

Beweis. Der Beweis erfolgt durch vollständige Induktion. Angenommen wir hätten den Satz für $k - 1$ gezeigt. Dann ist wegen

$$\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = 0$$

auch

$$0 = D_{\mathbf{y}}\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = D_1\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y})D_{\mathbf{y}}\mathbf{g}(\mathbf{y}) + D_2\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}).$$

Damit ist

$$D_{\mathbf{y}}\mathbf{g}(\mathbf{y}) = -(D_1\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}))^{-1} D_2\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}).$$

Also ist die rechte Seite in $C^{k-1}(B_\delta(\mathbf{y}_0), L(Y; X))$, dies ist aufgrund der angegebenen Gleichung die erste Ableitung von $\mathbf{g}$, also ist $\mathbf{g} \in C^k(B_\delta(\mathbf{y}_0); X)$. $\square$

Fragt man sich, wann differenzierbare Funktionen invertierbar sind, erlaubt einem die eben hergeleitete Beziehung eine einfache notwendige Bedingung zu finden. Existiert $\mathbf{f}^{-1}$ und ist dies differenzierbar, so gilt

$$\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x})) = \mathbf{x}$$

und die Differentiation dieser Formel liefert sofort

$$D(\mathbf{f}^{-1})(\mathbf{f}(\mathbf{x}))D\mathbf{f}(\mathbf{x}) = \mathbb{1},$$

also $D(\mathbf{f}^{-1})(\mathbf{f}(\mathbf{x}))$ ist invertierbar und

$$D(\mathbf{f}^{-1})(\mathbf{f}(\mathbf{x})) = (D\mathbf{f}(\mathbf{x}))^{-1}.$$

Dass die Invertierbarkeit des Differentials auch hinreichend für die Umkehrbarkeit der Funktion ist, folgt aus dem nachfolgenden Satz.

Satz 9.2.7 (Umkehrsatz)

Es seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ ein Banachräume. Ferner sei $U \subset X$ offen, $\mathbf{f} : U \rightarrow Y$ differenzierbar. Ist $D\mathbf{f}(\mathbf{x}_0)$ stetig invertierbar, so gibt es eine Umgebung $U_1 \subset U$ von $\mathbf{x}_0$ und $V \subset Y$ von $\mathbf{f}(\mathbf{x}_0)$, so dass $\mathbf{f} : U_1 \rightarrow V$ eine differenzierbare Bijektion ist mit $\mathbf{f}^{-1} : V \rightarrow U_1$ differenzierbar und es gilt

$$D(\mathbf{f}^{-1})(\mathbf{f}(\mathbf{x})) = (D\mathbf{f}(\mathbf{x}))^{-1}.$$

Beweis. Für $\mathbf{x} \in U$ und $\mathbf{y} \in Y$ setzen wir

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{y} - \mathbf{f}(\mathbf{x}).$$

Dann ist

$$\mathbf{F}(\mathbf{x}_0, \mathbf{f}(\mathbf{x}_0)) = 0$$

und

$$D_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \mathbf{y}) = -D\mathbf{f}(\mathbf{x}).$$

Daher ist $D_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \mathbf{y})$ im Punkt $(\mathbf{x}_0, \mathbf{f}(\mathbf{x}_0))$ stetig invertierbar, also gibt es Umgebungen Γ von $\mathbf{f}(\mathbf{x}_0) = \mathbf{y}_0$ und Ω von $\mathbf{x}_0$ und eine Abbildung $\mathbf{g} : \Gamma \rightarrow \Omega$ mit

$$F(\mathbf{g}(\mathbf{y}), \mathbf{y}) = 0$$

oder $\mathbf{x} = \mathbf{g}(\mathbf{y})$. Nach der Vorbemerkung oder der im Satz über implizite Funktionen angegebenen Formel gilt

$$D\mathbf{g}(\mathbf{y}_0) = (D\mathbf{f}(\mathbf{x}_0))^{-1}. \quad \square$$

Definition 9.2.8

Sind $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen und $\Psi : U \rightarrow V$ eine differenzierbare Bijektion, so dass Ψ , Ψ^{-1} stetig differenzierbar sind, dann heißt Ψ Diffeomorphismus.

Beispiel 9.2.9 (Polarkoordinaten)

Wir betrachten die Abbildung

$$\Phi : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}^2 : (r, \varphi) \mapsto \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}.$$

Für $r > 0$ gilt

$$\Phi'(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Die Determinante dieser Jacobimatrix ist

$$\det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = r(\cos^2 \varphi + \sin^2 \varphi) = r > 0.$$

Demzufolge ist der Umkehrsatz anwendbar, auf jeder Menge der Form $\mathbb{R}_+ \times (a - \pi, a + \pi)$ ist $\Phi_a = \Phi : \mathbb{R}_+ \times (a - \pi, a + \pi) \rightarrow \mathbb{R}^2$ ein Diffeomorphismus mit Bild

$$\mathbb{R}^2 \setminus \left\{ t \begin{pmatrix} \cos a \\ \sin a \end{pmatrix} \mid t \geq 0 \right\}.$$

Die Umkehrabbildung von Φ_a ist

$$\Phi_a^{-1}(\mathbf{x}) = (\sqrt{x_1^2 + x_2^2}, \varphi),$$

wobei $\cos \varphi = \frac{x_1}{\sqrt{x_1^2 + x_2^2}}$, $\sin \varphi = \frac{x_2}{\sqrt{x_1^2 + x_2^2}}$ und $\varphi \in (a - \pi, a + \pi)$. Für $a = 0$ nennt man die Paare (r, φ) die *Polarkoordinaten* des Punktes $\mathbf{x} \in \mathbb{R}^2$.

Typisch für den Satz über implizite Funktionen ist eine lokale Aussage, man beschreibt die Lösungsmenge einer Gleichung in der Nähe einer gegebenen Lösung,

insofern ist der folgende Satz als Prototyp einer anderen Entwicklung zu sehen. Es ist die einfachste Version eines solchen Satzes, formuliert wurden Sätze mit dieser Zielrichtung in der siebziger Jahren des 20. Jahrhunderts. Wir benötigen dazu noch einen Begriff.

Definition 9.2.10 (Zusammenhangskomponente)

Es sei (X, d) ein metrischer Raum $A \subset X$ eine nichtleere Teilmenge, $x \in A$ sei gegeben. Die Zusammenhangskomponente $C_x(A)$ von A des Punktes x ist gegeben, als

$$C_x(A) = \bigcup_{T \subset A: T \text{ ist zusammenhängend}} T.$$

Satz 9.2.11 (Satz über implizite Funktionen: globale Version)

Es sei $(X, \|\cdot\|_X)$ ein Banachraum, $U \subset X$ ein Gebiet, $h : U \rightarrow \mathbb{R}$ sei im Punkt $\mathbf{x}_0 \in U$ differenzierbar, $h(\mathbf{x}_0) = 0$ und $Dh(\mathbf{x}_0) \neq 0$. Dann zerlegt die Menge $h^{-1}(0)$, die den Punkt $\mathbf{x}_0$ enthält das Gebiet U in mindestens zwei Komponenten, d. h. sei $C = h^{-1}(0)$, so besitzt $U \setminus C$ mindestens zwei Komponenten.

Beweis. $C = h^{-1}(0) \neq U$ ist abgeschlossen in U (d. h. bzgl. des metrischen Raumes (U, d) , wobei d die zur Norm in X gehörende Metrik ist). C ist abgeschlossen. $U \setminus C \neq \emptyset$ ist daher offen. Damit ist gleichwertig (Aufgabe ???), dass $U \setminus C$ zusammenhängend und $U \setminus C$ wegezusammenhängend gleichwertig ist. Angenommen der Satz ist falsch. Dann ist $U \setminus C$ wegezusammenhängend. Es gibt Punkte $x_{1,2} \in U \setminus C$ mit $h(x_1) > 0 > h(x_2)$. Dann gibt es einen Weg $\gamma : [0, 1] \rightarrow U \setminus C$ mit $\gamma(0) = x_1$, $\gamma(1) = x_2$. Wegen der Stetigkeit von $h \circ \gamma$ und der Tatsache, dass $h \circ \gamma(0) > 0 > h \circ \gamma(1)$ gibt es ein $\tau \in (0, 1)$ mit

$$h \circ \gamma(\tau) = 0.$$

Dann ist $h(\gamma(\tau)) = 0$ und $\gamma(\tau) \in C$. Dies liefert einen Widerspruch, der den Satz beweist. $\square$

9.3 Extrema mit Nebenbedingungen

In diesem Abschnitt wollen wir die Maxima und Minima von Funktionen bestimmen, allerdings nicht auf einer offenen Menge, sondern auf einer durch Gleichungen definierten Teilmenge. Zur Motivation wieder ein paar Beispiele.

Beispiel 9.3.1 (Nebenbedingungen)

1. Zunächst beginnen wir mit dem Paradebeispiel der Wanderung im Gebirge. Wir beschreiben den Wanderweg als Bild einer Kurve auf der Landkarte und suchen nach dem Punkt der höchsten (geringsten) Höhe über NN längs des Weges.
2. Wir betrachten die Menge der Zahlenpaare $\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_i > 0 \right\}$ als Kantenlängen von Rechtecken. Suchen wir nun unter denen gleicher Fläche $h(\mathbf{x}) = x_1 x_2 = c$ das maximaler Kantenlänge $L = 2(x_1 + x_2)$, so lösen wir das Problem: Maximiere L unter der Nebenbedingung $h = c$.
3. Ein Zug fahre von A nach B , wobei der Abstand der beiden Stationen $d > 0$ sei. Als Steuermöglichkeit kann der Zugführer die Beschleunigungen $-a, 0, a$ für eine Zahl $a > 0$ wählen, wobei er die Zeitintervalle der Beschleunigung auswählt, d. h. er bestimmt eine Treppenfunktion φ , die nur die Werte $a, 0, -a$ annimmt, so dass mit $v'(t) = \varphi(t)$ gilt

$$\int_0^\infty v(t) \, dt = d,$$

(dies ist die Bedingung, dass die zurückgelegte Strecke d ist). In der ökologischen Variante will er nun unter der angegebenen Bedingung mit minimalen Energieverbrauch, in der ökonomischen Variante in möglichst kurzer Zeit ankommen. Dabei ist die Energie (proportional zu)

$$E = \int_0^\infty \varphi_+(t) \, dt$$

4. Wir wollen nun eine rein mathematische Bedingung angeben. Angenommen $\gamma(t)$ sei eine Kurve im $\mathbb{R}^2$, die $h = c$ parametrisiert. Wir suchen eine Extremalstelle $\mathbf{x}_0$ einer differenzierbaren Funktion unter der Nebenbedingung $h(\mathbf{x}) = c$. Also ist mit $\gamma(t_0) = \mathbf{x}_0$

$$(f \circ \gamma)'(t_0) = 0$$

oder

$$\langle \nabla f(\gamma(t_0)), \gamma'(t_0) \rangle = 0.$$

Da $h(\gamma(t)) = 0$ ist, folgt auch $\langle \nabla h(\gamma(t_0)), \gamma'(t_0) \rangle = 0$ und damit sind $\nabla f(\mathbf{x}_0)$ und $\nabla h(\mathbf{x}_0)$ linear abhängig. Dieses Beispiel wird Grundlage der weiteren Überlegungen.

Satz 9.3.2 (Lagrange-Multiplikator)

Es sei $(X, \|\cdot\|_X)$ ein Banachraum, ferner sei $U \subset V$ offen, $h : U \rightarrow \mathbb{R}$ sei eine stetig differenzierbare Abbildung, es sei $c \in \mathbb{R}$ mit $Dh(\mathbf{x}) \neq 0$ für alle $\mathbf{x}$ mit $h(\mathbf{x}) = c$. $f : U \rightarrow \mathbb{R}$ sei stetig differenzierbar. Sei $\mathbf{x}_0 \in U$ mit $h(\mathbf{x}_0) = c$, so dass ein $\delta > 0$ existiert mit $\|\mathbf{x} - \mathbf{x}_0\|_X < \delta$ und $h(\mathbf{x}) = c$ implizieren $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ (d. h. f hat in $\mathbf{x}_0$ ein lokales Maximum unter der Nebenbedingung $h(\mathbf{x}) = c$). Dann gibt es ein $\lambda \in \mathbb{R}$ mit

$$Df(\mathbf{x}_0) = \lambda Dh(\mathbf{x}_0).$$

Definition 9.3.3 (Lagrange-Multiplikator)

Die Zahl λ aus dem Satz 9.3.2 heißt Lagrange-Multiplikator.

Beweis. Sei $\mathbf{x}_0$ wie im Satz 9.3.2 angegeben. Da $h(\mathbf{x}_0) = c$ ist, ist laut Voraussetzung

$$Dh(\mathbf{x}_0) \neq 0.$$

Dann kann man schreiben

$$X = \ker Dh(\mathbf{x}_0) \oplus [\mathbf{y}_0] = K \oplus Y.$$

K, Y sind mit geeigneten Normen Banachräume, wir schreiben für $\mathbf{k} \in K, \mathbf{y} \in Y$

$$H(\mathbf{k}, \mathbf{y}) = h(\mathbf{x}_0 + \mathbf{k} + \mathbf{y}).$$

Dann ist für $H(0, 0) = c$ und $D_2H(0, 0) = Dh(\mathbf{x}_0) : Y \rightarrow \mathbb{R}$ (Ableitung bezüglich der zweiten Variablen) ein Isomorphismus, insbesondere gibt es aufgrund des Satzes über implizite Funktionen, Satz 9.2.2, Umgebungen $\Gamma \ni \mathbf{k}$, $\Omega \ni \mathbf{y}$ und $g : \Gamma \rightarrow \Omega$ mit

$$H(\mathbf{k}, g(\mathbf{k})) = c.$$

Betrachte nun

$$F(\mathbf{k}, g(\mathbf{k})) = f(\mathbf{x}_0 + \mathbf{k} + g(\mathbf{k})).$$

F ist auf einer Umgebung von $0 \in K$ definiert und hat bei 0 ein lokales Extremum. Also ist

$$D_{\mathbf{k}}F(0) = 0.$$

Dies impliziert $Df(\mathbf{x}_0)|_K = 0$. Betrachten wir nun

$$Df_{\mathbf{x}_0, \mathbf{y}_0}(t) = Df(\mathbf{x}_0)(\mathbf{y}_0) \in \mathbb{R}.$$

Da $Dh(\mathbf{x}_0)(\mathbf{y}_0) \neq 0$ gibt es eine reelle Zahl λ mit

$$\lambda Dh(\mathbf{x}_0)(\mathbf{y}_0) = Df(\mathbf{x}_0)(\mathbf{y}_0).$$

Für beliebiges $\mathbf{x} \in X$ ist nun $\mathbf{x} = \mathbf{k} + t\mathbf{y}_0$. Dann ist

$$Df(\mathbf{x}_0)(\mathbf{x}) = Df(\mathbf{x}_0)(\mathbf{k} + t\mathbf{y}_0) = tDf(\mathbf{x}_0)(\mathbf{y}_0) = t\lambda Dh(\mathbf{x}_0)(\mathbf{y}_0) = \lambda Dh(\mathbf{x}_0)(\mathbf{x}).$$

□

Ganz entsprechend kann man mehrere Nebenbedingungen behandeln. Man erhält dann die nachfolgende Aussage.

Satz 9.3.4 (Lokale Extrema und Nebenbedingungen)

Es sei $(X, \|\cdot\|_X)$ ein Banachraum, $U \subset X$ offen, $\mathbf{h} : U \rightarrow \mathbb{R}^k$ eine stetig differenzierbare Abbildung mit

$$\left\{ Dh_i(\mathbf{x}) \mid i = 1, \dots, k \right\} \text{ ist linear unabhängig } \forall \mathbf{x} \text{ mit } \mathbf{h}(\mathbf{x}) = \mathbf{c}.$$

Ferner sei $f : U \rightarrow \mathbb{R}$ stetig differenzierbar.

Sei $\mathbf{x}_0 \in U$ mit $\mathbf{h}(\mathbf{x}_0) = \mathbf{c}$, so dass ein $\delta > 0$ existiert mit $\|\mathbf{x} - \mathbf{x}_0\|_X < \delta$ und $\mathbf{h}(\mathbf{x}) = \mathbf{c}$ implizieren $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ (d. h. f hat in $\mathbf{x}_0$ ein lokales Maximum unter der Nebenbedingung $\mathbf{h}(\mathbf{x}) = \mathbf{c}$).

Dann gibt es $\lambda_i \in \mathbb{R}$, $i = 1, \dots, k$ mit

$$Df(\mathbf{x}_0) = \sum_{i=1}^k \lambda_i Dh_i(\mathbf{x}_0).$$

Definition 9.3.5 (Lagrange-Multiplikatoren)

Die Zahlen $\lambda_1, \dots, \lambda_k$ heißen Lagrange Multiplikatoren.

Korollar 9.3.6 (Lagrange-Multiplikatoren und Gradienten)

Es sei $(V, \langle \cdot, \cdot \rangle_V)$ ein Hilbertraum, bzw. $V = (\mathbb{R}^n, \langle \cdot, \cdot \rangle_2)$. Dann gilt unter den Voraussetzungen des Satzes 9.3.4

$$\nabla f(\mathbf{x}_0) = \sum_{i=1}^k \lambda_i \nabla h_i(\mathbf{x}_0)$$