Vereinbarung *Im Folgenden heiße eine nichtleere Teilmenge M des $\mathbb{R}^{n+m}$ eine f-definierte Mannigfaltigkeit, wenn $M = f^{-1}(0)$ mit einer stetig differenzierbaren Funktion $f\colon \mathbb{R}^{n+m} \hookrightarrow \mathbb{R}^m$ mit regulärem Wert 0.* ⋊

▪ Tangentialraum und Normalraum

Jedem Punkt p einer Mannigfaltigkeit M können wir *Tangentialvektoren* zuordnen. Dazu betrachten wir beliebige Kurven auf M durch p und deren Geschwindigkeitsvektoren in diesem Punkt.

18 **Definition** *Ein Vektor $v \in \mathbb{R}^{n+m}$ heißt Tangentialvektor an M im Punkt p, wenn es eine C^1-Kurve $\gamma\colon \mathbb{R} \hookrightarrow M$ gibt mit*

$$\gamma(0) = p, \qquad \dot\gamma(0) = v.$$

Die Menge aller Tangentialvektoren an M im Punkt p heißt der Tangentialraum von M an p und wird mit T_pM bezeichnet. ⋊

Man stellt sich Tangentialvektoren üblicherweise als im Punkt p angeheftete ›gebundene‹ Ortsvektoren vor. Tatsächlich handelt es sich um ›ungebundene‹ Vektoren, also Elemente eines Vektorraumes. Der Raum T_pM ist ein *Vektorraum* derselben Dimension wie M:

19 **Satz** *Sei M eine f-definierte Mannigfaltigkeit. Dann ist*

$$T_pM = \ker Df(p), \qquad p \in M.$$

Insbesondere ist jeder Tangentialraum T_pM ein Vektorraum derselben Dimension wie M. ⋊

⟨⟨⟨⟨⟨ Sei $F\colon U \to V$ ein lokaler trivialisierender Diffeomorphismus um p $_{14}$, so dass

$$F(M \cap U) = \mathbb{R}^n \times 0^m \cap V.$$

Für die Umkehrabbildung $\Phi = F^{-1}$ um $z = F(p)$ gilt dann $f \circ \Phi \equiv 0$ entlang aller Kurven in $\mathbb{R}^n \times 0^m \cap V$. Insbesondere gilt

$$(f \circ \Phi)(z + te_i) \equiv 0, \qquad 1 \leqslant i \leqslant n,$$

auf einem t-Intervall um 0, und damit

$$Df(p)D\Phi(z)e_i = 0, \qquad 1 \leqslant i \leqslant n.$$

Die Vektoren $D\Phi(z)e_i$ liegen somit sämtlich im Kern von $Df(p)$. Außerdem sind sie linear unabhängig. Da der Kern von $Df(p)$ aber Dimension n hat, folgt

$$\operatorname{span}\{D\Phi(z)e_i : 1 \leqslant i \leqslant n\} = \ker Df(p).$$

Auf der anderen Seite ist klar, dass der Raum links die Geschwindigkeitsvektoren sämtlicher möglicher Kurven auf M durch p enthält, also

$$T_pM = \operatorname{span}\{D\Phi(z)e_i : 1 \leqslant i \leqslant n\}$$

gilt. ⟫⟫⟫⟫⟫

Versehen wir den $\mathbb{R}^{n+m}$ mit einem beliebigen Skalarprodukt $\langle\cdot,\cdot\rangle$, so ist auch die Orthogonalität zweier Vektoren und das orthogonale Komplement einer Teilmenge erklärt. Damit können wir auch den *Normalraum* einer Mannigfaltigkeit definieren.

Definition *Sei M eine Mannigfaltigkeit im $\mathbb{R}^{n+m}$. Ist dieser mit einem Skalarprodukt versehen, so heißt das orthogonale Komplement zum Tangentialraum T_pM der Normalraum von M in p und wird mit $T_p^{\perp}M$ bezeichnet. Seine Elemente heißen die Normalenvektoren von M in p.* ⋊

Für jeden Punkt $p \in M$ gilt also

$$T_pM \oplus T_p^{\perp}M = \mathbb{R}^{n+m}.$$

20 **Satz** *Sei M eine f-definierte Mannigfaltigkeit der Dimension n im $\mathbb{R}^{n+m}$. Ist dieser mit einem Skalarprodukt versehen, so ist*

$$T_p^{\perp}M = \operatorname{span}\{\nabla f_1(p), .., \nabla f_m(p)\}, \qquad p \in M.$$ ⋊

Die Gradienten der Komponentenfunktionen von f bezüglich dieses Skalarproduktes stehen also überall senkrecht auf der Niveaumenge $M = f^{-1}(0)$ und spannen in jedem Punkt deren Normalraum auf.

⟪⟪⟪⟪⟪ Nach dem letzten Satz ist $T_pM = \ker Df(p)$. Für jede Komponentenfunktion f_k von f gilt also

$$Df_k(p)v = \langle\nabla f_k(p), v\rangle = 0, \qquad v \in T_pM.$$

Also ist

$$\nabla f_k(p) \in T_p^{\perp}M, \qquad 1 \leqslant k \leqslant m.$$

Andererseits sind die Gradienten $\nabla f_1(p), .., \nabla f_m(p)$ linear unabhängig, da f im Punkt p regulär ist. Da der Normalraum $T_p^{\perp}M$ genau die Dimension m hat, müssen diese m linear unabhängigen Vektoren den gesamten Normalraum aufspannen. ⟫⟫⟫⟫⟫

▹ Wir betrachten noch einmal

$$f : \mathbb{R}^{n+1} \to \mathbb{R}, \quad f(x) = \|x\|^2.$$

Der Gradient von f verschwindet nur im Nullpunkt, und dort ist $f(0) = 0$. Also ist 1 ein regulärer Wert von f, und $\mathbb{S}^n = f^{-1}(1)$ eine Mannigfaltigkeit der Kodimension 1 und Dimension n. Ein Normalenvektor im Punkt $x \in \mathbb{S}^n$ ist x selbst. Somit gilt

$$T_x^{\perp}\mathbb{S}^n = \operatorname{span}\{x\}, \qquad T_x\mathbb{S}^n = \{x\}^{\perp}.$$

Die *Tangentialebene* an $\mathbb{S}^n$ im Punkt x ist übrigens die zu T_xM parallele affine Ebene durch den Punkt x, also

$$\begin{aligned} x + T_x\mathbb{S}^n = x + \{x\}^{\perp} &= x + \{v : \langle v, x\rangle = 0\} \\ &= \{v : \langle v - x, x\rangle = 0\} \\ &= \{v : \langle v, x\rangle = 1\}. \end{aligned}$$

18.4 Extrema mit Nebenbedingungen

Ein oft auftretendes Problem ist, Extremwerte einer skalaren Funktion zu bestimmen, während gleichzeitig eine Anzahl von Nebenbedingungen in Form von Gleichungen einzuhalten sind.

Ein anschauliches Beispiel ist die Aufgabe, auf einer Fläche diejenigen Punkte mit dem größten oder kleinstem Abstand zu einem Referenzpunkt zu finden. Als Zielfunktion dient hier bequemerweise das *Quadrat* des euklidischen Abstands, also

$$f(x) = x_1^2 + .. + x_n^2.$$

Die Fläche selbst soll durch m Gleichungen

$$g_i(x) = 0, \qquad 1 \leqslant i \leqslant m.$$

beschrieben werden.

Ist 0 ein regulärer Wert der Vektorfunktion

$$g = (g_1, .., g_m)^{\top} : \mathbb{R}^n \hookrightarrow \mathbb{R}^m,$$

so handelt es sich bei dieser Fläche um eine durch g bestimmte Mannigfaltigkeit M der Kodimension m. Es geht dann darum, Extremwerte der *Einschränkung* $f|_M$ von f auf M zu bestimmen. Die primäre Aufgabe ist dabei, deren *kritischen Punkte* zu finden. Die Frage, ob dort ein Maximum oder Minium vorliegt, ergibt sich meist aus geometrischen Überlegungen.

Naheliegend ist die Idee, die Mannigfaltigkeit M mehr oder weniger geschickt zu parametrisieren und dadurch die Nebenbedingungen aufzulösen. Die Zielfunktion könnte man dann wie gewohnt untersuchen 15.14. Dies ist allerdings im Allgemeinen sehr mühsam – und auch garnicht nötig. Denn solche kritischen Punkte lassen sich geometrisch sehr einfach charakterisieren.

Im Folgenden hat der Gesamtraum die Dimension n. Eine Mannigfaltigkeit der Kodimension m hat jetzt also die Dimension $n - m$.

21 **Satz** *Sei M eine Mannigfaltigkeit und f eine in einer Umgebung von M stetig differenzierbare skalare Funktion. Dann besitzt $f|_M$ einen kritischen Punkt in $p \in M$ genau dann, wenn*

$$\nabla f(p) \in T_p^{\perp} M. \qquad \rtimes$$

⟨⟨⟨⟨⟨ Die Einschränkung der Funktion f auf M besitzt im Punkt $p \in M$ einen kritischen Punkt genau dann, wenn sie entlang aller Kurven auf M durch p dort einen kritischen Punkt besitzt. Somit müssen sämtliche Richtungsableitungen in *tangentialen* Richtungen an M verschwinden:

$$Df(p)v = \langle \nabla f(p), v \rangle = 0, \qquad v \in T_p M.$$

Also gilt $\nabla f(p) \perp T_p M$ und damit $\nabla f(p) \in T_p^{\perp} M$. ⟩⟩⟩⟩⟩

Dies ist ein abstraktes geometrisches Resultat. Es ist unabhängig davon, *wie* die Mannigfaltigkeit M gegeben ist. Konkreter wird es für gleichungsdefinierte Mannigfaltigkeiten.

22 **Korollar** *Sei M eine g-definierte Mannigfaltigkeit im $\mathbb{R}^n$ der Kodimension m und $f : \mathbb{R}^n \hookrightarrow \mathbb{R}$ in einer Umgebung von M stetig differenzierbar. Dann besitzt $f|_M$ einen kritischen Punkt in $p \in M$ genau dann, wenn es reelle Zahlen $\lambda_1, .., \lambda_m$ gibt, genannt Lagrangemultiplikatoren, so dass*

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + .. + \lambda_m \nabla g_m(p). \qquad \rtimes$$

⟨⟨⟨⟨⟨ Für einen kritischen Punkt p von $f|_M$ gilt ja 20, 21

$$\nabla f(p) \in T_p^{\perp} M = \operatorname{span}\{\nabla g_1(p), .., \nabla g_m(p)\}.$$

Also ist $\nabla f(p)$ eine Linearkombination von $\nabla g_1(p), .., \nabla g_m(p)$. ⟩⟩⟩⟩⟩

Für die praktische Anwendung formulieren wir dieses Ergebnis noch einmal für den Standardfall. Aus kritischen Punkten der Funktion f *mit* Nebenbedingungen werden kritische Punkte einer Hilfsfunktion F *ohne* Nebenbedingungen, wobei allerdings die Dimension des Problems vergrößert wird.

Satz *Es seien $f\colon \mathbb{R}^n \hookrightarrow \mathbb{R}$ und $g\colon \mathbb{R}^n \hookrightarrow \mathbb{R}^m$ auf einem gemeinsamen Definitionsbereich stetig differenzierbar und 0 ein regulärer Wert von g. Dann besitzt f unter der Nebenbedingung $g = 0$ einen kritischen Punkt p genau dann, wenn die erweiterte Funktion*

$$F\colon\ \mathbb{R}^n \times \mathbb{R}^m \hookrightarrow \mathbb{R}, \quad F(x,\lambda) = f(x) + \langle \lambda, g(x) \rangle$$

einen kritischen Punkt (p,μ) besitzt. ⋊

⟨⟨⟨⟨⟨ Besitzt F einen kritischen Punkt (p,μ), so ist erstens

$$0 = \partial_{\lambda_k} F(p,\mu) = g_k(p), \qquad 1 \leqslant k \leqslant m,$$

und zweitens

$$0 = \partial_{x_l} F(p,\mu) = \partial_{x_l} f(p) + \sum_{k=1}^{m} \mu_k \, \partial_{x_l} g_k(p), \qquad 1 \leqslant l \leqslant n.$$

Ersteres ist äquivalent mit $p \in M = g^{-1}(0)$, und Letzteres ist äquivalent mit

$$0 = \nabla F(p,\mu) = \nabla f(p) + \sum_{k=1}^{m} \mu_k \nabla g_k(p),$$

also mit $\nabla f(p) \in T_p^{\perp} M$. Also ist p ein kritischer Punkt von f unter den Nebenbedingungen $g = 0$ [22]. – Die Umkehrung beweist man mit denselben Argumenten. ⟩⟩⟩⟩⟩

Damit erhalten wir folgendes Verfahren zur Lösung von Extremalaufgaben mit Nebenbedingungen. Sind die Zielfunktion f und die Nebenbedingungen $g = (g_1, .., g_m) = 0$ gegeben, und ist 0 ein regulärer Wert von g, so bilden wir die erweiterte Funktion

$$F = f + \langle \lambda, g \rangle$$

und bestimmen deren kritische Punkte bezüglich x *und* λ. Dies führt zu einem System aus $n + m$ Gleichungen für ebensoviele Unbekannte $x_1, .., x_n$ und $\lambda_1, .., \lambda_m$, und zwar

$$\begin{aligned} 0 &= \partial_{x_k} f(x) + \lambda_1 \partial_{x_k} g_1(x) + .. + \lambda_m \partial_{x_k} g_m(x), && 1 \leqslant k \leqslant n, \\ 0 &= g_l(x), && 1 \leqslant l \leqslant m. \end{aligned}$$

Dieses ist zu lösen. Der Erfolgt ist allerdings nicht garantiert, da es sich im Allgemeinen um *nichtlineare* Gleichungen handelt. — Dazu zwei Beispiele.

23 ▷ *Youngsche Ungleichung* Betrachte

$$f(u,v) = \frac{u^p}{p} + \frac{v^q}{q},$$

wobei

$$p, q > 1, \qquad \frac{1}{p} + \frac{1}{q} = 1.$$

Auf dem positiven Quadranten $\{(u,v) : u, v > 0\}$ besitzt f keine lokalen Extrema. Wir beschränken uns deshalb auf die Teilmenge mit $uv = 1$, stellen also die Nebenbedingung

$$g(u,v) = uv - 1 = 0.$$

Die erweiterte Funktion ist dann

$$F(u,v,\lambda) = \frac{u^p}{p} + \frac{v^q}{q} + \lambda(uv-1).$$

Für die kritischen Punkte erhalten wir das Gleichungssystem

$$\begin{aligned} u^{p-1} + \lambda v &= 0, \\ v^{q-1} + \lambda u &= 0, \\ uv &= 1. \end{aligned}$$

Multiplizieren der ersten Gleichung mit u und der zweiten mit v ergibt $u^p = v^q$, und zusammen mit $uv = 1$ folgt $u = v = 1$. Offensichtlich handelt es sich hierbei um ein absolutes Minimum. Als Ergebnis erhalten wir

$$\frac{u^p}{p} + \frac{v^q}{q} \geqslant \frac{1}{p} + \frac{1}{q} = 1$$

für alle $u, v > 0$ mit $uv = 1$.

Um die allgemeine Ungleichung für $u, v > 0$ zu erhalten, setzen wir

$$\tilde{u} = \frac{u}{(uv)^{1/p}}, \qquad \tilde{v} = \frac{v}{(uv)^{1/q}}.$$

Es ist $\tilde{u}\tilde{v} = 1$, also

$$\frac{\tilde{u}^p}{p} + \frac{\tilde{v}^q}{q} = \frac{u^p}{puv} + \frac{v^q}{quv} \geqslant 1.$$

Also gilt allgemein

$$\frac{u^p}{p} + \frac{v^q}{q} \geqslant uv, \qquad u, v > 0. \quad ◂$$

▹ *Eigenwerte* Sei $A \in L(\mathbb{R}^{n+1})$ ein symmetrischer Operator bezüglich eines beliebigen Skalarprodukts $\langle \cdot , \cdot \rangle$ auf dem $\mathbb{R}^{n+1}$. Die quadratische Form

$$f : \mathbb{R}^{n+1} \to \mathbb{R}, \quad f(x) = \langle Ax, x \rangle$$

ist auf dem Gesamtraum unbeschränkt, wenn sie nicht identisch Null ist. Auf der kompakten Einheitssphäre $\mathbb{S}^n = \{x : \langle x, x \rangle = 1\}$ nimmt sie aber ihr Minimum und Maximum an, besitzt also dort im Fall $n \geqslant 2$ mindestens zwei kritische Punkte.

Um diese zu charakterisieren, schreiben wir die Nebenbedingung als

$$g(x) = 1 - \langle x, x \rangle \,.$$

Die erweiterte Funktion lautet dann

$$F(x, \lambda) = \langle Ax, x \rangle + \lambda(1 - \langle x, x \rangle).$$

Das zugehörige Gleichungssystem ist

$$\begin{aligned} Ax - \lambda x &= 0, \\ \langle x, x \rangle &= 1. \end{aligned}$$

Ein kritischer Punkt der quadratischen Form f auf $\mathbb{S}^n$ ist demnach ein normierter *Eigenvektor* von A, sein Lagrangemultiplikator der zugehörige *Eigenwert*.

Da die quadratische Form kritische Punkte besitzen *muss*, erhalten wir damit einen Beweis, dass jeder symmetrische Operator mindestens einen reellen Eigenwert besitzt.

Eine Erweiterung dieses Arguments liefert sogar die Existenz einer orthonormalen Basis für A. Ist v der im ersten Schritt gefundene normierte Eigenvektor, so suchen wir im nächsten Schritt ein Maximum auf $\mathbb{S}^n \cap \{v\}^\perp$. Mit anderen Worten, wir betrachten nun die Nebenbedingungen

$$\begin{aligned} g_0(x) &= 1 - \langle x, x \rangle = 0, \\ g_1(x) &= \langle v, x \rangle = 0. \end{aligned}$$

Eine Maximalstelle auf $g^{-1}(0) = \mathbb{S}^n \cap \{v\}^\perp$ tritt dann an einem weiteren Eigenvektor auf, der aufgrund der Konstruktion orthogonal zu v ist. Dabei spielt es keine Rolle, ob der zugehörige Eigenwert vom ersten verschieden ist oder nicht.

Fährt man so fort, so erhält man eine orthogonale Basis des Gesamtraumes, bestehend auf Eigenvektoren des Operators A. ◂