TECHNISCHE UNIVERSITÄT DRESDEN

Fakultät Mathematik Institut für Numerik, Professur für Numerik der Optimalen Steuerung

OPTIMIERUNG UND NUMERIK

Dr. John Martinovic

Wintersemester 2019/20

Autor : Eric Kunze
E-Mail : eric.kunze@mailbox.tu-dresden.de

Inhaltsverzeichnis

Kapitel 1
EINFÜHRUNG

1.1 Aufgabenstellung und Grundbegriffe

Es seien $G \subseteq \mathbb{R}^n$ und $f\colon G \to \mathbb{R}$ gegeben. In dieser Vorlesung betrachten wir Optimierungsaufgaben (OA) der Form

$$f(x) \to \min \qquad \text{bei } x \in G \tag{1.1}$$

Man nennt

- f die **Zielfunktion**,
- G den **zulässigen Bereich** und
- ein $x \in G$ **zulässigen Punkt** (oder zulässige Lösung).

Ein zulässiger Punkt $x^* \in G$ heißt **optimal** (oder Lösung oder optimale Lösung), wenn für alle $x \in G$ die Ungleichung

$$f(x^*) \leq f(x) \tag{1.2}$$

gilt. Falls das Problem (1.1) lösbar ist, so wird mit $f^* = f(x^*)$ der **Optimalwert** bezeichnet. Das Problem (1.1) ist ein

- **unrestringiertes** (oder freies) Optimierungsproblem, wenn $G = \mathbb{R}^n$ gilt,
- andernfalls (d.h. für $G \neq \mathbb{R}^n$) ein **restringiertes** Problem

und außerdem eine

- **diskrete** (oder ganzzahlige) OA (engl. integer program), falls jede Variable eine diskreten Menge angehört
- **kontinuierliche** (oder stetige) OA, falls alle Variablen stetige Werte annehmen
- gemischt ganzzahlige OA, wenn sowohl stetige als auch diskrete Variablen vorkommen.

Gilt in (1.1) $f(x) = c^\top x$ für ein $c \in \mathbb{R}^n$ und ist G durch lineare Bedingungen beschreibbar, so heißt (1.1) **linear**. In diesem Fall lässt sich (1.1) schreiben als

$$c^\top x \to \min \qquad \text{bei } Ax = a, Bx \leq b \tag{1.3}$$

mit geeigneten Matrizen A und B sowie Vektoren a und b.

Gerade für (gemischt) ganzzahlige OA kann die Lösung der Originalaufgabe schwierig sein. Eine verwandte, jedoch im Allgemeinen leichter zu lösende Aufgabe kann in diesen Fällen wie folgt erhalten werden:

Definition 1.1
Wir betrachten die Optimierungsaufgaben
(P) $f(x) \to \min$ bei $x \in D \cap E$
(Q) $g(x) \to \min$ bei $x \in E$
(Q) heißt **Relaxation** zu (P) falls $g(x) \leq f(x)$ für alle $x \in D \cap E$ gilt. In vielen Fällen wird dabei $g = f$ gewählt.

Der Optimalwert der Relaxation kann als Näherung (bzw. untere Schranke) für den tatsächlichen Optimalwert von (P) genutzt werden. Meistens liefert die Lösung von (Q) jedoch keinen zulässigen Puntk für (P).

Satz 1.1
Ist $\overline{x}$ eine Lösung von (Q) und gilt $\overline{x} \in D$ sowie $f(\overline{x}) = g(\overline{x})$, dann löst $\overline{x}$ auch (P).

Beweis. siehe Übung □

Definition 1.2
Seien (Q1) und (Q2) Relaxationen zu (P). (Q1) heißt **stärker** (oder strenger) als (Q2), wenn die Schranke (d.h. der Optimalwert) von (Q1) größer oder gleich der Schranke (Optimalwert) von (Q2) für jede Instanz von (P) ist.

Anmerkung. *Zur Erklärung des Begriffes "Instanz" betrachte das folgende Beispiel.*
- *Problemklasse:* $c^\top x \to \min$
- *Instanz der Problemklasse:* $x_1 + 2x_2 - 3x_3 \to \min$

Eine Instanz ist also eine konkrete Belegung.

1.2 Beispiele zur kontinuierlichen Optimierung

1.2.1 Transportoptimierung

$\to$ lineare Optimierung

Es gebe Erzeuger $i \in I = \{0, \dots, n\}$ und Verbraucher $j \in J = \{1, \dots, n\}$. Weiterhin seien die Kosten c_{ij} für den Transport einer Einheit von i nach j sowie der Vorrat $a_i > 0$ und der Bedarf $b_j > 0$ für alle i und j gegeben. Wie muss der Transport organisiert werden, damit die Gesamtkosten minimal sind?

Für jedes mathematische Modell einer OA braucht man
- geeignete Variablen ($\to x$)
- Zielfuntkion ($\to f$)
- Nebenbedingungen ($\to G$)

Variablen $x_{ij} \geq 0$ für alle $i \in I$ und $j \in J$ beschreibe die Einheiten, die von i nach j transportiert werden.

Zielfunktion $f(x) = \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} \to \min$

Nebenbedingungen

- Kapazitätsbeschränkung der Erzeuger $i \in I$: $\sum_{j\in J} x_{ij} \leq a_i \quad (i \in I)$
- Bedarfserfüllung von Verbrauchern $j \in J$: $\sum_{i\in I} x_{ij} \geq b_j \quad (j \in J)$

Somit können wir als Modell formulieren:

$$f(x) = \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} \to \min \qquad \text{bei } \begin{aligned} &\sum_{j\in J} x_{ij} \leq a_i \;\; (i \in I), \\ &\sum_{i\in I} x_{ij} \geq b_j \;\; (j \in J), \\ &x_{ij} \geq 0 \;\; ((i,j) \in I \times J) \end{aligned}$$

1.2.2 Kürzeste euklidische Entfernung

$\to$ nichtlineare Optimierung

Gegeben seien ein Punkt $\widetilde{x} \in \mathbb{R}^n$ und eine Menge $G \subseteq \mathbb{R}^n$ mit $x \notin G$. Wir betrachten die folgende OA:

$$f(x) = \|x - \widetilde{x}\|_2^2 \to \min \qquad \text{bei } x \in G$$

Ist $G \neq \emptyset$ und abgeschlossen, so existiert eine Lösung. Ist G zusätzlich konvex, so ist die Lösung sogar eindeutig.

Weitere Beispiele und Theorie sind in der Vorlesung "Kontinuierliche Optimierung“ im Master Mathematik zu erfahren.

1.3 Beispiele zur diskreten Optimierung

1.3.1 Das Rucksackproblem

Gegeben seien ein Behälter ("Rucksack“) mit Kapazität $b \in \mathbb{Z}_+ := \{0, 1, \ldots\}$ sowie m Teile, die jeweils durch ein Gewicht $a_i \in \mathbb{Z}_+$ und einen Nutzen $c_i \in \mathbb{Z}_+$ beschrieben werden ($i = 1, \ldots, m$). Aus dieser Menge von Objekten ist eine nutzenmaximale Teilmenge auszuwählen.

Variablen

$$x_i := \begin{cases} 1 & \text{wenn Teil } i \text{ eingepackt wird} \\ 0 & \text{sonst} \end{cases} \qquad (i = 1, \ldots, m)$$

Zielfunktion $f(x) = \sum_{i=1}^{m} c_i x_i \to \max$

Nebenbedingungen Kapazitätsbedingung: $\sum_{i=1}^{m} a_i x_i \leq b$

Als Modell können wir somit formulieren:

$$f(x) = \sum_{i=1}^{m} c_i x_i \to \max \qquad \text{bei } \sum_{i=1}^{m} a_i x_i \leq b \text{ und } x_i \in \{0,1\} \quad (i = 1, \ldots, m)$$

Aufgrund der binären Gestalt der Variablen wird das Problem auch als 0/1-Rucksackproblem bezeichnet. Im Gegensatz dazu ist beim klassischen Rucksackproblem jedes Teil mehrfach nutzbar. In diesem Fall ist $x_i \in \mathbb{Z}_+$ zu fordern.

1.3.2 Das Bin-Packing-Problem

Gegeben seien (sehr große) Anzahl an Behältern der Kapazität L sowie b_i Teile des Gewichts oder Volumens ℓ_i mit $i \in I = \{1, \ldots, m\}$. Man ermittle die minimale Anzahl an Behältern, die benötigt wird, um alle Objekte zu verstauen. Jede Packung (eines Behälters) kann als Vektor $a = (a_1, \ldots, a_m) \in \mathbb{Z}_+^m$ geschrieben werden, wobei a_i angibt, wie oft das Teil i benutzt wird. Ein solcher Vektor ist eine zulässige Packung, wenn

$$\sum_{i=1}^{m} \ell_i a_i \leq L$$

ist.

Modell nach Kantorovich Wir benötigen

- eine obere Schranke $u \in \mathbb{Z}_+$ für die maximal benötigte Anzahl an Behältern
- $y_k = \begin{cases} 1 & \text{wenn Rucksack } k \text{ benutzt wird} \\ 0 & \text{sonst} \end{cases} \qquad (k = 1, \ldots, u)$
- $x_{ik} \in \mathbb{Z}_+$, die angeben, wieviele Objekte vom Typ i in Rucksack k gepackt werden $((i,k) \in \{1, \ldots, m\} \times \{1, \ldots, u\})$

Daraus ergibt sich nun folgendes Modell:

$$\begin{aligned} f^{\text{Kant}}(x,y) = \sum_{k=1}^{u} y_k \to \min \text{ bei } \quad & \sum_{k=1}^{u} x_{ik} = b_i && (i = 1, \ldots, m) \\ & \sum_{i=1}^{m} x_{ik} \ell_i \leq L \cdot y_k && (k = 1, \ldots, u) \\ & y_k \in \{0,1\} && (k = 1, \ldots, u) \\ & x_{ik} \in \mathbb{Z}_+ && ((i,k) \in \{1, \ldots, m\} \times \{1, \ldots, u\}) \end{aligned}$$

Die erste Nebenbedingung sorgt dafür, dass alle Teile gepackt werden; die zweite Nebenbedingung liefert die Einhaltung der Kapazität unter Berücksichtigung, dass nur bepackte Behälter gezählt werden.

Es kann stets $u = \sum_{i=1}^{m} b_i$ gewählt werden. Das Auffinden besserer Schranken ist im Allgemeinen schwierig. Eine Relaxation kann z.B. durch $y_k \in [0,1]$ und $x_{ik} \in \mathbb{R}_+$ erhalten werden. Diese liefert jedoch keine guten Näherungen.

Modell von Gilmore & Gomory Es seien J eine Indexmenge aller zulässigen Packungen und $x_j \in \mathbb{Z}_+$ $(j \in J)$ die Häufigkeit, wie oft ein Behälter nach dem durch j angegebenen Schema $a^j = (a_1^j, \dots, a_m^j)$ mit $\ell^\top a^j \leq L$ gefüllt wird. Daraus ergibt sich folgendes Modell:

$$f^{GG}(x) = \sum_{j \in J} x_j \to \min \quad \text{bei} \quad \begin{aligned} \sum_{j \in J} a_i^j \cdot x_j &= b_i \qquad (i = 1, \dots, m) \\ x_j &\in \mathbb{Z}_+ \qquad (j \in J) \end{aligned}$$

Die Nebenbedingung sorgt dafür, dass alle Teile gepackt werden.

Es gibt im Allgemeinen exponentiell viele zulässige Packungen a^j $(j \in J)$, deren Koeffizienten allesamt in den Nebenbedingungen benötigt werden.

Eine Relaxation erhält man zum Beispiel durch $x_j \in \mathbb{R}_+$. Diese stetige Relaxation ist sehr gut; man vermutet, dass folgende Bedingung gilt:

$$f^{GG,*} - f^{GG,*}_{\text{relax}} < 2$$

Erfreulicherweise gibt es zum Gilmore-Gomory-Modell äquivalente Formulierungen, die mit einer polynomiellen Zahl von Variablen arbeiten und eine ebenso gute stetige Relaxation besitzen (z.B. Flussmodelle).

1.3.3 Standortplanung

Ein Dienstleister möchte neue Filialen aufbauen, um seine Kunden $k \in K := \{1, \dots, m\}$ zu versorgen. Dabei sind aus der Menge $S := \{1, \dots, n\}$ mögliche Standorte, die neuen Standorte so auszuwählen, dass der Bedarf aller Kunden befriedigt wird und die Gesamtkosten minimial sind.

Wir benötigen

- $c_s > 0$ ... Fixkosten für den Aufbau von Standort $s \in S$
- $d_{ks} > 0$ ... Kosten, um den Kunden $k \in K$ (vollständig) von Standort $s \in S$ zu beliefern.

Variablen:

- $x_s = \begin{cases} 1 & \text{wenn Standort } s \in S \text{ gebaut wird} \\ 0 & \text{sonst} \end{cases}$
- $y_{ks} \geq 0$ ... Anteil des Bedarfs des Kunden $k \in K$, der vom Standort $s \in S$ bedient wird (implizit: $y_{ks} \in [0, 1]$)

Modell zur Standortplanung:

$$f(x, y) = \underbrace{\sum_{s \in S} x_s c_s}_{\text{Fixkosten}} + \underbrace{\sum_{s \in S} \sum_{k \in K} y_{ks} d_{ks}}_{\text{variable Kosten}} \to \min$$

bei

$$\begin{aligned} \sum_{s\in S} y_{ks} &= 1 \qquad && (k \in K) \\ y_{ks} &\le x_s && (s \in S, k \in K) \\ x_s &\in \{0,1\} && (s \in S) \\ y_{ks} &\ge 0 && (k \in K, s \in S) \end{aligned}$$

1.3.4 Quadratisches Zuordnungsproblem

Es sollen n Personen auf n Räume verteilt werden. Person i muss Person j c_{ij} mal am Tag treffen. Außerdem habe Büro k von Büro ℓ die Entfernung $d_{k\ell} > 0$. Wird Person i das Büro k zugewiesen und Person j das Büro ℓ, so ergibt sich eine Gesamtwegstrecke von $2c_{ij}d_{kl}$ (beachte Hin- und Rückweg). Gesucht ist die wegminimale Belegung der Büros.

Variablen: $x_{ik} = \begin{cases} 1 & \text{wenn Person } i \text{ das Büro } k \text{ zugewiesen wird} \\ 0 & \text{sonst} \end{cases} \qquad (i,k) \in \{1,\dots,n\}\times\{1,\dots,n\}$

Zielfunktion:

$$f(x) = \sum_{i=1}^{n} \sum_{k=1}^{n} \sum_{\substack{j=1 \\ j\ne i}}^{n} \sum_{\substack{\ell=1 \\ \ell\ne k}}^{n} x_{j\ell} x_{ik} \cdot 2c_{ij} d_{kl} \to \min$$

bei

$$\begin{aligned} \sum_{i=1}^{n} x_{ik} &= 1 \qquad && (k = 1,\dots,n) && \text{Büro } k \text{ bekommt genau einen Einwohner} \\ \sum_{k=1}^{n} x_{ik} &= 1 && (i = 1,\dots,n) && \text{Person } i \text{ bekommt genau ein Büro} \\ x_{ik} &\in \{0,1\} && (i,k) \in \{1,\dots,n\} \times \{1,\dots,n\} \end{aligned}$$

Weitere Beispiele sind in der Vorlesung ”Diskrete Optimierung“ (Master Mathe) zu finden..

Kapitel 2

GRUNDLAGEN

2.1 Existenz von Lösungen

Wir betrachten die Optimierungsaufgabe

$$f(x) \to \min \qquad \text{bei } x \in G \tag{2.1}$$

wobei folgende Bedingungen erfüllt seinen:

- f ist stetig (zumindest auf G)
- G ist kompakt
- $G \neq \emptyset$

Satz 2.1 (Weierstrass)
Unter diesen Voraussetzungen existiert ein $\overline{x} \in G$ mit

$$f^* := f(\overline{x}) \leq f(x) \quad \forall x \in G$$

Beweis. Sei $f^* := \inf_{x \in G} f(x)$. Wegen $G \neq \emptyset$, finden wir eine Folge $\{f_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}$ mit $f_k = f(x_k) \geq f^*$ mit $x_k \in G$ für alle $k \in \mathbb{N}$ und $\lim_{k\to\infty} f_k = f^*$. Die daraus resultierende Folge $\{x_k\}_{k\in\mathbb{N}}$ besitzt wegen der Kompaktheit von G eine konvergente Teilfolge $\{\widetilde{x_k}\}_{k\in\mathbb{N}} \subseteq \{x_k\}_{k\in\mathbb{N}}$ mit $\lim_{k\to\infty} \widetilde{x_k} = \overline{x} \in G$ (Abgeschlossenheit von G). Die Stetigkeit von f ergibt nun $\lim_{k\to\infty} f(\widetilde{x_k}) = f(\overline{x}) = f^*$ (insbesondere $f^* \in \mathbb{R}$) $\square$

Beispiel 2.1

(1) Satz 2.1 anwendbar (G kompakt, Minimum existiert):

$$f(x_1, x_2) = x_1 - x_2 \to \min \text{ bei } x_1^2 + 4x_2^2 \leq 1$$

Der zulässige Bereich ist eine Ellipse mit Rand.

(2) Satz 2.1 nicht anwendbar (G unbeschränkt, kein Minimum, $f^* = -\infty$):

$$f(x_1, x_2) = x_1 - x_2 \to \min \text{ bei } x_1^2 + 4x_2^2 \geq 1$$

(3) Satz 2.1 nicht anwendbar (G unbeschränkt, kein Minimum, $f^* = 0$)

$$f(x_1, x_2) = \frac{1}{x_1} \to \min \text{ bei } x_2 \leq \frac{1}{x_1}, x_1 \geq 1, x_2 \geq 0$$

(4) Satz 2.1 nicht anwendbar (G unbeschränkt, Minimum existiert, $f^* = -1$)

$$f(x_1, x_2) = -\frac{1}{x_1} \to \min \text{ bei } x_2 \leq \frac{1}{x_1}, x_1 \geq 1, x_2 \geq 0$$

Offensichtlich besitzen also nicht alle Optimierungsaufgaben eine (globale) Lösung, insbesondere deshalb, weil Bedingung (**??**) ziemlich stark ist. Stattdessen hat sich in der Literatur auch der folgende ”schwächere“ Lösungsbegriff etabliert.

Definition 2.1

Ein zulässiger Punkt $\overline{x} \in G$ heißt lokale Lösung von (2.1), falls ein $\rho > 0$ existiert mit

$$f(\overline{x}) \leq f(x) \quad \forall x \in G \cap B_\rho(\overline{x})$$

wobei $B_\rho(\overline{x}) := \{x \in \mathbb{R}^n : \|x - \overline{x}\|_2 \leq \rho\}$ die offene Kugel vom Radius ρ um $\overline{x}$ ist.

Bemerkung 2.1

Jede globale Lösung ist auch lokale Lösung. Die Umkehrung ist im Allgemeinen nicht korrekt.

Sofern eine globale Lösung existiert, ist diese in der Menge der lokalen Lösungen enthalten. Die Betrachtung lokaler Lösungen ist damit im Allgemeinen ausreichend. Für eine spezielle Klasse von Optimierungsaufgaben sind beide Lösungskonzepte sogar äquivalent. Dazu betrachten wir die folgenden Definitionen:

Definition 2.2 (Konvexität)

(1) $G \subseteq \mathbb{R}^n$ ist konvex, falls für alle $x, y \in G$ gilt

$$[x, y] := \{x(\lambda) \in \mathbb{R}^n : x(\lambda) = (1 - \lambda)x + \lambda y, \lambda \in [0, 1]\} \subseteq G$$

(2) Sei G konvex. Die Funktion $f : G \to \mathbb{R}$ heißt konvex, wenn gilt

$$f(x + \lambda(y - x)) \leq f(x) + \lambda\,(f(y) - f(x))$$

für alle $x, y \in G$ und $\lambda \in [0, 1]$.

(3) Sei G konvex. Eine Funktion $f : G \to \mathbb{R}$ heißt streng konvex, wenn gilt

$$f(x + \lambda(y - x)) < f(x) + \lambda\,(f(y) - f(x))$$

für alle $x, y \in G$ und $\lambda \in [0, 1]$.

Ausgehend von diesen Begrifflichkeiten erhalten wir das folgende Resultat:

Satz 2.2

Sei $G \subseteq \mathbb{R}^n$ eine konvexe Menge und $f : G \to \mathbb{R}$ eine konvexe Funktion.

(1) Jede lokale Lösung ist gleichzeitig auch globale Lösung von (2.1).

(2) Falls f sogar streng konvex ist, dann existiert höchstens eine Lösung.

Beweis. (1) Sei $\tilde{x} \in G$ eine lokale Lösung von (2.1). Wir nehmen an, dass dies jedoch keine globale Lösung ist, d.h. es existiert ein $\overline{x} \in G$ mit $f(\overline{x}) < f(\tilde{x})$. Wegen der Konvexität von G gilt dann $x(\lambda) = \tilde{x} + \lambda(\overline{x} - \tilde{x}) \in G$ für alle $\lambda \in [0, 1]$. Mit der Konvexität von f folgt letzlich

$$f(x(\lambda)) = f(\widetilde{x} + \lambda(\overline{x} - \widetilde{x})) \overset{f \text{ konvex}}{\leq} f(\widetilde{x}) + \underbrace{\lambda}_{>0} \underbrace{(f(\overline{x}) - f(\widetilde{x}))}_{<0} < f(\widetilde{x}) \qquad \forall \lambda \in (0,1]$$

Somit ist $\widetilde{x}$ keine lokale Lösung im Widerspruch zur Annahme.

(2) Seien x, y zwei voneinander verschiedene Lösungen., d.h. $f(x) = f(y) = f^*$. Wir erhalten $x(\lambda) \in G$ für alle $\lambda \in [0,1]$ und

$$f(x(\lambda)) = f(x + \lambda(y - x)) \overset{f \text{ streng konvex}}{<} f(x) + \lambda \underbrace{(f(y) - f(x))}_{=0}$$

Somit ist x keine Lösung. □

Für konvexe Optimierungsaufgaben sind lokale und globale Lösungen also äquivalent. Als wichtigen Spezialfall konvexer Mengen halten wir die folgende Darstellung fest.

Aussage 1
Sei G gegeben durch

$$G := \{x \in \mathbb{R}^n : g_i(x) \leq 0, i \in I, h_j(x) = 0, j \in J\}$$

Dann gilt: falls alle Funktionen g_i $(i \in i)$ konvex und alle Funktionen h_j $(j \in J)$ affin-linear sind, dann ist G konvex.

Beweis. Seien $x, y \in G$ und $\lambda \in (0,1)$. Zur Klärung der Konvexität, stellt sich die Frage, ob $x(\lambda) \in G$?

$$\begin{aligned} g_i(x(\lambda)) &= g_i(x + \lambda(y-x)) \leq g_i(x) + \lambda\,(g_i(y) - g_i(x)) = \underbrace{1-\lambda}_{>0} \underbrace{g_i(x)}_{\leq 0} + \underbrace{\lambda}_{>0} \underbrace{g_i(y)}_{\leq 0} \\ &\leq 0 \\ h_j(x(\lambda)) &= h_j(x + \lambda(y-x)) = A_j(x + \lambda(y-x)) + b_j = (1-\lambda)A_j x + \lambda A_j y + b_j \\ &= (1-\lambda)[\underbrace{A_j x + b_j}_{h_j(x)=0}] + \lambda[\underbrace{A_j y + b_j}_{h_j(y)=0}] \\ &= 0 \end{aligned}$$

Somit ist $x(\lambda) \in G$ und G also konvex. □

Jeder zulässige Bereich einer linearen Optimierungsaufgabe (↗ Kapitel 3) hat diese Gestalt.

2.2 Optimalitätsbedingungen

Definition 2.3

Eine Menge $K \subseteq \mathbb{R}^n$ heißt **Kegel**, falls gilt:

$$x \in K \quad \Rightarrow \quad \lambda x \in K \quad \forall \lambda \geq 0$$

Ein Kegel K ist ein **konvexer Kegel**, falls K eine konvexe Menge bzw. falls gilt

$$x, y \in K \quad \Rightarrow \quad x + y \in K$$

für alle $x, y \in K$. Der **Kegel der zulässigen Richtungen** $Z(\widetilde{x})$ ist definiert durch

$$Z(\widetilde{x}) := \{d \in \mathbb{R}^n \mid \exists \bar{t} := \bar{t}(\widetilde{x}, d) > 0 \text{ sodass } \widetilde{x} + td \in G \;\; \forall t \in [0, \bar{t}]\}$$

Für Optimierungsaufgaben ist der Kegel der zulässigen Richtungen von großer Bedeutung.

Aussage 2 (notwendiges Optimalitätskriterium)

Ist f auf G stetig differenzierbar und $\widetilde{x} \in G$ ein lokales Minimum. Dann gilt

$$\nabla f(\widetilde{x})^\top \cdot d \geq 0 \qquad \forall d \in Z(\widetilde{x}) \tag{2.2}$$

Ist G konvex, dann erhält man die Bedingung

$$\nabla f(\widetilde{x})^\top (x - \widetilde{x}) \geq 0 \qquad \forall x \in G \tag{2.3}$$

Beweis. Sei $\widetilde{x}$ ein lokales Minimum und $d \in Z(\widetilde{x})$ eine zulässige Richtung. Dann existiert gemäß Definition ein $\bar{t}$, sodass $\widetilde{x} + td \in G$ für alle $t \in [0, \bar{t}]$ gilt. Weil außerdem $\overline{x}$ eine lokale Lösung ist, gibt es $\rho > 0$ mit $\rho < \bar{t}$ sodass $f(\widetilde{x} + td) \geq f(\widetilde{x})$ für $t \in (0, \rho)$ gilt. Aus dieser Ungleichung folgt

$$\frac{f(\widetilde{x} - td) - f(\widetilde{x})}{t} \geq 0 \qquad \forall t \in (0, \rho)$$

Durch Grenzwertbildung $t \to 0$ auf beiden Seiten erhält man mithilfe der Definition der Richtungsableitung und der Stetigkeit von f die Behauptung (2.2). Für konvexe Mengen gilt stets $x - \widetilde{x} \in Z(\widetilde{x})$ für $x \in G$, also folgt (2.3). □

Dieses Kriterium sagt aus, dass im Punkt $\widetilde{x}$ alle Richtungsableitungen (bezüglich zulässiger Richtungen) nicht-negativ sind, d.h. es keine zulässige Abstiegsrichtung gibt.

Bemerkung 2.2

Ein Punkt, der die Bedingung (2.2) erfüllt, heißt **stationärer Punkt**.

Bemerkung 2.3

Bei der freien Minimierung (d.h. für $G = \mathbb{R}^n$) ergibt sich wegen $Z(\widetilde{x}) = \mathbb{R}^n$ für alle $\widetilde{x} \in G$ die notwendige Bedingung

$$\widetilde{x} \text{ ist lokales Minimum} \quad \Rightarrow \quad \nabla f(\widetilde{x}) = 0$$

Wähle dafür $d \in \{\pm e^i\}_{i=1}^n$.

Für konvexe Optimierungsaufgaben gilt auch die Umkehrung des Resultats der vorherigen Aussage.

Aussage 3 (hinreichendes Optimalitätskriterium)
Es seien $G \subseteq \mathbb{R}^n$ sowie $f\colon G \to \mathbb{R}$ konvex und stetig differenzierbar. Falls ein $\widetilde{x} \in G$ existiert, welches der Bedingung (2.3) genügt, dann ist $\widetilde{x}$ (globales) Minimum von (2.1).

Beweis. Wenn f konvex und stetig differenzierbar ist und gilt

$$f(x) \geq f(\widetilde{x}) + \nabla f(\widetilde{x})^\top (x - \widetilde{x}) \qquad \forall x \in G$$

Wegen (2.3) folgt unmittelbar die (globale) Optimalität. Ausführlicher: siehe Übung. □

Im Fall polyedrischer zulässiger Mengen $G \subseteq \mathbb{R}^n$ (wie z.B. in der linearen Optimierung) kann die Bedingung (2.2) präzisiert werden, da dann $Z(x)$ eine einfache Struktur besitzt.

Definition 2.4
$G \subseteq \mathbb{R}^n$ heißt **polyedrisch**, falls eine Darstellung $G = \{x \in \mathbb{R}^n \colon Ax \leq b\}$ für eine geeignete Matrix A und einen geeigneten Vektor b existiert. Hierbei gilt

$$Ax \leq b \quad :\Leftrightarrow \quad \forall i \in I = \{1, \ldots, n\} : a_i^\top x = \sum_{j=1}^{n} a_{ij} x_j \leq b_i$$

Bemerkung 2.4
Eine polyedrische Menge G ist konvex und abgeschlossen, aber im Allgemeinen nicht beschränkt. Implizit können in der Beschreibung von G aus Definition 2.4 auch Gleichungsrestriktionen enthalten sein.

Definition 2.5
Für $x \in G$ ist die **Indexmenge der aktiven Restriktionen** definiert durch

$$I_0(x) := \left\{ i \in I \colon a_i^\top x = b_i \right\}$$

Sei nun ein zulässiger Punkt $x \in G$ gegeben. Damit eine beliebige Richtung $d \in \mathbb{R}^n$ zulässig ist, nuss ein $\overline{t} > 0$ existieren, sodass $x + td \in G$ für alle $t \in [0, \overline{t}]$ gilt. Für einen polyedrischen Bereich G ist dies äquivalent zu

$$\forall i \in I \colon a_i^\top (x + td) \leq b_i \quad \Leftrightarrow \quad \forall i \in I \colon t a_i^\top d \leq b_i - a_i^\top x$$

für alle $t \in [0, \overline{t}]$.

- Für alle inaktiven Restriktionen (also solche $a_i^\top x < b_i$) wäre $t a_i^\top d \leq b_i - a_i^\top x$ zu erfüllen. Egal, welchen Wert $a_i^\top d$ annimmt, es kann stets eine hinreichend kleine Schrittweite (im Sinne der Definition einer zulässigen Richtung) gefunden werden. Somit schränken inaktive Restriktionen die möglichen Richtungen $d \in \mathbb{R}^n$ *nicht* ein.
- Für aktive Restriktionen (also $a_i^\top x = b_i$) erhält man $t a_i^\top d \leq 0$, also (wegen $t > 0$) $a_i^\top d \leq 0$.

Diese Bedingung lässt sich geometrisch interpretieren: das Skalarprodukt der zulässigen Richtungen und des Normalenvektors (nach außen gerichtet) a_i der begrenzenden Hyperebene muss kleiner oder gleich Null sein, d.h. der Schnittwinkel beider Vektoren liegt im Bereich $[\frac{\pi}{2}, \pi]$. Folglich zeigt die zulässige Richtung $d \in \mathbb{R}^n$ tatsächlich in das Innere von G.

Für einen zulässigen Punkt $x \in G$ kann somit folgende Beobachtung angegeben werden:

$$d \in Z(x) \quad \Leftrightarrow \quad \forall u \in I_0(x) \colon a_i^\top d \leq 0 \tag{2.4}$$

Außerdem ist die Größe $\widetilde{t}$ (maximale Schrittweite) wohldefiniert.

$$\widetilde{t} := \widetilde{t} := \min \left\{ \frac{b_i - a_i^\top x}{a_i^\top d} \colon i \in I(x,d) \right\} \tag{2.5}$$

wobei $I(x,d) := \left\{ i \in I \colon a_i^\top d > 0 \right\}$.

Bemerkung 2.5

Falls $I(x,d) = \emptyset$, setzen wir $\widetilde{t} := \infty$.

Beispiel 2.2

Wir betrachten $x := (1,1,1)^\top$ und die polyedrische Menge

$$G := \left\{ (x_1, x_2, x_3)^\top \in \mathbb{R}^3 \colon x_1 + 2x_2 + x_3 \leq 4, 3x_1 + x_2 + x_3 \leq 6, x_i \geq 0, i = 1,2,3 \right\}$$

Offenbar gilt $x \in G$. Wir betrachten die Richtungen

$$d^1 = (1,1,1)^\top \text{ und } d^2 = (-1,-2,-1)^\top$$

Als aktive Restriktionen erkennen wir $I_0(x) = \{1\}$ (da nur die erste Nebenbedingung von G mit Gleichheit erfüllt ist).

- Für $d = d^1$ gilt
$$a_i^\top d = \left(\begin{smallmatrix} 1 \\ 2 \\ 1 \end{smallmatrix}\right)^\top \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right) = 4 > 0$$
Somit ist d^1 keine zulässige Richtung wegen (2.4).

- Für $d = d^2$ gilt
$$a_i^\top d = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}^\top \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix} = -6 \leq 0$$
Somit ist d^2 eine zulässige Richtung wegen (2.4). Zur maximalen Schrittweite: Die Ungleichung $(3,1,1)^\top (x + td) \leq 6$ liefert die Bedingung $t \in [-\frac{1}{6}, \infty)$. Aus $x + td \geq 0$ folgt die Bedingung $t \leq \frac{1}{2}$. Insgesamt gilt $\widetilde{t} = \frac{1}{2}$.

Zusammengefasst erhalten wir das folgende Resultat:

Folgerung 2.1
Sei G polyedrisch, d.h. $G = \{x \in \mathbb{R}^n \colon Ax \leq b\}$ und $f \colon G \to \mathbb{R}$ stetig differenzierbar. Ist $\widetilde{x}$ eine lokale Lösung von (2.1), so gilt

$$\nabla f(\widetilde{x})^\top \cdot d \geq 0 \qquad \forall d \in \mathbb{R}^n \text{ mit } a_i^\top d \leq 0 \ \forall i \in I_0(\widetilde{x}) \tag{2.6}$$

Ist f zusätzlich konvex, dann gilt auch die Umkehrung.

2.3 Das Lemma von Farkas

Das folgende Resultat besitzt vielfältige Anwendungen in der Optimierung ($\nearrow$ Dualität).

Lemma 2.1 (Farkas)
Es seien $A \in \mathbb{R}^{m \times n}$ und $a \in \mathbb{R}^m$. Von den Systemen
(i) $Az \leq 0$, $a^\top z > 0$
(ii) $A^\top u = a$, $u \geq 0$.
ist *genau* eines lösbar.

Beweis. ▪ *höchstens* eines der Systeme ist lösbar: Seien (i) und (ii) lösbar. Dann gilt

$$0 < a^\top z = A^\top u^\top z = \underbrace{u^\top}_{\geq 0} \underbrace{Az}_{\leq 0} \leq 0 \quad \grave{}$$

- *mindestens* eines der Systeme ist lösbar — die Unlösbarkeit von (ii) impliziert die Lösbarkeit von (i): Sei (ii) nicht lösbar. Dann gilt $a \notin K := \left\{x = A^\top u \colon u \geq 0\right\}$, wobei K ein konvexer, abgeschlossener Kegel ist. Wir betrachten die Optimierungsaufgabe

$$f(x) = \frac{1}{2} \|a - x\|_2^2 = \frac{1}{2} (a - x)^\top (a - x) \to \min \text{ bei } x \in K$$

Dann existiert eine eindeutige (und globale) Lösung $\overline{x} \in K$ mit

$$(1) \quad \nabla f(\overline{x})^\top \overline{x} = 0 \qquad\qquad (2) \quad \nabla f(\overline{x})^\top x \geq 0 \ \forall x \in K$$

Zunächst folgt gemäß **??** 2, dass $\nabla f(\overline{x})^\top (x - \overline{x}) \geq 0$ für alle $x \in K$. Durch Einsetzen von $x = \frac{1}{2}\overline{x} \in K$ und $x = 2 \cdot \overline{x} \in K$ (beachte: K ist Kegel) erhält man (1). Dies wiederum lässt sich zur notwendigen Bedingung ddazu addieren und man erhält (2). Nun zeigen wir, dass $z := a - \overline{x} \neq 0$ (wegen $a \notin K$) das System (i) löst. Es gilt $\nabla f(\bar{}) = -z$ und damit folgt

$$0 = \nabla f(\overline{x})^\top \overline{x} = -z^\top \cdot (\overline{x} - a + a) = z^\top (z - a) \ \Rightarrow \ a^\top z = z^\top z \overset{z \neq 0}{>} 0$$

Weiter gilt $x \in K$ genau dann, wenn ein $u \geq 0$ existiert mit $x = A^\top u$. Aus (2) folgt dann

$$\begin{aligned} \nabla f(\overline{x})^\top x \geq 0 \ \forall x \in K \quad &\Rightarrow \quad -z^\top A^\top u \geq 0 \quad \forall u \geq 0 \\ &\Rightarrow \quad (Az)^\top u \leq 0 \quad \forall u \geq 0 \\ &\Rightarrow \quad Az \leq 0 \quad (\text{wähle z.B. wieder } u = e^1, e^2, \dots) \end{aligned}$$

Damit löst z das System (1). □

Damit können die notwendigen Optimalitätsbedingungen (2.6) bzw. äquivalent dazu

$$\forall i \in I_0(x)\colon a_i^\top \cdot d \leq 0 \ \Rightarrow \ \nabla f(\widetilde{x})^\top \cdot d \geq 0 \tag{2.7}$$

wie folgt umformuliert werden: Offenbar ist (2.7) gleichbedeutend mit der Unlösbarkeit von

$$\nabla f(\widetilde{x})^\top \cdot d < 0, \qquad a_i^\top \cdot d \leq 0 \ \forall i \in I_0(\widetilde{x})$$

Wählt man also im Lemma von Farkas $a = -\nabla f(\widetilde{x})$ und A bestehend aus den Zeilen $a_i^\top$, so folgt die Lösbarkeit des Systems

$$\nabla f(\widetilde{x}) + \sum_{i \in I_0(\widetilde{x})} u_i a_i = 0 \qquad (u \geq 0) \tag{2.8}$$

Für konvexe Optimierungsaufgaben ist die Lösbarkeit von (2.8) sogar äquivalent dazu, dass $\widetilde{x}$ Lösung der betrachteten Aufgabe ist.