix.
- Behauptung: A ist genau dann positiv definit, wenn alle Eigenwerte von A positiv sind.
- Wir beweisen die beiden Richtungen:
  - Positive Definitheit $\Rightarrow$ Positive Eigenwerte
  - Positive Eigenwerte $\Rightarrow$ Positive Definitheit

- Angenommen, A ist positiv definit. Das bedeutet:

$$h^\top A h > 0 \quad \text{für alle } h \in \mathbb{R}^n \setminus \{0\}.$$

- Sei λ ein Eigenwert von A mit zugehörigem Eigenvektor $v \neq 0$, sodass $Av = \lambda v$.
- Setze $h = v$ und erhalte:

$$v^\top A v = v^\top (\lambda v) = \lambda v^\top v.$$

- Da $v^\top v > 0$, folgt $\lambda > 0$.

- Da dies für jeden Eigenwert von A gilt, sind alle Eigenwerte von A positiv.
- Dies zeigt: Wenn A positiv definit ist, sind alle Eigenwerte positiv.

Richtung 2: Positive Eigenwerte $\Rightarrow$ Positive Definitheit

- Angenommen, alle Eigenwerte von A sind positiv.
- Da A symmetrisch ist, können wir A diagonalisieren:

$$A = Q\Lambda Q^\top,$$

wobei Λ eine Diagonalmatrix mit positiven Diagonaleinträgen ist.

- Für $h \in \mathbb{R}^n \setminus \{0\}$:

$$h^\top A h = h^\top Q \Lambda Q^\top h = (Q^\top h)^\top \Lambda (Q^\top h).$$

- Setze $y = Q^\top h$. Da Q orthogonal ist, gilt $y \neq 0$ wenn $h \neq 0$.
- Dann ist:

$$h^\top A h = y^\top \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2.$$

- Da $\lambda_i > 0$ für alle i und $y \neq 0$, folgt $\sum_{i=1}^{n} \lambda_i y_i^2 > 0$.
- Somit ist $h^\top A h > 0$ für alle $h \neq 0$.
- Also ist A positiv definit.