Kapitel 3

Konvexität

3.1 Konvexe Mengen

Der Begriff der konvexen Menge ist bereits aus Definition 1.4, Teil I, bekannt.

Definition 3.1 Konvexer Kegel. Eine Menge $\Omega \subset \mathbb{R}^n$ heißt konvexer Kegel, wenn mit $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ auch $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \in \Omega$ für alle $\lambda_1, \lambda_2 \geq 0$. □

Beispiel 3.2 Für $\lambda_1 = \lambda_2 = 0$ folgt, dass insbesondere $\mathbf{0}$ in einem konvexen Kegel enthalten sein muss.

Nicht jede konvexe Menge ist ein konvexer Kegel, zum Beispiel sind Kreise keine konvexen Kegel.

Der Winkelraum mit Scheitel $\mathbf{0}$ ist ein konvexer Kegel. Der Winkelraum mit Scheitel $\neq \mathbf{0}$ ist kein konvexer Kegel.

Sei $\mathbf{a} \in \mathbb{R}^n, \mathbf{a} \neq \mathbf{0}$, $\alpha \in \mathbb{R}$ und definiere die Hyperebene

$$H(\mathbf{a}, \alpha) = \left\{ \mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{a}^T \mathbf{x} = \alpha \right\}.$$

Diese Hyperebene ist eine konvexe Menge. *Übungsaufgabe* Ist $\alpha \neq 0$, dann ist $\mathbf{0} \notin H(\mathbf{a}, \alpha)$ und $H(\mathbf{a}, \alpha)$ ist kein konvexer Kegel. Im Falle $\alpha = 0$ ist $H(\mathbf{a}, \alpha)$ ein konvexer Kegel $H(\mathbf{a}, \alpha) = \mathbf{a}^\perp$. *Übungsaufgabe*

Analog gilt für Halbräume

$$H_+(\mathbf{a}, \alpha) = \left\{ \mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{a}^T \mathbf{x} \geq \alpha \right\}, \ H_-(\mathbf{a}, \alpha) = \left\{ \mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{a}^T \mathbf{x} \leq \alpha \right\},$$

dass diese für $\alpha = 0$ konvexe Kegel sind und sonst nicht.

Eine Menge, die durch Hyperebenen berandet ist, die allesamt den Punkt $\mathbf{0}$ enthalten, wird auch polyhedrischer Kegel genannt. Polyhedrische Kegel sind konvexe Kegel, insbesondere

$$\mathbb{R}^n_+ = \{ \mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{x} \geq \mathbf{0} \}.$$

□

Definition 3.3 Trennung von Mengen. Die Hyperebene $H(\mathbf{a}, \alpha)$ trennt die nichtleeren Mengen $\Omega_1, \Omega_2 \subset \mathbb{R}^n$, wenn gilt

$$\mathbf{a}^T \mathbf{x} \leq \alpha \ \forall \mathbf{x} \in \Omega_1 \quad \text{und} \quad \mathbf{a}^T \mathbf{x} \geq \alpha \ \forall \mathbf{x} \in \Omega_2.$$

Falls in beiden Relationen die echten Ungleichungen stehen, dann wird die Trennung streng genannt. □

Lemma 3.4 *Sei $\Omega \subset \mathbb{R}^n$ nichtleer, abgeschlossen und konvex und gelte $\mathbf{0} \notin \Omega$. Dann existiert eine Hyperebene $H(\mathbf{a}, \alpha)$ mit $\alpha > 0$, so dass $\mathbf{a}^T \mathbf{x} > \alpha$ für alle $\mathbf{x} \in \Omega$.*

Beweis: Wir werden eine konkrete Hyperebene angeben.

Seien $\mathbf{x}_1 \in \Omega$ und $B_1 = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \leq \|\mathbf{x}_1\|_2\}$ eine Kugel mit Radius $\|\mathbf{x}_1\|_2$ und Mittelpunkt $\mathbf{0}$. Dann ist $\Omega \cap B_1 \neq \emptyset$ und der Durchschnitt ist kompakt (abgeschlossen und beschränkt). Damit nimmt die stetige Funktion $d(\mathbf{x}) = \|\mathbf{x}\|_2$ ihr Minimum über $\Omega \cap B_1$ in einem Punkt $\mathbf{x}_0$ an (Satz von Bolzano–Weierstrass), dass heißt es gilt

$$\|\mathbf{x}\|_2 \geq \|\mathbf{x}_0\|_2 \quad \forall\, \mathbf{x} \in \Omega \cap B_1. \tag{3.1}$$

Wegen $\mathbf{0} \notin \Omega$ ist $\|\mathbf{x}_0\|_2 > 0$. Da die Elemente von Ω, die nicht in $\Omega \cap B_1$ liegen, eine größere Norm als $\|\mathbf{x}_1\|_2 \geq \|\mathbf{x}_0\|_2$ besitzen, gilt (3.1) für alle $\mathbf{x} \in \Omega$.

Wegen der Konvexität von Ω ist $\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0) \in \Omega$ für alle $\mathbf{x} \in \Omega$ und $\lambda \in (0, 1]$ und es folgt

$$\|(1-\lambda)\mathbf{x}_0 + \lambda\mathbf{x}\|_2 = \|\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0)\|_2 \geq \|\mathbf{x}_0\|_2 > 0.$$

Quadratur dieser Ungleichung und Ausschreiben der Skalarprodukte ergibt

$$2\lambda(\mathbf{x} - \mathbf{x}_0)^T \mathbf{x}_0 + \lambda^2 (\mathbf{x} - \mathbf{x}_0)^T (\mathbf{x} - \mathbf{x}_0) \geq 0.$$

Division durch λ und dann $\lambda \to 0$ ergibt

$$\mathbf{x}^T \mathbf{x}_0 \geq \mathbf{x}_0^T \mathbf{x}_0 = \|\mathbf{x}_0\|_2^2 > \frac{\|\mathbf{x}_0\|_2^2}{2} > 0 \quad \forall\, \mathbf{x} \in \Omega.$$

Die Hyperebene mit $\mathbf{a} = \mathbf{x}_0$ und $\alpha = \|\mathbf{x}_0\|_2^2 / 2$ erfüllt nun die Bedingungen des Lemmas. ■

Eine Folgerung ist der nachfolgende Satz.

Satz 3.5 Erster Trennungssatz. *Seien $\Omega_1 \neq \emptyset$ abgeschlossen und konvex und Ω_2 kompakt und konvex, $\Omega_1, \Omega_2 \subset \mathbb{R}^n$, mit $\Omega_1 \cap \Omega_2 = \emptyset$. Dann existiert eine Hyperebene $H(\mathbf{a}, \alpha)$, die Ω_1 und Ω_2 streng trennt.*

Beweis: Seien

$$\begin{aligned} -\Omega_1 &= \{\mathbf{x} : \mathbf{x} = -\mathbf{y}, \mathbf{y} \in \Omega_1\}, \\ \Omega_2 - \Omega_1 &= \{\mathbf{x} : \exists \mathbf{x}_1 \in \Omega_1, \mathbf{x}_2 \in \Omega_2 \text{ mit } \mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1\}. \end{aligned}$$

Diese Mengen sind konvex und abgeschlossen. Da Ω_1 und Ω_2 einen leeren Durchschnitt haben, gilt $\mathbf{0} \notin \Omega_2 - \Omega_1$. Nach Lemma 3.4 gibt es eine Hyperebene $H(\mathbf{a}, \alpha)$ mit $\alpha > 0$ und $\mathbf{a}^T(\mathbf{x}_2 - \mathbf{x}_1) > \alpha > 0$ für alle $\mathbf{x}_1 \in \Omega_1, \mathbf{x}_2 \in \Omega_2$. Durch Umstellen dieser Beziehung erhält man

$$\inf_{\mathbf{x}_2 \in \Omega_2} \mathbf{a}^T \mathbf{x}_2 \geq \sup_{\mathbf{x}_1 \in \Omega_1} \mathbf{a}^T \mathbf{x}_1 + \alpha > \sup_{\mathbf{x}_1 \in \Omega_1} \mathbf{a}^T \mathbf{x}_1.$$

Jede Hyperebene $H(\mathbf{a}, \beta)$ mit

$$\beta \in \left(\sup_{\mathbf{x}_1 \in \Omega_1} \mathbf{a}^T \mathbf{x}_1, \inf_{\mathbf{x}_2 \in \Omega_2} \mathbf{a}^T \mathbf{x}_2 \right)$$

trennt damit Ω_1 und Ω_2 streng. ■

Übungsaufgabe: Notwendigkeit der Kompaktheitsvoraussetzung für strenge Trennung an Beispiel zeigen.

Satz 3.6 *Jede abgeschlossene konvexe Menge $\Omega \subset \mathbb{R}^n$ ist der Durchschnitt aller abgeschlossenen Halbräume, die Ω enthalten.*

Beweis: Beweisskizze. Bezeichnen wir den Durchschnitt aller abgeschlossenen Halbräume, die Ω enthalten, mit D. Dann ist $\Omega \subset D$ klar (Durchschnitt beliebig vieler konvexer Mengen ist konvex). Zu zeigen bleibt $D \subset \Omega$. Das geschieht indirekt, wobei der Widerspruch mit einer trennenden Hyperebene hergeleitet wird (sogenannter zweiter Trennungssatz, Details siehe [ERSD77, Abschnitt 2.1.4]). ■

Der Eckpunkt oder Extrempunkt einer konvexen Menge wurde bereits in Definition 1.10, Teil I, definiert als ein Punkt, für den es keine echte konvexe Linearkombination gibt. Dafür, dass $\mathbf{x}_0 \in \Omega$ ein Extrempunkt ist, reicht es bereits aus, dass man keine $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ findet, so dass

$$\mathbf{x}_0 = \frac{1}{2}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2. \quad \mathbf{x}_1, \mathbf{x}_2 \neq \mathbf{x}_0, \quad \textit{Übungsaufgabe}$$

Satz 3.7 *Jede nichtleere, kompakte, konvexe Menge $\Omega \subset \mathbb{R}^n$ besitzt mindestens einen Extrempunkt.*

Beweis: Da Ω kompakt ist, gibt es ein $\mathbf{x}_0 \in \mathbb{R}^n$ mit $\|\mathbf{x}_0\|_2 \geq \|\mathbf{x}\|_2$ für alle $\mathbf{x} \in \Omega$. Sei

$$\mathbf{x}_0 = \frac{1}{2}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2 = \mathbf{x}_1 + \frac{1}{2}(\mathbf{x}_2 - \mathbf{x}_1) = \mathbf{x}_2 + \frac{1}{2}(\mathbf{x}_1 - \mathbf{x}_2)$$

für $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$, $\mathbf{x}_1, \mathbf{x}_2 \neq \mathbf{x}_0$. Damit folgt

$$\|\mathbf{x}_1\|_2^2 \leq \|\mathbf{x}_0\|_2^2 = \|\mathbf{x}_1\|_2^2 + \underbrace{\frac{1}{4}\|\mathbf{x}_2 - \mathbf{x}_1\|_2^2 + (\mathbf{x}_2 - \mathbf{x}_1)^T \mathbf{x}_1}_{\geq 0, \text{ da } \|\mathbf{x}_0\|_2 \geq \|\mathbf{x}_1\|_2}.$$

Das heißt

$$\|\mathbf{x}_2 - \mathbf{x}_1\|_2^2 \geq -4(\mathbf{x}_2 - \mathbf{x}_1)^T \mathbf{x}_1$$

und analog

$$\|\mathbf{x}_2 - \mathbf{x}_1\|_2^2 \geq -4(\mathbf{x}_1 - \mathbf{x}_2)^T \mathbf{x}_2 = 4(\mathbf{x}_2 - \mathbf{x}_1)^T \mathbf{x}_2.$$

Addition der Ungleichungen ergibt

$$2\|\mathbf{x}_2 - \mathbf{x}_1\|_2^2 \geq 4(\mathbf{x}_2 - \mathbf{x}_1)^T (\mathbf{x}_2 - \mathbf{x}_1) = 4\|\mathbf{x}_2 - \mathbf{x}_1\|_2^2$$

und damit $\mathbf{x}_2 = \mathbf{x}_1 = \mathbf{x}_0$. Somit ist $\mathbf{x}_0$ ein Extrempunkt. ■

3.2 Konvexe und konkave Funktionen

Definition 3.8 Konvexe, konkave Funktion. Die Funktion $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ heißt konvex, wenn für beliebige $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ gilt

$$f((1-\lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2) \leq (1-\lambda)f(\mathbf{x}_1) + \lambda f(\mathbf{x}_2) \quad \forall\, \lambda \in [0,1].$$

Sie heißt konkav, wenn

$$f((1-\lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2) \geq (1-\lambda)f(\mathbf{x}_1) + \lambda f(\mathbf{x}_2) \quad \forall\, \lambda \in [0,1].$$

□

Gilt die echte Ungleichung spricht man von strenger Konvexität (Konkavität).

Beispiel 3.9 Die lineare Funktion $f(\mathbf{x}) = \mathbf{c}^T\mathbf{x}$ mit $\mathbf{c} \in \mathbb{R}^n$ ist konvex und konkav in $\mathbb{R}^n$, jedoch nicht streng. □

Beispiel 3.10 Jede über $\mathbb{R}^n$ positiv semidefinite Bilinearform $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ mit $A = A^T \in \mathbb{R}^{n\times n}$ ist konvex. Sei nämlich

$$\mathbf{x} = (1-\lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2 = \mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1), \quad \mathbf{x}_1 \neq \mathbf{x}_2 \in \mathbb{R}^n,\ \lambda \in [0,1],$$

dann gilt

$$\begin{aligned}
\mathbf{x}^T A\mathbf{x} &= \mathbf{x}_1^T A\mathbf{x}_1 + 2\lambda(\mathbf{x}_2-\mathbf{x}_1)^T A\mathbf{x}_1 + \underbrace{\lambda^2}_{\le\lambda}\underbrace{(\mathbf{x}_2-\mathbf{x}_1)^T A(\mathbf{x}_2-\mathbf{x}_1)}_{\ge 0} \\
&\le \mathbf{x}_1^T A\mathbf{x}_1 + 2\lambda(\mathbf{x}_2-\mathbf{x}_1)^T A\mathbf{x}_1 + \lambda(\mathbf{x}_2-\mathbf{x}_1)^T A(\mathbf{x}_2-\mathbf{x}_1) \\
&= \mathbf{x}_1^T A\mathbf{x}_1 + \lambda(\mathbf{x}_2-\mathbf{x}_1)^T A\,(\mathbf{x}_1+\mathbf{x}_2) \\
&= (1-\lambda)\mathbf{x}_1^T A\mathbf{x}_1 + \lambda\mathbf{x}_2^T A\mathbf{x}_2.
\end{aligned}$$

Analog zeigt man, dass

- A ist positiv definit, dann ist $f(\mathbf{x})$ streng konvex,
- A ist negativ semidefinit, dann ist $f(\mathbf{x})$ konkav,
- A ist negativ definit, dann ist $f(\mathbf{x})$ streng konkav.

□

Beispiel 3.11 Sind die Funktionen $f_j(\mathbf{x}), j = 1,\ldots,k$, über $\Omega \in \mathbb{R}^n$ konvex, dann ist auch die Linearkombination

$$f(\mathbf{x}) = \sum_{j=1}^{k} \alpha_j f_j(\mathbf{x}), \quad \alpha_j \ge 0\ \forall\ j$$

konvex. □

Zur Stetigkeit von konvexen Funktionen gibt es folgende Aussage.

Satz 3.12 *Seien $\Omega \subset \mathbb{R}^n$ konvex und das Innere der Menge,* $\mathrm{int}(\Omega)$*, nichtleer. Dann ist jede konvexe Funktion $f : \Omega \to \mathbb{R}$ stetig in* $\mathrm{int}(\Omega)$.

Beweis: Siehe Literatur, zum Beispiel [ERSD77, Satz 2.65]. Man prüft die ε–δ–Definition nach. ■

Beispiel 3.13 *Nichtstetige konvexe Funktion über konvexer Menge.* Stetigkeit bis auf den Rand muss bei einer konvexen Funktion im allgemeinen nicht vorliegen. Die konvexe Funktion $f\ :\ [1,2] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 1/x & \text{für } x \in [1,2) \\ 2 & \text{für } x = 2 \end{cases}$$

ist ein Beispiel. □

Jetzt werden einige Bedingung dafür gegeben, dass ein Funktion auf einer konvexen Menge konvex ist. Zunächst wird eine Monotonieaussage für den Differenzenquotienten gegeben.

Lemma 3.14 *Seien $\Omega \subset \mathbb{R}^n$ eine offene konvexe Menge, $f : \Omega \to \mathbb{R}$ mit $f \in C^1(\Omega)$ eine konvexe Funktion, $\mathbf{x}_0 \in \Omega$, $\mathbf{h} \in \mathbb{R}^n$, $\mathbf{x}_0 - \nu\mathbf{h} \in \Omega$ und $\mathbf{x}_0 + \rho\mathbf{h} \in \Omega$ für gewisse $\nu, \rho > 0$. Dann gilt*

$$\begin{aligned}
\frac{f(\mathbf{x}_0) - f(\mathbf{x}_0 - \nu\mathbf{h})}{\nu} &\le \frac{f(\mathbf{x}_0) - f(\mathbf{x}_0 - \mu\mathbf{h})}{\mu} \\
&\le \frac{f(\mathbf{x}_0 + \lambda\mathbf{h}) - f(\mathbf{x}_0)}{\lambda} \le \frac{f(\mathbf{x}_0 + \rho\mathbf{h}) - f(\mathbf{x}_0)}{\rho}
\end{aligned}$$

für alle $\mu \in (0,\nu], \lambda \in (0,\rho]$.

Beweis: *Übungsaufgabe* ∎

Damit ist für konvexe Funktionen der Differenzenquotient im Punkt $\mathbf{x}_0$ in Richtung $\mathbf{h}$ eine monoton wachsende Funktion von λ.

Satz 3.15 *Sei $\Omega \subset \mathbb{R}^n$ eine offene konvexe Menge und $f : \Omega \to \mathbb{R}$ mit $f \in C^1(\Omega)$. Dann ist f auf Ω genau dann konvex, wenn für alle $\mathbf{x}_0, \mathbf{x}_1 \in \Omega$ mit $\mathbf{x}_0 \neq \mathbf{x}_1$ gilt*

$$f(\mathbf{x}_1) - f(\mathbf{x}_0) \geq (\mathbf{x}_1 - \mathbf{x}_0)^T \nabla f(\mathbf{x}_0). \tag{3.2}$$

Beweis: 1) *Notwendigkeit von (3.2).* Da f differenzierbar ist, gilt für die Richtungsableitung in Richtung $\mathbf{h} \in \mathbb{R}^n$ mit $\|\mathbf{h}\|_2 = 1$

$$\frac{\partial f}{\partial \mathbf{h}}(\mathbf{x}_0) = \mathbf{h}^T \nabla f(\mathbf{x}_0) = \lim_{\lambda \to 0} \frac{f(\mathbf{x}_0 + \lambda \mathbf{h}) - f(\mathbf{x}_0)}{\lambda} = - \lim_{\lambda \to 0} \frac{f(\mathbf{x}_0 - \lambda \mathbf{h}) - f(\mathbf{x}_0)}{\lambda},$$

wobei bei der letzten Gleichung λ durch $-\lambda$ ersetzt wurde. Nach Lemma 3.14 gilt alle $\rho > 0$ mit $\mathbf{x}_0 + \rho\mathbf{h} \in \Omega$ (nehme $\lambda \to 0$ in Lemma 3.14)

$$\frac{f(\mathbf{x}_0 + \rho \mathbf{h}) - f(\mathbf{x}_0)}{\rho} \geq \mathbf{h}^T \nabla f(\mathbf{x}_0)$$

oder

$$f(\mathbf{x}_0 + \rho \mathbf{h}) \geq f(\mathbf{x}_0) + \rho \mathbf{h}^T \nabla f(\mathbf{x}_0).$$

Wählt man $\mathbf{x}_1 = \mathbf{x}_0 + \rho\mathbf{h}$, so ergibt sich (3.2).

2) *Hinlänglichkeit von (3.2).* Sei $\lambda_0 > 0$ beliebig gewählt, setzte $\lambda_1 = 1 - \lambda_0$. Dann gilt für $\mathbf{x}_2 = \lambda_0 \mathbf{x}_0 + \lambda_1 \mathbf{x}_1$ nach Voraussetzung

$$f(\mathbf{x}_0) - f(\mathbf{x}_2) \geq (\mathbf{x}_0 - \mathbf{x}_2)^T \nabla f(\mathbf{x}_2), \quad f(\mathbf{x}_1) - f(\mathbf{x}_2) \geq (\mathbf{x}_1 - \mathbf{x}_2)^T \nabla f(\mathbf{x}_2).$$

Damit folgt

$$\begin{aligned} \lambda_0 f(\mathbf{x}_0) + \lambda_1 f(\mathbf{x}_1) &\geq f(\mathbf{x}_2) + \Big(\lambda_0(\mathbf{x}_0 - \mathbf{x}_2) + \lambda_1(\mathbf{x}_1 - \mathbf{x}_2)\Big)^T \nabla f(\mathbf{x}_2) \\ &= f(\mathbf{x}_2) + \underbrace{\Big(\lambda_0 \mathbf{x}_0 + \lambda_1 \mathbf{x}_1 - \mathbf{x}_2\Big)^T}_{=\mathbf{0}} \nabla f(\mathbf{x}_2) \\ &= f(\mathbf{x}_2) = f(\lambda_0 \mathbf{x}_0 + \lambda_1 \mathbf{x}_1) \end{aligned}$$

für alle $\mathbf{x}_0, \mathbf{x}_1 \in \Omega$. ∎

Für eine zweimal stetig differenzierbare Funktion $f : \Omega \to \mathbb{R}, \Omega \subset \mathbb{R}^n$, offen, ist die Hesse–Matrix

$$H_f(\mathbf{x}) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1,\dots,n} (\mathbf{x})$$

symmetrisch (Satz von Schwarz).

Liegt zusammen mit $\mathbf{x}_0, \mathbf{x}_1$ die Strecke $[\mathbf{x}_0, \mathbf{x}_1]$ vollständig in Ω, so gilt für eine geeignete Zahl $\lambda \in (0, 1)$ nach dem Satz von Taylor

$$f(\mathbf{x}_1) = f(\mathbf{x}_0) + (\mathbf{x}_1 - \mathbf{x}_0)^T \nabla f(\mathbf{x}_0) + \frac{1}{2}(\mathbf{x}_1 - \mathbf{x}_0)^T H_f(\mathbf{x}_0 + \lambda(\mathbf{x}_1 - \mathbf{x}_0))(\mathbf{x}_1 - \mathbf{x}_0). \tag{3.3}$$

Satz 3.16 *Seien $\Omega \subset \mathbb{R}^n$ eine offene konvexe Menge und $f : \Omega \to \mathbb{R}$ mit $f \in C^2(\Omega)$. Dann ist f auf Ω genau dann konvex, wenn $H_f(\mathbf{x})$ positiv semidefinit ist für alle $\mathbf{x} \in \Omega$.*

Beweis: 1.) Sei $H_f(\mathbf{x})$ positiv semidefinit. Dann folgt aus (3.3) für beliebige $\mathbf{x}_0, \mathbf{x}_1 \in \Omega$ und eine geeignete Zahl $\lambda \in (0,1)$

$$f(\mathbf{x}_1) - f(\mathbf{x}_0) - (\mathbf{x}_1 - \mathbf{x}_0)^T \nabla f(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_1 - \mathbf{x}_0)^T H_f(\mathbf{x}_0 + \lambda(\mathbf{x}_1 - \mathbf{x}_0))(\mathbf{x}_1 - \mathbf{x}_0) \geq 0.$$

Aus Satz 3.15 folgt, dass $f(\mathbf{x})$ konvex ist.

2) Sei $f(\mathbf{x})$ konvex auf Ω. Seien $\mathbf{x}_0 \in \Omega$ und $\mathbf{h} \in \mathbb{R}^n$ beliebig gegeben. Da Ω offen ist, gibt es eine Zahl $\theta_0 > 0$ so dass $\mathbf{x}_0 + \theta \mathbf{h} \in \Omega$ für alle $\theta \in (0, \theta_0]$. Wählt man in (3.2) $\mathbf{x}_1 = \mathbf{x}_0 + \theta \mathbf{h}$, erhält man

$$f(\mathbf{x}_0 + \theta \mathbf{h}) - f(\mathbf{x}_0) - \theta \mathbf{h}^T \nabla f(\mathbf{x}_0) \geq 0 \quad \forall\, \theta \in (0, \theta_0].$$

Aus (3.3) folgt nun

$$\mathbf{h}^T H_f(\mathbf{x}_0 + \lambda \theta \mathbf{h}) \mathbf{h} \geq 0 \quad \forall\, \theta \in (0, \theta_0]$$

mit $\lambda \in (0,1)$. Da die Hesse–Matrix nach Voraussetzung stetig ist, folgt für $\theta \to +0$

$$\mathbf{h}^T H_f(\mathbf{x}_0) \mathbf{h} \geq 0.$$

Da $\mathbf{x}_0 \in \Omega$ und $\mathbf{h} \in \mathbb{R}^n$ beliebig gewählt wurden, ist $H_f(\mathbf{x})$ in jedem Punkt $\mathbf{x} \in \Omega$ positiv semidefinit. ■

Eine skalare Funktion einer skalaren Veränderlichen $f : (a,b) \to \mathbb{R}$ mit $f \in C^2(a,b)$ ist konvex, wenn $f''(x) \geq 0$ in (a,b), zum Beispiel die Funktion $f(x) = x^2$.

Bemerkung 3.17 Es gibt Verallgemeinerungen des Begriffs der konvexen Funktion, zum Beispiel quasi–konvex und explizit konvex, siehe Literatur. □

3.3 Ungleichungen und konvexe Mengen

Satz 3.18 *Die Funktion $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ sei konvex, Ω sei konvex und sei $b \in \mathbb{R}$. Dann ist*

$$V = \{\mathbf{x} \in \Omega \ : \ f(\mathbf{x}) \leq b\}$$

konvex.

Beweis: Zu zeigen ist: $\mathbf{x}_1, \mathbf{x}_2 \in V \implies \mathbf{x} = (1-\lambda)\mathbf{x}_1 + \lambda \mathbf{x}_2 \in V$ für alle $\lambda \in [0,1]$. Aus $f(\mathbf{x}_1), f(\mathbf{x}_2) \leq b$ folgt aus der Konvexität von f

$$f(\mathbf{x}) \leq (1-\lambda) f(\mathbf{x}_1) + \lambda f(\mathbf{x}_2) \leq b.$$

Damit ist $\mathbf{x} \in V$. ■

Bemerkung 3.19 1.) Die Menge V aus Satz 3.18 heißt Unterhalbmenge von $f(\mathbf{x})$ in Ω zum Wert b.

2.) Die Niveaumenge $\{\mathbf{x} \in \Omega \ : \ f(\mathbf{x}) = b\}$ der konvexen Funktion $f(\mathbf{x})$ ist im allgemeinen nicht konvex. Zum Beispiel besteht die Niveaumenge zu $b = 1$ der konvexen Funktion $f(x) = x^2, x \in \mathbb{R}$, aus den Punkten -1 und 1.

Folgerung 3.20 *Sei $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ konkav, sei Ω konvex und sei $b \in \mathbb{R}$. Dann ist $U = \{\mathbf{x} \in \Omega \ : \ f(\mathbf{x}) \geq b\}$ konvex.*

Beweis: Nutze Satz 3.18 mit $-f(\mathbf{x})$, da diese Funktion konvex ist. ■

3.4 Extrema von konvexen Funktionen

Es wird die Frage untersucht, unter welchen Bedingungen eine konvexe Funktion ihr globales Minimum über einer gegebenen konvexen Menge Ω annimmt.

Satz 3.21 *Sei $f : \Omega \to \mathbb{R}$ eine konvexe Funktion über einer konvexen Menge Ω. Dann ist ein lokales Minimum von $f(\mathbf{x})$ in Ω zugleich das globale Minimum über Ω.*

Beweis: Indirekt. Nehme $f(\mathbf{x})$ ein lokales Minimum in $\mathbf{x}_0 \in \Omega$ an. Sei weiter $\mathbf{x}^* \in \Omega$ mit $f(\mathbf{x}^*) < f(\mathbf{x}_0)$. Aus der Konvexität von $f(\mathbf{x})$ folgt

$$f((1-\lambda)\mathbf{x}_0 + \lambda\mathbf{x}^*) \leq (1-\lambda)f(\mathbf{x}_0) + \lambda f(\mathbf{x}^*) = f(\mathbf{x}_0) + \lambda\left(f(\mathbf{x}^*) - f(\mathbf{x}_0)\right) < f(\mathbf{x}_0)$$

für alle $\lambda \in (0,1)$, da $f(\mathbf{x}^*) < f(\mathbf{x}_0)$. Aus der Konvexität von Ω folgt, dass der Funktionswert im Zwischenpunkt wohldefiniert ist. Wählt man $\lambda < \varepsilon$ so ist der Zwischenpunkt beliebig nahe an $\mathbf{x}_0$ und die obige Ungleichung steht im Widerspruch dazu, dass $f(\mathbf{x})$ in $\mathbf{x}_0$ ein lokales Minimum annimmt. ■

Beispiel 3.22 *Konvexe Funktion über konvexer Menge ohne Annahme eines lokalen Minimums.* Der obige Satz kann nur angewendet werden, wenn es ein lokales Minimum gibt, das angenommen wird. Das muss nicht der Fall sein. Die konvexe Funktion aus Beispiel 3.13 nimmt ihr Minimum nicht an. □

Folgerung 3.23 *Die Menge Ω_0, über der eine konvexe Funktion $f : \Omega \to \mathbb{R}$ über einer konvexen Menge Ω das globale Minimum annimmt, ist konvex.*

Beweis: Folgt aus Satz 3.18. ■

Satz 3.24 *Sei $f(\mathbf{x}) \in C^1(\Omega)$ über einer konvexen Menge Ω konvex. Dann nimmt $f(\mathbf{x})$ in jedem $\mathbf{x}_1 \in \Omega$ mit $\nabla f(\mathbf{x}_1) = \mathbf{0}$ sein globales Minimum über Ω an.*

Beweis: Es ist zu zeigen, dass alle Punkte $\mathbf{x}_1 \in \Omega$ mit $\nabla f(\mathbf{x}_1) = \mathbf{0}$ wirklich Extremalpunkte sind und zudem Minima.

Sei $\mathbf{x}_1 \in \Omega$ mit $\nabla f(\mathbf{x}_1) = \mathbf{0}$. Da $f(\mathbf{x})$ konvex ist, gilt

$$f\left(\lambda\mathbf{x}_2 + (1-\lambda)\mathbf{x}_1\right) \leq (1-\lambda)f(\mathbf{x}_1) + \lambda f(\mathbf{x}_2) = f(\mathbf{x}_1) + \lambda\left(f(\mathbf{x}_2) - f(\mathbf{x}_1)\right)$$

für alle $\mathbf{x}_2 \in \Omega$ und $\lambda \in [0,1]$. Für $\lambda \in (0,1]$ erhält man daraus

$$\frac{f\left(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\right) - f(\mathbf{x}_1)}{\lambda} \leq f(\mathbf{x}_2) - f(\mathbf{x}_1).$$

Nach dem Mittelwertsatz der Differentialrechnung gibt es ein $\theta \in [0,1]$ mit

$$(\mathbf{x}_2 - \mathbf{x}_1)^T \nabla f\left(\mathbf{x}_1 + \theta\lambda(\mathbf{x}_2 - \mathbf{x}_1)\right) \leq f(\mathbf{x}_2) - f(\mathbf{x}_1).$$

Für $\lambda \to 0$ folgt nun mit $\nabla f(\mathbf{x}_1) = \mathbf{0}$

$$0 = (\mathbf{x}_2 - \mathbf{x}_1)^T \nabla f\left(\mathbf{x}_1\right) \leq f(\mathbf{x}_2) - f(\mathbf{x}_1),$$

also $f(\mathbf{x}_1) \leq f(\mathbf{x}_2)$ für alle $\mathbf{x}_2 \in \Omega$. ■