Martin Schmidt
Ross Ogilvie

Analysis II

28. April 2021

7. Übung

20. Interpretation von ∇f

Es sei $f : \mathbb{R}^n \to \mathbb{R}$ in $x \in \mathbb{R}^n$ differenzierbar. Dann ist die Zeilenvektor $\nabla f(x)$ eine Darstellungmatrix von $f'(x)$ auf die Standardbasis von $\mathbb{R}^n$. *Zeige*, dass für jeden Vektor $v \in \mathbb{R}^n$ mit $\|v\| = 1$ gilt:

$$- \|\nabla f(x)\| \ \le \ \nabla f(x) \cdot v \ \le \ \|\nabla f(x)\| \,. \tag{$\star$}$$

Zeige zudem, dass für $\nabla f(x) \neq 0$ in $(\star)$ in der linken bzw. rechten Ungleichung Gleichheit gilt, wenn $v = -\frac{\nabla f(x)}{\|\nabla f(x)\|}$ bzw. $v = \frac{\nabla f(x)}{\|\nabla f(x)\|}$ ist. *(2 Punkte)*

[*Hinweis:* Cauchy-Schwarzsche Ungleichung]

Interpretation: $\nabla f(x)$ gibt die Richtung und den Betrag des größten Anstiegs von f.

Lösung. Wegen der Cauchy-Schwarzschen Ungleichung (9.7) und da f in x differenzierbar ist gilt

$$|f'(x) \cdot v| \le \|\nabla f(x)\| \underbrace{\|v\|}_{=1} = \|\nabla f(x)\|$$

und somit

$$-\|\nabla f(x)\| \le f'(x)v \le \|\nabla f(x)\|.$$

Gleichheit gilt in der Cauchy-Schwarzschen Ungleichung, wenn $f'(x)$ und v linear abhängig sind (vgl. lineare Algebra). Da $\|v\| = 1$ gilt dies nur für $e^{\pm} = \pm \frac{f'(x)}{\|f'(x)\|} = \pm \frac{\nabla f(x)}{\|\nabla f(x)\|} \in \mathbb{R}^n$. Denn dann ist

$$\|v^{\pm}\| = \left\| \frac{\nabla f(x)}{\|\nabla f(x)\|} \right\| = \frac{\|\nabla f(x)\|}{\|\nabla f(x)\|} = 1.$$

Somit gilt in der linken Ungleichung Gleichheit, falls $f'(x)v^{\pm} \le 0$ und in der rechten Ungleichung Gleichheit, falls $f'(x)v^{\pm} \ge 0$.

21. Weitere Untersuchung.

Betrachten Sie die Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$ aus Aufgabe 18 (b):

$$f(x, y) = 2x^3 + xy^2 + 5x^2 + y^2.$$

Die kritischen Punkte von f sind $(0, 0), (-5/3, 0), (-1, 2), (-1, -2)$. Hier ist es als eine Graph: https://www.math3d.org/QaCIDWpL. Schreiben wir in dieser Aufgabe $\mathbf{x} = (x, y) \in \mathbb{R}^2$.

Das Taylorpolynom ist im Satz 10.29 definiert. Das Taylorpolynom in $\mathbf{x}_0$ der Ordnung 0 ist einfach das Wert $f(\mathbf{x}_0)$. Das Taylorpolynom in $\mathbf{x}_0$ der Ordnung 1 ist

$$\begin{aligned} T_{1,\mathbf{x}_0}(\mathbf{x}) &:= \frac{1}{0!} f^{(0)}(\mathbf{x}_0) + \frac{1}{1!} f^{(1)}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \\ &= f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \\ &= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \end{aligned}$$

(a) Zeige, dass der Taylorpolynom von unserem f in $\mathbf{x}_0 = (0,1)$ der Ordnung 1 ist $T_{1,(0,1)}(x,y) = -1 + x + 2y$. *(2 Punkte)*

(b) Berechne $T_{1,(-5/3,0)}(x,y)$. *(1 Punkt)*

(c) Was ist die geometrische Interpretation von $T_{1,\mathbf{x}_0}(\mathbf{x})$? Man könnte math3d.org und (a) und (b) zum Hilfe benutzen. *(1 Punkt)*

Das Taylorpolynom in $\mathbf{x}_0$ der Ordnung 2 ist

$$\begin{aligned} T_{2,\mathbf{x}_0}(\mathbf{x}) &:= \frac{1}{0!} f^{(0)}(\mathbf{x}_0) + \frac{1}{1!} f^{(1)}(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) + + \frac{1}{2!} f^{(2)}(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0, \mathbf{x}-\mathbf{x}_0) \\ &= f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) + \frac{1}{2} f''(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0, \mathbf{x}-\mathbf{x}_0). \end{aligned}$$

Nach Definition 10.26 ist die Darstellungmatrix von $f''(x)$ auf die Standardbasis von $\mathbb{R}^n$ die Hessematrix:

$$\operatorname{Hess}(f)(\mathbf{x}_0) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) \right) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial x \partial y}(\mathbf{x}_0) \\ \frac{\partial^2 f}{\partial y \partial x}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial y^2}(\mathbf{x}_0) \end{pmatrix}$$

Also gilt mit Matrizen

$$T_{2,\mathbf{x}_0}(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x}-\mathbf{x}_0) + \frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T \cdot \operatorname{Hess}(f)(\mathbf{x}_0) \cdot (\mathbf{x}-\mathbf{x}_0)$$

(d) Zeige, dass in $\mathbf{x}_0 = (a,b)$ ist $\operatorname{Hess}(f)(\mathbf{x}_0) = \begin{pmatrix} 12a+10 & 2b \\ 2b & 2a+2 \end{pmatrix}$ *(4 Punkte)*

(e) Berechne daher $T_{2,\mathbf{x}_0}(\mathbf{x})$ in $\mathbf{x} = (0,1)$, $\mathbf{x}_0 = (-5/3, 0)$ und $\mathbf{x}_0 = (0,0)$. *(3 Punkte)*

(f) Was kann man über $T_{2,\mathbf{x}_0}(\mathbf{x})$ sagen, wenn $\mathbf{x}_0$ ein kritischer Punkt ist? *(1 Punkt)*

Lösung.

(a) Nach Aufgabe 18(b) ist $\nabla f(x,y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (6x^2 + y^2 + 10x, 2xy + 2y)$. Berechnen wir $f(0,1) = 1$ und $\nabla f(0,1) = (1,2)$. Dann folgt

$$T_{1,(0,1)}(x,y) = 1 + (1,2) \cdot (x-0, y-1) = 1 + 1(x-0) + 2(y-1) = -1 + x + 2y.$$

(b) $(-5/3, 0)$ ist ein kritischer Punkt, d.h. $\nabla f(-5/3,0) = 0$. Deshalb $T_{1,(-5/3,0)}(x,y) = f(-5/3,0) + 0 = 125/27 = 4.63$, eine konstante Funktion.

(c) $T_{1,\mathbf{x}_0}(x)$ ist die Tangentfläche zu f in x_0. Es ist auch die lineare Approximation.

(d)

$$\begin{aligned} \frac{\partial f}{\partial x} &= 6x^2 + y^2 + 10x \\ \frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(6x^2 + y^2 + 10x) = 12x + 10 \\ \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial}{\partial y}(6x^2 + y^2 + 10x) = 2y \end{aligned}$$

$$\frac{\partial f}{\partial y} = 2xy + 2y$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}(2xy + 2y) = 2y$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}(2xy + 2y) = 2x + 2$$

Bemerke, dass $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$. Welcher Satz sagt, dass das so sein muss?

$$\text{Hess}(f)(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial x \partial y}(\mathbf{x}_0) \\ \frac{\partial^2 f}{\partial y \partial x}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial y^2}(\mathbf{x}_0) \end{pmatrix} = \begin{pmatrix} 12a + 10 & 2b \\ 2b & 2a + 2 \end{pmatrix}$$

(e)

$$\begin{aligned} T_{2,(0,1)}(x,y) &= T_{1,(0,1)}(x,y) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \cdot \text{Hess}(f)(\mathbf{x}_2) \cdot (\mathbf{x} - \mathbf{x}_0) \\ &= 1 + x + 2(y-1) + \frac{1}{2}(x, y-1) \cdot \begin{pmatrix} 10 & 2 \\ 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y-1 \end{pmatrix} \\ &= 1 + x + 2(y-1) + \Big(5x + (y-1), x + (y-1)\Big) \cdot \begin{pmatrix} x \\ y-1 \end{pmatrix} \\ &= 1 + x + 2(y-1) + 5x^2 + x(y-1) + x(y-1) + 2(y-1)^2 \\ &= 1 + x + 2(y-1) + 5x^2 + 2x(y-1) + 2(y-1)^2 \end{aligned}$$

$$\begin{aligned} T_{2,(-5/3,0)}(x,y) &= \frac{125}{27} + \frac{1}{2}(x + 5/3, y) \begin{pmatrix} -10 & 0 \\ 0 & -4/3 \end{pmatrix} \begin{pmatrix} x + 5/3 \\ y \end{pmatrix} \\ &= \frac{125}{27} - 5(x+5/3)^2 - \frac{2}{3}y^2. \end{aligned}$$

$$\begin{aligned} T_{2,(0,0)}(x,y) &= 0 + 0 + \frac{1}{2}(x,y) \begin{pmatrix} 10 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= 5x^2 + y^2 \end{aligned}$$

(f) Natürlich kann man viele Sache sagen:
- $\nabla f(\mathbf{x}_0) = 0$.
- der "Ordnung 1" Term des Taylorpolynoms ist null.
- die Tangentfläche ist horizontal.
- der Taylorpolynom der Ordnung 2 ist ein konstant plus quadratisch auf $x - a$ und $y - b$.
- der Taylorpolynom der Ordnung 2 ist quadratisch Fläche mit einem eindeutigen kritischen Punkt in $\mathbf{x}_0$.

22. Zu Satz 10.27: Hinreichendes Kriterium für lokale Extremwerte.

Es sei $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine symmetrische Bilinearform nach Definition 10.20 und Satz 10.22. Wir wollen zeigen, dass es ein $\delta > 0$ gibt, so dass $\beta(x,x) \geq \delta \|x\|^2$ für alle $x \in \mathbb{R}^n$, wenn β *positiv definit* ist. Positiv definit heißt $\beta(x,x) > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.

(a) Warum muss β stetig sein? *(1 Bonuspunkt)*

(b) Setze voraus, dass β positiv definit ist. Betrachten $W = \{\beta(v,v) \mid v \in \mathbb{R}^n \text{ mit } \|v\| = 1\}$. Warum hat diese Menge ein Minimum δ mit $\delta > 0$? *(2 Punkte)*

(c) Zeige dann, dass $\beta(x,x) \geq \delta \|x\|^2$ für alle $x \in \mathbb{R}^n$. *(2 Punkte)*

Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, $f : U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion und $x_0 \in U$ ein kritischer Punkt von f. Wir wissen, dass $f''(x_0)$ eine symmetrische Bilinearform ist.

(d) *Beweise* man: Wenn $f''(x_0)(x,x) > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ gilt, so besitzt f in x_0 ein striktes lokales Minimum. *(2 Punkte)*

Bemerkung 1. Wenn $f''(x_0)(x,x) < 0$ für einen kritischen Punkt x_0, betrachte $g(x) = -f(x)$. Dann ist $g' = -f'$ und $g'' = -f''$. Es folgt dann, dass x_0 auch ein kritischer Punkte von g ist und ein lokales Minimum von g. Deshalb ist es ein striktes lokales Maximum von f.

Bemerkung 2. Diese Verschärfung von Satz 10.27 gilt nicht, wenn f auf einem unendlichdimensionalen normierten Vektorraum definiert ist, siehe Beispiel 10.28.

Lösung.

(a) Nach Lemma 10.21 ist $B(v)(w) := \beta(v,w)$ eine lineare Abbildung zwischen $\mathbb{R}^n$ und $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$, die endlichdimensionale Vektorräume sind. Es folgt wegen Satz 9.57 (alle lineare Abbildungen von $\mathbb{R}^n$ in einen normierten Vektorraum sind stetig), dass B und daher β stetig sind.

(b) Weil β stetig ist und $\{v \mid v \in \mathbb{R}^n \text{ mit } \|v\| = 1\}$ kompakt ist, folgt die Kompaktheit von $W \subset \mathbb{R}$. Daher hat es ein Minimum δ und $\delta = \beta(v_0, v_0) > 0$ für eine $v_0 \in \mathbb{R}^n$.

(c) Sei $x \in \mathbb{R}^n$ beliebig. Wenn $x = 0$, gilt $\beta(0,0) = \delta \cdot 0$. Wenn $x \neq 0$, setze $\hat{x} = x/\|x\|$, so dass $\|\hat{x}\| = 1$. Nach (b) haben wir

$$\beta(x,x) = \beta(\|x\|\hat{x}, \|x\|\hat{x}) = \|x\|\, \beta(\hat{x}, \|x\|\hat{x}) = \|x\|^2\, \beta(\hat{x}, \hat{x}) \geq \|x\|^2\, \delta.$$

ODER man kann (a)-(c) wie so zeigen: Sei β positiv definit. Dann wird durch $\|x\|_\beta := \sqrt{\beta(x,x)}$ mit $x \in X$ eine Norm auf X definiert. Denn da β eine positiv definite Bilinearform ist, ist β ein Skalarprodukt (vgl. Lineare Algebra 1) und induziert daher durch $\sqrt{\beta(\cdot,\cdot)}$ eine Norm. Da im Endlichdimensionalen alle Normen äquivalent sind, folgt sofort, dass

$$\|x\|_\beta \geq c\|x\|$$

für eine Konstante $c > 0$. Also folgt mit $\delta = c^2$ die Behauptung, denn

$$\beta(x,x) = \|x\|_\beta^2 \geq \delta\|x\|^2.$$

(d) Setze $f''(x_0)(x,x) > 0$ für x_0 einen kritischen Punkt und für alle $x \in \mathbb{R}^n \setminus \{0\}$ voraus. Dann wissen wir jetzt, dass es ein $\delta > 0$ gibt mit $f''(x_0)(x,x) = \beta(x,x) \geq \delta\|x\|^2$. Dann erfüllt das die Voraussetzungen von Satz 10.27, und also ist x_0 ein striktes lokales Minimum.

23. Extremwertsuche.

Bestimme alle lokalen Extrema der Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$ aus Aufgabe 18 (b) und 21, und *entscheide* jeweils, ob es sich um ein Maximum oder ein Minimum handelt. *(5 Punkte)*

Lösung. Nach Aufgabe 18(b) und 21 ist $\nabla f(x,y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (6x^2 + y^2 + 10x, 2xy + 2y)$.
Kritische Punkte: $\{(0,0), (-\frac{5}{3}, 0), (-1,2), (-1,-2)\}$
Hesse-Matrix:

$$\operatorname{Hess}(f)(a,b) = \begin{pmatrix} 12a+10 & 2b \\ 2b & 2a+2 \end{pmatrix}$$

Einsetzen der kritischen Punkte ergibt

$$\operatorname{Hess}(f)(0,0) = \begin{pmatrix} 10 & 0 \\ 0 & 2 \end{pmatrix}$$

Dies ist eine Diagonalmatrix mit positiven Einträgen, also positiv definit:

$$\begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} 10 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 10v_1^2 + 2v_2^2 > 0$$

für $(v_1, v_2) \neq 0$. Daher liegt in $(0,0)$ ein lokales Minimum vor.

$$\operatorname{Hess}(f)\left(-\frac{5}{3}, 0\right) = \begin{pmatrix} -10 & 0 \\ 0 & -\frac{4}{3} \end{pmatrix}$$

Dies ist eine Diagonalmatrix mit negativen Einträgen, also negativ definit. Daher liegt in $\left(-\frac{5}{3}, 0\right)$ ein lokales Maximum vor.

$$\operatorname{Hess}(f)\,(-1,2) = \begin{pmatrix} -2 & 4 \\ 4 & 0 \end{pmatrix}$$

Hier ist es nicht so klar, ob es positiv oder negativ definit ist. Berechnen wir

$$\begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} -2 & 4 \\ 4 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = -2v_1^2 + 8v_1v_2 = -2(v_1 - 2v_2)^2 + 8v_2^2.$$

Das zeigt, dass die Matrix indefinit ist und $(-1,2)$ ist nicht ein lokales Extremum. Denn für $(v_1, v_2) = (1,1)$ ist $-2(v_1 - 2v_2)^2 + 8v_2^2 = 6 > 0$ und für $(v_1, v_2) = (1,0)$ ist $-2(v_1 - 2v_2)^2 + 8v_2^2 = -2 < 0$.

$$\operatorname{Hess}(f)\,(-1,-2) = \begin{pmatrix} -2 & -4 \\ -4 & 0 \end{pmatrix}$$

Hier ist es ähnlich zum obenen Fall:

$$\begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} -2 & -4 \\ -4 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = -2v_1^2 - 8v_1v_2$$

ergibt für $(v_1, v_2) = (1, 1)$, dass $-2v_1^2 - 8v_1v_2 = -10$ und für $(v_1, v_2) = (1, -1)$ ist $-2v_1^2 - 8v_1v_2 = 6 > 0$, also ist die Matrix auch indefinit.