Vo 18.11.09

Taylor (1dim)

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \dots$$

$\alpha = (\alpha_1, \dots, \alpha_n)$ $\quad |\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$

$$D^\alpha f(x_0) = D^{(\alpha_1, \dots, \alpha_n)} f(x_0) = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}} f(x_0)$$

$$\alpha ! := \alpha_1 ! \, \alpha_2 ! \dots \alpha_n !$$

$$(x-x_0)^\alpha := (x_1 - x_{10})^{\alpha_1} \dots (x_n - x_{n0})^{\alpha_n}$$

Nov 18-12:31

Satz von Taylor:

Sei $D \subset \mathbb{R}^n$ offen und konvex, $f : D \to \mathbb{R}$ eine C^{m+1}–Funktion und sei $\mathrm{x}_0 \in D$.

Dann gilt für alle $\mathrm{x} \in D$ die folgenden Entwicklung nach Taylor:

$$\boxed{f(\mathrm{x}) \overset{!}{=} T_m(\mathrm{x}; \mathrm{x}_0) + R_m(\mathrm{x}; \mathrm{x}_0)}$$

$$T_m(\mathrm{x}; \mathrm{x}_0) = \sum_{|\alpha| \le m} \frac{D^\alpha f(\mathrm{x}_0)}{\alpha!}(\mathrm{x} - \mathrm{x}_0)^\alpha$$

$$R_m(\mathrm{x}; \mathrm{x}_0) = \sum_{|\alpha| = m+1} \frac{D^\alpha f(\mathrm{x}_0 + \theta(\mathrm{x} - \mathrm{x}_0))}{\alpha!}(\mathrm{x} - \mathrm{x}_0)^\alpha$$

mit einem geeigneten $\theta \in (0, 1)$.

Beachte: Summation über $|\alpha| \le m$ und $|\alpha| = m + 1$.

$$\underbrace{\underbrace{\underbrace{f(x_0)}_{T_0} + \sum_{|\alpha|=1} \frac{D^\alpha f(x_0)}{\alpha!}(x-x_0)}_{T_1} + \sum_{|\alpha|=2}}_{T_2} + R_2$$

48

Nov 18-12:43

Herleitung der Taylorformel:

Wir definieren eine skalare Funktion einer Variablen $t \in [0, 1]$ als

$g : [0,1] \to \mathbb{R}^n$ $\qquad g(t) := f(\mathrm{x}_0 + t(\mathrm{x} - \mathrm{x}_0))$ $\qquad g(0) = f(x_0)$, $g(1) = f(x)$

und berechnen die Taylor–Entwicklung um $t = 0$.

Es gilt: 1dim Taylor

$$f(x) = g(1) = \underbrace{g(0)}_{f(x_0)} + g'(0) \cdot (1 - 0) + \frac{1}{2} g''(\xi) \cdot (1 - 0)^2, \qquad \xi \in (0, 1)$$

Berechnung von $g'(0)$:

$$\begin{aligned} g'(0) &= \frac{d}{dt} f(x_1^0 + t(x_1 - x_1^0), x_2^0 + t(x_2 - x_2^0), \dots, x_n^0 + t(x_n - x_n^0))_{|t=0} \\ &= D_1 f(\mathrm{x}_0)(x_1 - x_1^0) + \dots + D_n f(\mathrm{x}_0)(x_n - x_n^0) \\ &= \sum_{|\alpha|=1} \frac{D^\alpha f(\mathrm{x}_0)}{\alpha!}(\mathrm{x} - \mathrm{x}_0)^\alpha \end{aligned}$$

$$\frac{d}{dt} \underbrace{f(x_0 + t(x - x_0))}_{h(t)} \Big|_{t=0} = \operatorname{grad} f(x_0 + t(x - x_0)) \cdot (x - x_0) \Big|_{t=0} = \operatorname{grad} f(x_0)(x-x$$

49

Nov 18-12:49

$$\frac{1}{2} \frac{d}{dt} \frac{d}{dt} f(x_0 + t(x - x_0)) \Big|_{t=0} =$$

$$= \frac{1}{2} \frac{d}{dt} \left(\operatorname{grad} f(x_0 + t(x - x_0)) \cdot (x - x_0) \right) =$$

$$= \frac{1}{2} \frac{d}{dt} \sum_{i=1}^n \frac{\partial f}{\partial x_i} (\qquad) \cdot (x_i - x_{i0})$$

$$= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (x_j - x_{j0}) \cdot \underbrace{\frac{\partial^2 f}{\partial x_i \partial x_j} (x_0 + t(x - x_0))} (x_i - x_{i0})$$

$$= \frac{1}{2} (x - x_0)^T Hf(x_0 + t(x - x_0)) \cdot (x - x_0) \Big|_{t=0}$$

$$= \frac{1}{2} (x - x_0)^T Hf(x_0)(x - x_0)$$

Nov 18-12:50

Beispiel:

Berechne das Taylor–Polynom $T_2(\mathrm{x}; \mathrm{x}_0)$ zweiten Grades der Funktion

$$\boxed{f(x, y, z) = xy^2 \sin z}$$

zum Entwicklungspunkt $(x, y, z) = (1, 2, 0)^T$.

Zur Berechnung von $T_2(\mathrm{x}; \mathrm{x}_0)$ benötigen wir die partiellen Ableitungen bis zur zweiten Ordnung.

Diese Ableitungen müssen am Punkt $(x, y, z) = (1, 2, 0)^T$ ausgewertet werden.

Als Ergebnis erhält man $T_2(\mathrm{x}; \mathrm{x}_0)$ in der Form

$$T_2(\mathrm{x}; \mathrm{x}_0) = 4z(x + y - 2)$$

Berechnung auf Folie

$$T_2(x, x_0) = \underbrace{\underbrace{0}_{f(x_0)} + 4z}_{T_1} + \underbrace{4(x-1)z}_{\frac{\partial^2}{\partial x \partial z}} + \underbrace{4(y-2)z}_{\frac{\partial^2}{\partial y \partial z}} = 4z(x+y-2)$$

51

Nov 18-13:00

Beispiel:

Man berechne das Taylor–Polynom $T_3(\mathrm{x}; \mathrm{x}_0)$ dritten Grades der Funktion

$$\boxed{f(x, y) = e^y \cos x}$$

zum Entwicklungspunkt $\mathrm{x}_0 = (0, 0)^T$:

1) unter Verwendung des Taylorschen Satzes,

2) unter Verwendung bekannter Reihenentwicklungen.

Berechnung auf Folie

$$f(x, y) = \left(1 + y + \frac{y^2}{2} + \frac{y^3}{6}\right)\left(1 - \frac{x^2}{2}\right) =$$

$$= \underbrace{\underbrace{\underbrace{\underbrace{1}_{T_0} + y}_{T_1} + \frac{y^2}{2} - \frac{x^2}{2}}_{T_2} + \frac{y^3}{6} - \frac{yx^2}{2}}_{T_3}$$

52

Nov 18-13:03

Bemerkung: (Fortsetzung)

2) Man nennt die Matrix

$$Hf(x_0) := \begin{pmatrix} f_{x_1x_1}(x_0) & \dots & f_{x_1x_n}(x_0) \\ \vdots & & \vdots \\ f_{x_nx_1}(x_0) & \dots & f_{x_nx_n}(x_0) \end{pmatrix}$$

die **Hesse–Matrix** von $f(x)$ im Punkt x_0.

Hesse–Matrix = Jacobi–Matrix des Gradienten ∇f

Die Taylor–Entwicklung einer C^3–Funktion lautet daher

$$f(x) = f(x_0) + \text{grad}\, f(x_0)(x-x_0) + \frac{1}{2}(x-x_0)^T Hf(x_0)(x-x_0) + O(\|x-x_0\|^3)$$

Die Hesse–Matrix einer C^2–Funktion ist **symmetrisch**.

$n=2$

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

54

Nov 18-13:08

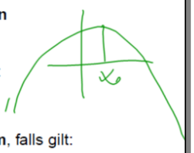
Kapitel 2: Anwendung der Differentialrechnung mehrerer Variablen

2.1 Extrema von Funktionen mehrerer Variablen

Definition: Sei $f: D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, $x^0 \in D$.

1) $f(x)$ hat in x^0 ein **globales Maximum**, falls gilt:

$$\forall x \in D: \quad f(x) \le f(x^0)$$

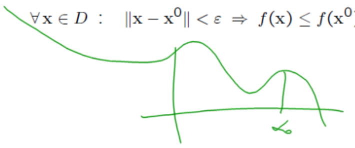


2) $f(x)$ hat in x^0 ein **strenges globales Maximum**, falls gilt:

$$\forall x \in D \setminus \{x^0\}: \quad f(x) < f(x^0)$$

3) $f(x)$ hat in x^0 ein **lokales Maximum**, falls es ein ε gibt mit:

$$\forall x \in D: \quad \|x - x^0\| < \varepsilon \Rightarrow f(x) \le f(x^0)$$



55

Nov 18-13:12

Satz: (Notwendige Bedingung I)

Besitzt $f(x)$ mit $f \in C^2$ in einem Punkt $x^0 \in D^0$ ein lokales Extremum (Minimum oder Maximum), so gilt $\boxed{\text{grad}\, f(x^0) = (0,\dots,0)^T}$.

Beweis:

Für ein beliebiges $v \in \mathbb{R}^n$, $v \neq 0$ ist die Funktion

$\varphi: \mathbb{R} \to \mathbb{R}$

$$\varphi(t) := f(x^0 + tv)$$

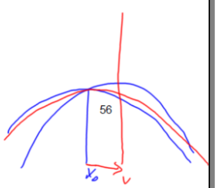
in einer Umgebung von $t^0 = 0$ stetig differenzierbar.

Gleichzeitig hat $\varphi(t)$ bei $t^0 = 0$ ein lokales Extremum. Damit folgt:

$$\varphi'(0) = \text{grad}\, f(x^0) v = 0$$

Da dies für alle $v \neq 0$ gilt, folgt die Bedingung:

$$\text{grad}\, f(x^0) = (0,\dots,0)^T$$

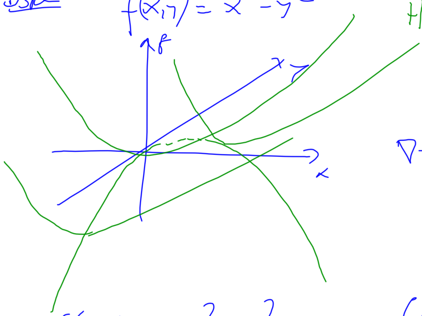


56

Nov 18-13:17

Bsp

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



$$Hf(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$ indef $\Downarrow$ Sattel

$\nabla f(0,0) = 0$

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

$\nabla f(0,0) = 0$

$$Hf(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$ posit. def.

Nov 18-13:18

2.2 Implizit definierte Funktionen

Untersuche die Lösungsmengen von nichtlinearen Gleichungssystemen der Form

$$g(x) = 0$$

mit $g: D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$, d.h. wir betrachten m Gleichungen für n Unbekannte.

Insbesondere gelte:

$$m < n$$

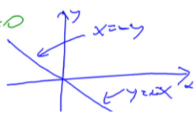
d.h. wir haben **weniger** Gleichungen als Unbekannte.

Man nennt dann das Gleichungssystem **unterbestimmt** und die Lösungsmenge $G \subset \mathbb{R}^n$ enthält gewöhnlich unendlich viele Punkte.

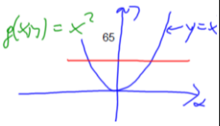
$n=2$, $m=1$

$f(x,y) = y + x = 0$

$\Rightarrow y = -x$



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



65

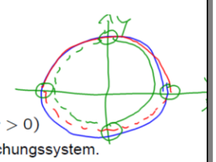
Nov 18-13:44

Beispiel:

$n=2$, $m=1$

Die Kreisgleichung

$$\boxed{g(x,y) = x^2 + y^2 - r^2 = 0} \qquad (r > 0)$$



definiert ein **unterbestimmes** nichtlineares Gleichungssystem.

Wir haben **zwei** Unbekannte (x,y), aber nur eine Gleichung.

Die Kreisgleichung lässt sich **lokal** auflösen und definiert dabei die folgenden vier Funktionen:

$$\begin{aligned} y &= \sqrt{r^2 - x^2}, & -r \le x \le r \\ y &= -\sqrt{r^2 - x^2}, & -r \le x \le r \\ x &= \sqrt{r^2 - y^2}, & -r \le y \le r \\ x &= -\sqrt{r^2 - y^2}, & -r \le y \le r \end{aligned}$$

$\frac{\partial f}{\partial y}(x,y) = 2y = 0 \qquad y = 0$

$\frac{\partial f}{\partial x}(x,y) = 2x = 0 \qquad x = 0$

67

Nov 18-13:51

$g(x,y) = 0$

falls $\underbrace{\text{nach } y \text{ aufl.}}_{y = y(x)}$ $\Rightarrow$ $g(x, y(x)) = 0$

$0 = \frac{d}{dx} g(x,y(x)) = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial y} y'(x)$

$y'(x) = -\frac{\partial g}{\partial x} \cdot \frac{1}{\frac{\partial g}{\partial y}}$ falls $\frac{\partial g}{\partial y} \neq 0$

Nov 18-13:51

$g(x) = Cx + b = (\underbrace{c_1 \dots c_{n-m}}_{B}\ \underbrace{c_{n-m+1}, \dots, c_n}_{A}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + b$

Beispiel:

Sei $g(x)$ eine affin–lineare Funktion, d.h.

$$g(x) = Cx + b, \quad C \in \mathbb{R}^{(m,n)}, \quad b \in \mathbb{R}^m$$

Wir spalten die Variablen x in zwei Vektoren auf

$$x^{(1)} = (x_1, \dots, x_{n-m})^T \in \mathbb{R}^{n-m}$$
$$x^{(2)} = (x_{n-m+1}, \dots, x_n)^T \in \mathbb{R}^n$$

Aufspaltung der Matrix C ergibt die Darstellung

$$g(x) = \boxed{Bx^{(1)}} + \boxed{Ax^{(2)}} + b$$

mit $B \in \mathbb{R}^{(m,n-m)}$, $A \in \mathbb{R}^{(m,m)}$.

A $m \times m$

B $m \times (n-m)$

Das Gleichungssystem $g(x) = 0$ ist genau dann nach den Variablen $x^{(2)}$ (eindeutig) auflösbar, falls A regulär ist:

$$g(x) = 0 \quad \Leftrightarrow \quad \boxed{x^{(2)} = -A^{-1}(Bx^{(1)} + b)}$$

68

Nov 18-13:58

Beispiel: (Fortsetzung)

Wie kann man die Matrix A in Abhängigkeit von g schreiben? Aus der Darstellung

$$g(x) = Bx^{(1)} + Ax^{(2)} + b$$

erkennt man direkt, dass

$$\boxed{A = \frac{\partial g}{\partial x^{(2)}}(x^{(1)}, x^{(2)})} \quad m \times m$$

gilt, d.h. A ist die Jacobi–Matrix der Abbildung $x^{(2)} \to g(x^{(1)}, x^{(2)})$ für festes $x^{(1)}$!

Auflösbarkeit ist also möglich, falls die Jacobi–Matrix regulär ist.

Dies führt auf den

Satz über implizite Funktionen

69

Nov 18-14:00

$f(x) = f(x_0) + \underbrace{\cancel{\text{grad} f(x_0)}}_{=0} (x-x_0) + \frac{1}{2}(x-x_0)^T Hf(x_0)(x-x_0)$

$x = x_0 + \varepsilon v$

Beweis: (zu Teil 1))

Sei x^0 ein lokales Minimum. Für $v \neq 0$ und $\varepsilon > 0$ hinreichend klein folgt aus der Taylor–Formel

(1) $f(x^0 + \varepsilon v) - f(x^0) = \underbrace{\frac{1}{2}(\varepsilon v)^T H f(x^0 + \theta\varepsilon v)(\varepsilon v)}_{R_2} \geq 0$

mit $\theta = \theta(\varepsilon, v) \in (0,1)$.

Der Gradient in der Taylorentwicklung entfällt, da grad $f(x^0) = (0, \dots, 0)^T$ gilt.

Wegen (1) gilt auch

(2) $\quad v^T H f(x^0 + \theta\varepsilon v) v \geq 0$

Da $f(x)$ eine C^2–Funktion ist, ist die Hesse–Matrix eine **stetige** Abbildung. Im Grenzwert $\varepsilon \to 0$ folgt daher aus (2)

$$v^T H f(x^0) v \geq 0$$

d.h. $H f(x^0)$ ist positiv semidefinit.

59

Nov 18-13:28

Bemerkung:

1) Ein stationärer Punkt x^0 mit det $Hf(x^0) = 0$ heißt **ausgeartet**.
Die Hesse–Matrix besitzt dann einen Eigenwert $\lambda = 0$.
Ist x^0 **nicht** ausgeartet, so existieren genau drei Fälle:

pos def Alle Eigenwerte $> 0 \Rightarrow x^0$ ist strenges lokales Minimum

neg def Alle Eigenwerte $< 0 \Rightarrow x^0$ ist strenges lokales Maximum

indef. $\exists$ Eigenwerte $\lambda_1 \cdot \lambda_2 < 0 \Rightarrow x^0$ Sattelpunkt

2) Es gelten die folgenden Implikationen:

$$\begin{array}{ccc} x^0 \text{ lokales Minimum} & \Leftarrow & x^0 \text{ strenges lokales Minimum} \\ \Downarrow & & \Uparrow \\ Hf(x^0) \text{ positiv semidefinit} & \Leftarrow & Hf(x^0) \text{ positiv definit} \end{array}$$

62

Nov 18-13:34