Aufgabe 27: Sei $x_0 \in \mathbb{R}^3$ und

$$Z = \{x \in \mathbb{R}^3 \,|\, x_1^2 + x_2^2 = 1\}.$$

Finden Sie $x_Z \in Z$, so dass

$$\|x_0 - x_Z\| \leq \|x_0 - x\|$$

für alle $x \in Z$.

a) Stellen Sie x_0 und ein beliebiges $x \in Z$ in Zylinderkoordinaten dar.

b) Geben Sie den Abstand $\|x_0 - x\|^2$ als Funktion $d(\varphi, z)$ an.

c) Bestimmen Sie die kritischen Punkte der Funktion $d(\varphi, z)$.

d) Berechnen Sie die Hessematrix von $d(\varphi, z)$.

e) Bestimmen Sie das Minimum der Funktion $d(\varphi, z)$.

LÖSUNG:

a)

$$x_0 = \begin{pmatrix} r_0 \cos\varphi_0 \\ r_0 \sin\varphi_0 \\ z_0 \end{pmatrix}, \qquad x = \begin{pmatrix} \cos\varphi \\ \sin\varphi \\ z \end{pmatrix}$$

b)

$$\begin{aligned} d(\varphi, z) &= \|x_0 - x\|^2 \\ &= \left\| \begin{pmatrix} r_0 \cos\varphi_0 - \cos\varphi \\ r_0 \sin\varphi_0 - \sin\varphi \\ z_0 - z \end{pmatrix} \right\|^2 \\ &= (r_0 \cos\varphi_0 - \cos\varphi)^2 + (r_0 \sin\varphi_0 - \sin\varphi)^2 + (z_0 - z)^2 \end{aligned}$$

c)

$$\begin{aligned} \partial_\varphi d(\varphi, z) &= 2\,(r_0 \cos\varphi_0 - \cos\varphi)\sin\varphi + 2\,(r_0 \sin\varphi_0 - \sin\varphi)\,(-\cos\varphi) \\ &= 2r_0 \cos\varphi_0 \sin\varphi - 2\cos\varphi\sin\varphi - 2r_0 \sin\varphi_0 \cos\varphi + 2\sin\varphi\cos\varphi \\ &= 2r_0\,(\cos\varphi_0 \sin\varphi - \sin\varphi_0 \cos\varphi) \\ &= 2r_0\,(\cos(-\varphi_0)\sin\varphi + \sin(-\varphi_0)\cos\varphi) \\ &= 2r_0 \sin(\varphi - \varphi_0) \\ \partial_z d(\varphi, z) &= -2\,(z_0 - z) \end{aligned}$$

Betrachte nun

$$\nabla d(\varphi, z) = 0.$$

Für $r_0 \neq 0$ ist dies äquivalent zu

$$\begin{pmatrix} 2r_0 \sin(\varphi - \varphi_0) \\ 2(z - z_0) \end{pmatrix} = 0 \qquad \Leftrightarrow \qquad \varphi = \varphi_0 \text{ oder } \varphi_0 + \pi \text{ und } z = z_0$$

$\Rightarrow$ (φ_0, z_0) und $(\varphi_0 + \pi, z_0)$ sind kritische Punkte der Funktion $d(\varphi, z)$.

Im Fall $r_0 = 0$ gilt

$$\begin{pmatrix} 2r_0 \sin(\varphi - \varphi_0) \\ 2(z - z_0) \end{pmatrix} = 0 \qquad \Leftrightarrow \qquad \begin{pmatrix} 0 \\ 2(z - z_0) \end{pmatrix} = 0 \qquad \Leftrightarrow \qquad z = z_0$$

Kritische Punkte sind in diesem Fall also alle Punkte (φ, z_0) mit $\varphi \in [0, 2\pi]$.

d)

$$D^2 d(\varphi, z) = \begin{pmatrix} 2r_0 \cos(\varphi - \varphi_0) & 0 \\ 0 & 2 \end{pmatrix}$$

e) $\underline{r_0 \neq 0}$

$$D^2 d(\varphi_0, z_0) = \begin{pmatrix} 2r_0 & 0 \\ 0 & 2 \end{pmatrix} \qquad \text{ist positiv definit} \Rightarrow \text{Minimum}$$

$$D^2 d(\varphi_0 + \pi, z_0) = \begin{pmatrix} -2r_0 & 0 \\ 0 & 2 \end{pmatrix} \qquad \text{ist indefinit} \Rightarrow \text{Sattelpunkt}$$

Im Fall $r_0 \neq 0$ gilt

$$x_Z = \begin{pmatrix} \cos\varphi_0 \\ \sin\varphi_0 \\ z_0 \end{pmatrix}.$$

Im Fall $r_0 = 0$ ist die Matrix $D^2 d$ positiv semidefinit, so dass wir keine allgemeine Aussage machen können. Allerdings gilt in diesem Fall

$$\begin{aligned} d(\varphi, z) &= \cos^2\varphi + \sin^2\varphi + (z_0 - z)^2 \\ &= 1 + (z_0 - z)^2, \end{aligned}$$

d.h. der Abstand hängt nicht mehr von φ sondern nur noch von z ab. Da für $d(z) = d(\varphi, z)$ die zweite Ableitung $d''(z) = 2$ größer Null ist, handelt es sich bei allen kritischen Punkten um Minima. Es gibt in diesem Fall also nicht nur einen Punkt x_z, der auf dem Zylinder Z liegt und minimalen Abstand zum Punkt x_0 hat sondern eine Menge M_Z von Punkten, die alle auf Z liegen und minimalen Abstand zum Punkt x_0 haben.

$$M_Z = \{(\varphi, z_0) \,|\, \varphi \in [0, 2\pi]\}$$

Aufgabe 28: Betrachten Sie die Funktion

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

a) Schreiben Sie f in der Form

$$f(x,y) = \frac{1}{2}A\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + b \cdot \begin{pmatrix} x \\ y \end{pmatrix},$$

wobei $A \in \mathbb{R}^{2,2}$ und $b \in \mathbb{R}^2$.

b) Berechnen und plotten bzw. skizzieren Sie die 1-Niveaulinie von f, d.h.

$$\{(x,y) \in \mathbb{R}^2 \mid f(x,y) = 1\}.$$

c) Berechnen Sie die exakte Lösung des Minimierungsproblems

$$\min_{(x,y)\in\mathbb{R}^2} f(x,y).$$

d) Führen Sie 4 Schritte des Gradientenverfahrens für den Startwert $\hat{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ durch. Benutzen Sie hierfür einen Taschenrechner, Matlab etc.

e) Führen Sie 4 Schritte des Gradientenverfahrens für den Startwert $\tilde{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ durch. Benutzen Sie hierfür einen Taschenrechner, Matlab etc.

f) Plotten bzw. skizzieren Sie die 2- und 7-Niveaulinie.

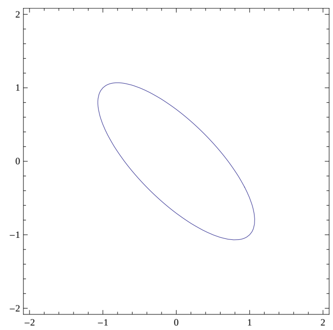
g) Fügen Sie nun die Iterationsschritte der Gradientenverfahren und deren Abstiegsrichtungen zu Ihren Skizzen hinzu. Welches typische Verhalten stellen Sie fest?

LÖSUNG:

a) Es gilt

$$f(x,y) = \frac{1}{2}\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}.$$

b) Durch Hauptachsentransformation erkennt man, dass es sich um die Ellipse mit Hauptachsen der Länge 1 in Richtung $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ und $\frac{1}{\sqrt{7}}$ in Richtung $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ handelt.



c) Aus der Vorlesung ist bekannt: x^* ist Minimierer von

$$f(x,y) = \frac{1}{2} A \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + b \begin{pmatrix} x \\ y \end{pmatrix},$$

dann löst x^*: $Ax^* = b$. In diesem Fall bedeutet dies:

$$\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad x = y = 0.$$

d) Wir folgen dem Algorithmus aus der Vorlesung und verwenden deren Notation hier.

Wir starten mit $\hat{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

1. Schritt:

$$\hat{p}_1 = A\hat{x}_0 - b = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

$$\hat{t}_1 = \frac{\hat{p}_1 \cdot \hat{p}_1}{A\hat{p}_1 \cdot \hat{p}_1} = 0.14535$$

$$\hat{x}_1 = \hat{x}_0 - \hat{t}_1 \cdot \hat{p}_1 = \begin{pmatrix} 0.41860 \\ -0.43605 \end{pmatrix}$$

2. Schritt:

$$\hat{p}_2 = A\hat{x}_1 - b = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 0.41860 \\ -0.43605 \end{pmatrix} = \begin{pmatrix} 0.36628 \\ -0.48837 \end{pmatrix}$$

$$\hat{t}_2 = \frac{\hat{p}_2 \cdot \hat{p}_2}{A\hat{p}_2 \cdot \hat{p}_2} = 0.89286$$

$$\hat{x}_2 = \hat{x}_1 - \hat{t}_2 \cdot \hat{p}_2 = \begin{pmatrix} 9.1570e - 02 \\ -5.5511e - 17 \end{pmatrix}$$

3. Schritt:

$$\hat{p}_3 = A\hat{x}_2 - b = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 9.1570e - 02 \\ -5.5511e - 17 \end{pmatrix} = \begin{pmatrix} 0.36628 \\ 0.27471 \end{pmatrix}$$

$$\hat{t}_3 = \frac{\hat{p}_3 \cdot \hat{p}_3}{A\hat{p}_3 \cdot \hat{p}_3} = 0.14535$$

$$\hat{x}_3 = \hat{x}_2 - \hat{t}_3 \cdot \hat{p}_3 = \begin{pmatrix} 0.038332 \\ -0.039929 \end{pmatrix}$$

4. Schritt:

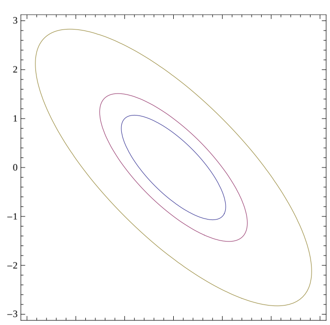
$$\hat{p}_4 = A\hat{x}_3 - b = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 0.038332 \\ -0.039929 \end{pmatrix} = \begin{pmatrix} 0.033540 \\ -0.044720 \end{pmatrix}$$
$$\hat{t}_4 = \frac{\hat{p}_4 \cdot \hat{p}_4}{A\hat{p}_4 \cdot \hat{p}_4} = 0.89286$$
$$\hat{x}_4 = \hat{x}_3 - \hat{t}_4 \cdot \hat{p}_4 = \begin{pmatrix} 8.3850e-03 \\ 6.9389e-18 \end{pmatrix}$$

e) Wir starten mit $\tilde{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

1. Schritt:

$$\tilde{p}_1 = A\hat{x}_0 - b = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \end{pmatrix}$$
$$\tilde{t}_1 = \frac{\tilde{p}_1 \cdot \tilde{p}_1}{A\tilde{p}_1 \cdot \tilde{p}_1} = \frac{1}{7},$$
$$\tilde{x}_1 = \tilde{x}_0 - \tilde{t}_1 \cdot \tilde{p}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir sind fertig!!! (D.h. alle weiteren Abstiegsrichtungen sind der Nullvektor und $\tilde{x}_1 = \tilde{x}_2 = \tilde{x}_3 = \dots$)

f) 

g) Für den Startwert $\hat{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ stellen wir das typische Zick-Zack-Muster fest und das Verfahren konvergiert. Für den Startwert $\tilde{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ liegen wir genau auf einer Hauptachsen und daher berechnet der Algorithmus nach nur einem Schritt die exakte Lösung. Dieser Fall ist für praktische Probleme nicht realistisch.

Aufgabe 29: a) Bestimmen Sie für die Funktion

$$f : x \mapsto f(x) = \cos(x)$$

die Taylor-Entwicklung an der Stelle $x_0 = \frac{\pi}{2}$ mit Restglied $O(|x - x_0|^{2n})$.

b) Bestimmen Sie für die Funktion

$$g : (x, y) \mapsto g(x, y) = \cos(x)\cos(y)$$

die Taylor-Entwicklung an der Stelle $(x_0, y_0) = (0, 0)$ mit Restglied der Ordnung 3.

LÖSUNG:

a) Die Formel für die Taylor-Entwicklung bis Ordnung n an einem Punkt x_0 ist:

$$\begin{aligned} f(x) &= \sum_{i=0}^{n} \frac{1}{i!} f^{(i)}(x_0)(x - x_0)^i + O(|x - x_0|^{n+1}) \\ &= f(x_0) + \sum_{i=1}^{n} \frac{1}{i!} f^{(i)}(x_0)(x - x_0)^i + O(|x - x_0|^{n+1}) \end{aligned}$$

wobei $f^{(i)}(x_0)$ die $i$te Ableitung an der Stelle x_0 ist.

$$\begin{aligned} f'(x_0) = f^{(1)}(x_0) &= -\sin(x_0) \\ f''(x_0) = f^{(2)}(x_0) &= -\cos(x_0) \\ f^{(3)}(x_0) &= \sin(x_0) \\ f^{(4)}(x_0) &= \cos(x_0) \end{aligned}$$

und dann für alle $t \in \mathbb{N}$

$$\begin{aligned} f^{(2t-1)}(x_0) &= (-1)^t \sin(x_0) \\ f^{(2t)}(x_0) &= (-1)^t \cos(x_0) \end{aligned}$$

Für $x_0 = \frac{\pi}{2}$,

$$\begin{aligned} f^{(2t-1)}(\frac{\pi}{2}) &= (-1)^t \\ f^{(2t)}(\frac{\pi}{2}) &= 0 \end{aligned}$$

und die Taylor-Entwicklung bis Ordnung $(2n - 1)$ an der Stelle $x = \frac{\pi}{2}$ ist:

$$f(x) = \sum_{t=1}^{n} \frac{1}{(2t-1)!}(-1)^t (x - \frac{\pi}{2})^{(2t-1)} + O\left(|x - \frac{\pi}{2}|^{2n}\right)$$

b) Die Taylor-Entwicklungsformel für g an der Stelle x_0 ist:

$$g(x_0+\xi, y_0+\zeta) = g(x_0,y_0) + \sum_n \sum_{|\alpha|=n} \frac{1}{\alpha!}\frac{\partial^\alpha g}{\partial(x,y)^\alpha}(x_0,y_0)(\xi,\zeta)^\alpha + O\left(\left\|\begin{pmatrix}\xi\\ \zeta\end{pmatrix}\right\|^{n+1}\right)$$

wobei

$$\begin{aligned}\alpha &= (\alpha_1,\alpha_2)\\ \alpha! &= \alpha_1!\alpha_2!\\ \frac{\partial^\alpha g}{\partial(x,y)^\alpha} &= \frac{\partial^{|\alpha|} g}{\partial x^{\alpha_1} y^{\alpha_2}}\\ (\xi,\zeta)^\alpha &= \xi^{\alpha_1}\zeta^{\alpha_2}\end{aligned}$$

Daraus folgt:

$$\begin{aligned}g(\xi,\zeta) = g(0,0) &+ \frac{\partial}{\partial x}g(0,0)\xi + \frac{\partial}{\partial y}g(0,0)\zeta\\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2}g(0,0)\xi^2 + \frac{\partial^2}{\partial x\partial y}g(0,0)\xi\zeta + \frac{1}{2}\frac{\partial^2}{\partial y^2}g(0,0)\zeta^2 + O\left(\left\|\begin{pmatrix}\xi\\ \zeta\end{pmatrix}\right\|^3\right)\end{aligned}$$

$$\begin{aligned}\frac{\partial g}{\partial x}(0,0) &= 0\\ \frac{\partial g}{\partial y}(0,0) &= 0\\ \frac{\partial^2 g}{\partial x^2}(0,0) &= -1\\ \frac{\partial^2 g}{\partial x\partial y}(0,0) &= 0\\ \frac{\partial^2 g}{\partial^2 y}(0,0) &= -1\end{aligned}$$

und somit

$$g(\xi,\zeta) = 1 - \frac{1}{2}\xi^2 - \frac{1}{2}\zeta^2 + O\left(\left\|\begin{pmatrix}\xi\\ \zeta\end{pmatrix}\right\|^3\right)$$

Aufgabe 30: Sei $a \in \mathbb{R}^n$ gegeben. Entwicklen Sie die Funktion $f(x) = \|x-a\|$ nach Taylor an der Stelle x_0 bis einschließlich Terme zweiter Ordnung.

LÖSUNG:

$$f(x) = \|x-a\| = \left((x_1-a_1)^2 + (x_2-a_2)^2 + \ldots + (x_n-a_n)^2\right)^{\frac{1}{2}}$$

setze $x = x_0 + h, h \in \mathbb{R}^n$

$$f(x_0+h) = f(x_0) + \operatorname{grad} f(x_0)\cdot h + \frac{1}{2}D^2 f(x_0)h\cdot h + O(\|h\|^3)$$

$$
\begin{aligned}
\text{grad } f(x) &= \left(\frac{\partial}{\partial x_1} f(x), \frac{\partial}{\partial x_2} f(x), \ldots, \frac{\partial}{\partial x_n} f(x)\right)^T \\
&= \left(\frac{1}{2} \cdot \frac{1}{\|x-a\|} \cdot 2(x_1 - a_1), \ldots, \frac{1}{2} \cdot \frac{1}{\|x-a\|} \cdot 2(x_n - a_n)\right)^T \\
&= \left(\frac{x_1 - a_1}{\|x-a\|}, \ldots, \frac{x_n - a_n}{\|x-a\|}\right)^T = \frac{1}{\|x-a\|}(x-a)
\end{aligned}
$$

$$
H(x) = \begin{pmatrix} \frac{\partial^2}{\partial x_1^2} f(x) & \frac{\partial^2}{\partial x_1 \partial x_2} f(x) & \ldots & \frac{\partial^2}{\partial x_1 \partial x_n} f(x) \\ \vdots & \vdots & \vdots & \vdots \\ \ldots & \ldots & \ldots & \frac{\partial^2}{\partial x_n^2} f(x) \end{pmatrix}
$$

$$
\begin{aligned}
\frac{\partial^2}{\partial x_i^2} f(x) &= \frac{\partial}{\partial x_i} \frac{(x_i - a_i)}{f(x)} \\
&= \frac{1}{f(x)} - \frac{(x_i - a_i)(x_i - a_i)}{f^3(x)} \\
&= \frac{1}{\|x-a\|} - \frac{1}{\|x-a\|^3}(x_i - a_i)^2
\end{aligned}
$$

$$
\begin{aligned}
& & \frac{\partial^2}{\partial x_i^2} f(x) &= \frac{1}{f(x)} - \frac{1}{f(x)^3}(x_i - a_i)^2 \\
i \neq j: & & \frac{\partial^2}{\partial x_i \partial x_j} f(x) &= \frac{\partial}{\partial x_j} \frac{(x_i - a_i)}{f(x)} \\
& & &= -\frac{1}{f^3(x)}(x_i - a_i)(x_j - a_j) \\
& & &= -\frac{(x_i - a_i)(x_j - a_j)}{\|x-a\|^3}
\end{aligned}
$$

$$
\Rightarrow H(x) = \frac{1}{\|x-a\|} \left(\mathbb{1} - \frac{(x-a)}{\|x-a\|} \frac{(x-a)^T}{\|x-a\|}\right)
$$

Insgesamt ergibt sich

$$
f(x_0 + h) = \|x_0 - a\| + \frac{(x_0 - a)}{\|x_0 - a\|} \cdot h + \frac{1}{2\|x_0 - a\|} \left(\mathbb{1} - \frac{(x-a)}{\|x-a\|} \frac{(x-a)^T}{\|x-a\|}\right) h \cdot h + O(\|h\|^3)
$$

Zusätzliche Erläuterung: (Nicht Teil der Lösung!)
Mit der Bezeichnung $g = \text{grad } f(x_0)$ erhalten wir

$$
f(x_0 + h) = f(x_0) + g \cdot h + \frac{1}{2\|x_0 - a\|} \left(\|h\|^2 - (g \cdot h)^2\right) + (\|h\|^3)\,.
$$

Dabei ist offenbar $\|g\| = 1$, also ist g der Einheitsvektor, der von a in Richtung x_0 zeigt. Der lineare Term (also die Approximation der Änderung in erster Ordnung) ist daher die Projektion von h auf die Gerade durch x_0 und a. Hier spielt also nur der Anteil von h eine Rolle, der auf a zu oder von a weg zeigt, nicht der Anteil „seitwärts“.

In ähnlicher Weise erklärt sich der quadratische Term: Mit $s^2 = \|h\|^2 - (g \cdot h)^2$ ist s der „Seitwärts-Anteil“ von h (Pythagoras!). Der Term zweiter Ordnung berücksichtigt also die Änderung „seitwärts“.

Die Skalierung überlegt man sich beispielweise folgendermaßen: Mit $a = (0,0)^T$, $x_0 = (1,0)^T$ und $h = (t,s)^T$ erhält man $f(x_0 + h) = 1 + t + \frac{1}{2}s^2 + O(\|h\|^3)$. Zu $x_0 = (L,0)^T$ erhält man die skalierte Gleichung $\frac{f(x_0+h)}{L} = 1 + \frac{t}{L} + \frac{1}{2}\left(\frac{s}{L}\right)^2 + O(\|h\|^3)$. Multiplikation mit $L = \|x_0 - a\|$ ergibt schließlich die obige Form.