

1 Einleitung

Optimierungsprobleme sind allgegenwärtig, und die Wurzeln zu ihrer Behandlung reichen weit zurück. In Sektion 1.2 geben wir daher ‘zum Warmwerden’ eine Reihe verschiedener Beispiele, die die große Vielfalt der Anwendungsbereiche verdeutlicht. Hierzu gehören verschiedene Probleme der Produktions- und Routenplanung, aber auch ein Grundproblem der diskreten Tomographie sowie eine aktuelle Frage aus dem Chip-Design. Wir beginnen in Sektion 1.1 mit der Definition der allgemeinen und einiger spezieller Optimierungsprobleme, die im Zentrum unserer Untersuchungen stehen werden. Sektion 1.3 fasst die wesentliche (meistens den üblichen Standards folgende) Notation zusammen, die in allen Kapiteln durchgängig Verwendung findet.

Literatur: Beispiele, Hinweise zur Modellierung sowie einführende Kommentare sind in jedem der in der Literaturliste angegebenen Bücher zur Optimierung enthalten; man beachte jedoch insbesondere [83].

1.1 Optimierungsprobleme

Wie kann man die Produktion von Gütern bestmöglich planen, wie neue Standorte für Auslieferungslager festlegen, so dass die Versorgung der Kunden kostengünstig und umweltschonend erfolgt? Wie lassen sich Arbeitskräften Aufgaben, Personen Partner oder Gütern Regalplätze optimal zuordnen? Wie kann man kristalline Strukturen aus Aufnahmen der Elektronenmikroskopie rekonstruieren, und wie lassen sich Computerchips energieeffizient auslegen? Trotz der Vielfalt ihrer Anwendungsfelder gehören zu solchen und anderen Optimierungsaufgaben gleiche ‘Zutaten’. Stets ist eine Grundmenge von zulässigen Objekten gegeben, aus der (mindestens) eines ausgewählt werden soll. Diese Menge kann endlich sein oder nicht, kann durch einfache Funktionen beschrieben werden oder nicht; in jedem Fall muss sie irgendwie ‘spezifiziert’ werden, um mit ihr arbeiten zu können. Die zweite typische Zutat ist eine Zielfunktion, mit deren Hilfe die zur Auswahl stehenden Objekte bewertet werden. Diese kann die Länge von Wegstrecken bei der Routenplanung, die CO₂-Belastung beim Einsatz verschiedener Fahrzeugtypen, die Passgenauigkeit der Hobbies von potentiellen Lebenspartnern messen, die Erträge von Investitionsstrategien schätzen oder auch auf ganz anderen Zielvorstellungen beruhen. In jedem Fall ist eine quantitative Bewertung der erwünschten Eigenschaften einer Lösung erforderlich. Wir werden in den einzelnen Kapiteln zahlreiche Anwendungsbeispiele ausführlich besprechen. Insbesondere werden bereits in Sektion 1.2 viele konkrete Beispiele vorgestellt. Wir beginnen aber zunächst mit einigen wenigen Begriffen, um diese Bestandteile der Optimierung formal zu beschreiben und besonders relevante Klassen von Optimierungsproblemen einzuführen.

1.1.1 Definition. (a) Eine allgemeine **Optimierungsaufgabe** [engl.: *instance (of an optimization problem)*] (über \mathbb{R}) ist durch folgende Daten spezifiziert:

$$n \in \mathbb{N} \quad \wedge \quad F \subset G \subset \mathbb{R}^n \quad \wedge \quad \varphi : G \rightarrow \mathbb{R} \quad \wedge \quad \text{opt} \in \{\min, \max\}.$$

Jeder Punkt $x^* \in F$ mit

$$x \in F \Rightarrow \varphi(x^*) \leq \varphi(x)$$

für $\text{opt} = \min$ bzw.

$$x \in F \Rightarrow \varphi(x^*) \geq \varphi(x)$$

für $\text{opt} = \max$ heißt **Optimalpunkt** (bzw. **Minimal-** oder **Maximalpunkt**). Die Aufgabe besteht darin, φ über F zu optimieren, d.h. zu entscheiden, ob F leer ist, ob φ nach unten (für $\text{opt} = \min$) bzw. nach oben (für $\text{opt} = \max$) beschränkt ist, ob ein Optimalpunkt existiert und, falls das der Fall ist, einen solchen zu bestimmen.

- (b) φ heißt **Zielfunktion** [engl.: objective function], F **zulässiger Bereich** [engl.: feasible region] der gegebenen Aufgabe, und jeder Punkt $x \in F$ heißt **zulässig**. Ist der zulässige Bereich leer, so wird die Aufgabe **unzulässig** genannt. Ist $\text{opt} = \min$, so liegt eine **Minimierungsaufgabe** vor; für $\text{opt} = \max$ eine **Maximierungsaufgabe**.
- (c) Ist $F = \mathbb{R}^n$, so spricht man von einer Aufgabe der **unrestringierten Optimierung** [engl.: unconstrained optimization], für $F \neq \mathbb{R}^n$ von einer Aufgabe der **restringierten Optimierung**.
- (d) Die Menge aller Optimierungsaufgaben heißt **Optimierungsproblem**, die Teilmengen mit $\text{opt} = \min$ bzw. $\text{opt} = \max$ heißen **Minimierungsproblem** bzw. **Maximierungsproblem**. Ist ein Optimierungsproblem gegeben, so bezeichnet man jede zugehörige Aufgabe oft auch als **Instanz** des Optimierungsproblems.
- (e) Besteht das Ziel darin, den (in $\mathbb{R} \cup \{\pm\infty\}$ liegenden) Wert $\inf\{\varphi(x) : x \in F\}$ (für $\text{opt} = \min$) bzw. $\sup\{\varphi(x) : x \in F\}$ (für $\text{opt} = \max$) zu bestimmen, so spricht man von einer **Auswertungs-** oder **Evaluationsaufgabe**. Entsprechend ist auch das **Auswertungs-** oder **Evaluationsproblem** definiert.
- (f) Besteht die Aufgabe darin, festzustellen, ob der zulässige Bereich einer gegebenen Aufgabe nicht leer ist¹, so spricht man von einer **Zulässigkeitsaufgabe**. Entsprechend ist auch das **Zulässigkeitsproblem** definiert.

Natürlich sind die so formulierten Probleme zu allgemein, um gute einheitliche Algorithmen zu ermöglichen, und wir werden in den folgenden Kapiteln verschiedene Problemklassen betrachten, deren spezielle Struktur zur Konstruktion ebensolcher Algorithmen ausgenutzt wird. Auch ist die vorgenommene Trennung in verschiedene Typen ohne weitere Einschränkungen nicht wirklich sinnvoll. Natürlich kann man durch Übergang von einer Zielfunktion φ_1 zu der neuen Zielfunktion $\varphi_2 := -\varphi_1$ eine Maximierungsaufgabe in eine äquivalente Minimierungsaufgabe überführen und umgekehrt. Das folgende Beispiel zeigt, dass auch die Unterscheidung zwischen restringierter und unrestringierter Optimierung willkürlich sein kann.

1.1.2 Beispiel. Seien $n \in \mathbb{N}$, $G := \mathbb{R}^n$, $F_1 \subset G$ kompakt, $F_2 := \mathbb{R}^n$, und $\varphi_1, \varphi_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ seien für $x := (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n$ definiert durch

$$\varphi_1(x) := \sum_{i=1}^n \xi_i^2 \quad \wedge \quad \varphi_2(x) := \begin{cases} \varphi_1(x) & \text{für } x \in F_1, \\ -1 & \text{für } x \notin F_1. \end{cases}$$

¹ Das entspricht der Wahl $\varphi \equiv 0$ der Zielfunktion.

Dann sind durch $(n, F_1, G, \varphi_1, \max)$ eine restringierte und durch $(n, F_2, G, \varphi_2, \max)$ eine unrestringierte Maximierungsaufgabe gegeben.

Für alle $x \in F_1$ ist $\varphi_1(x) \geq 0$. Daher gilt

$$F_1 = \emptyset \quad \Leftrightarrow \quad \max\{\varphi_2(x) : x \in F_2\} = -1.$$

Für $F_1 \neq \emptyset$ stimmen ferner die Mengen der Maximalpunkte der beiden Aufgaben überein. Die restringierte und unrestringierte Maximierungsaufgabe sind somit 'äquivalent'.

Beispiel 1.1.2 zeigt, dass man bei der Festlegung von Optimierungsaufgaben durchaus die Freiheit hat, Eigenschaften zwischen der Zielfunktion und dem zulässigen Bereich 'hin und her zu schieben'. Die Struktur von F_1 findet sich in der speziellen Struktur von φ_2 wieder. Will man also verschiedene Typen von Optimierungsproblemen voneinander unterscheiden, so muss man die sie spezifizierenden Mengen und Funktionen auf spezielle Klassen einschränken.

In der allgemeinen Definition ist nicht gefordert, dass F abgeschlossen und φ stetig ist, so dass selbst dann, wenn F nicht leer und beschränkt ist, Optimalpunkte nicht zu existieren brauchen. Dennoch schreibt man allgemeine Optimierungsaufgaben oftmals suggestiv in der folgenden Form.

1.1.3 Bezeichnung. Für eine durch (n, F, G, φ, \max) spezifizierte allgemeine Maximierungsaufgabe wird oftmals die Schreibweise

$$\begin{aligned} & \max \varphi(x) \\ & x \in F \end{aligned}$$

verwendet, ohne damit ausdrücken zu wollen, dass das Maximum auch angenommen wird.² Ferner bezeichnet $\operatorname{argmax}\{\varphi(x) : x \in F\}$ die **Menge der Maximalpunkte** der Aufgabe, d.h. wir schreiben

$$x^* \in \operatorname{argmax}\{\varphi(x) : x \in F\}$$

genau dann, wenn x^* ein Maximalpunkt der Aufgabe ist.

Analog werden auch bei der Minimierung die Schreibweisen

$$\begin{aligned} & \min \varphi(x) \\ & x \in F \end{aligned}$$

sowie

$$x^* \in \operatorname{argmin}\{\varphi(x) : x \in F\}$$

zur Bezeichnung der Aufgabe bzw. der **Menge der Minimalpunkte** verwendet. Oft schreiben wir auch

$$\operatorname{argmax}_{x \in F} \varphi(x), \quad \operatorname{argmin}_{x \in F} \varphi(x).$$

Von praktisch großer Relevanz ist die Frage, wie eigentlich die Zielfunktion und der zulässige Bereich genau gegeben sind. Wir formulieren in Definition 1.1.1 einfach kühn $F \subset \mathbb{R}^n$, aber das ist solange bloß ein theoretisches Konzept, wie wir nicht sagen, wie F konkret und operational, d.h. algorithmisch umsetzbar dargestellt wird.

² Es wäre also eigentlich konsequenter $\sup\{\varphi(x) : x \in F\}$ zu verwenden, und wenn immer es wirklich darauf ankommt, werden wir das auch tun. Die Konvention ist aber für viele Klassen von Optimierungsproblemen üblich.

1.1.4 Beispiel. Seien $n := 1$, $F := \{0\}$, $G := \mathbb{R}$, und sei $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ für jedes $\xi \in \mathbb{R}$ definiert durch $\varphi(\xi) := \xi^2 + 1$. Es gilt

$$F := \{0\} = \{\xi \in \mathbb{R} : \xi \geq 0 \wedge 2\xi \leq 1 \wedge \xi(1 - \xi) = 0\}.$$

Dieselbe Menge F wird hier durch Angabe ihres einzigen Elements oder als Lösungsmenge eines linearen Ungleichungssystems und einer quadratischen Gleichung dargestellt. Natürlich hat die zugehörige Maximierungsaufgabe stets dieselbe Lösung 1. Dieses ist für die erste Beschreibung von F sofort durch Einsetzen des einzigen zulässigen Punktes 0 in die Zielfunktion erkennbar. Die zweite Darstellung erfordert hingegen eine genauere Analyse der angegebenen Bedingungen, etwa, dass die quadratische Gleichung lediglich die Lösungen 0 und 1 besitzt, die erste lineare Bedingungen somit redundant, also überflüssig ist, und die zweite lineare Ungleichung den Punkt 1 ausschließt.

Anders als in Beispiel 1.1.4 kann man im Allgemeinen nicht alle zulässigen Punkte einzeln aufzählen. Man benötigt insbesondere für Mengen mit mehr als endlich vielen Punkten ‘kompaktere’ Darstellungen, mit denen man praktisch arbeiten kann. Für die Zielfunktion gilt das Gleiche.

Verschiedene Beschreibungen der spezifizierenden Daten können dabei durchaus sehr unterschiedliche Anforderungen an Lösungsverfahren für die gegebenen Optimierungsprobleme stellen. Dieses Phänomen zieht sich durch das gesamte Gebiet der Optimierung. Andererseits basieren viele, für die Entwicklung von Algorithmen zentrale Konzepte auf strukturellen Eigenschaften der zugrunde liegenden Probleme. Oft stehen dabei geometrische Charakteristika im Vordergrund. Da sich viele Ansätze durch das Wechselspiel zwischen strukturellen Eigenschaften der Probleme und Besonderheiten der Darstellungen der relevanten Mengen und Funktionen ergeben, werden wir stets versuchen, darstellungsunabhängige Strukturaussagen von darstellungsabhängigen Verfahren zu trennen. Daher werden wir oft von ‘theoretisch gegebenen’ Daten ausgehen, diese im Laufe der Entwicklung von Algorithmen aber immer weiter operationalisieren.

Von besonderer Bedeutung sind die Klassen der linearen und ganzzahligen (linearen) Optimierungsprobleme, in denen sich die Zielfunktion und die Restriktionen als Skalarprodukte darstellen lassen.

1.1.5 Definition. (a) Eine **lineare Optimierungsaufgabe** oder ein **lineares Programm** (über \mathbb{R}) ist durch folgende Daten spezifiziert:

$$m, n \in \mathbb{N} \quad \wedge \quad a_1, \dots, a_m \in \mathbb{R}^n \quad \wedge \quad \beta_1, \dots, \beta_m \in \mathbb{R} \quad \wedge \quad \gamma_1, \dots, \gamma_n \in \mathbb{R}.$$

Ferner sei das lineare Funktional $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ für $x := (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n$ durch

$$\varphi(x) := \sum_{i=1}^n \gamma_i \xi_i$$

definiert, und es sei

$$F := \{x \in \mathbb{R}^n : a_1^T x \leq \beta_1 \wedge \dots \wedge a_m^T x \leq \beta_m\}.$$

Ziel ist es, φ über F zu maximieren.

(b) Ist zusätzlich eine Teilmenge

$$J \subset \{1, \dots, n\}$$

spezifiziert, und ist für jedes $j \in J$ die Bedingung

$$\xi_j \in \mathbb{Z}$$

zu erfüllen, so spricht man von einer **gemischt-ganzzahligen (linearen) Optimierungsaufgabe**. Für $J := \{1, \dots, n\}$ heißt die Aufgabe **ganzzahlige (lineare) Optimierungsaufgabe**.

Sind hingegen zusätzlich die Bedingungen

$$\xi_1, \dots, \xi_n \in \{0, 1\}$$

zu erfüllen, so spricht man von einer **(linearen) 0-1-Optimierungsaufgabe**.

- (c) Die Menge aller linearen Optimierungsaufgaben heißt **lineares Optimierungsproblem**.

Analog sind auch das **gemischt-ganzzahlige (lineare) Optimierungsproblem**, das **ganzzahlige (lineare) Optimierungsproblem** und das **(lineare) 0-1-Optimierungsproblem** definiert. Kürzer werden auch die Bezeichnungen **LP-Problem** (*linear programming*), **MILP-Problem** (*mixed-integer linear programming*) bzw. **ILP-Problem** (*integer linear programming*) verwendet.

- (d) Zur Betonung der in (a) – (c) verwendeten Darstellung der Restriktionen durch Ungleichungen des Typs $a^T x \leq \beta$ und zur Abgrenzung von anderen Formen der Aufgaben und Probleme (vgl. Bezeichnung 1.4.1) wird bisweilen der Zusatz **‘in natürlicher Form’** verwendet.

In Definition 1.1.5 ist es erlaubt, dass einige der Vektoren a_1, \dots, a_m gleich Null sind. Für $a_{i_0} = 0$ und $\beta_{i_0} \geq 0$ wird die Bedingung $a_{i_0}^T x \leq \beta_{i_0}$ von jedem Vektor $x \in \mathbb{R}^n$ erfüllt, für $a_{i_0} = 0$ und $\beta_{i_0} < 0$ von keinem. Im ersten Fall ist die Restriktion also überflüssig, im zweiten Fall ist der zulässige Bereich leer. Da man dieses den Daten unmittelbar ansehen kann, wird daher im Folgenden meistens

$$a_1, \dots, a_m \in \mathbb{R}^n \setminus \{0\}$$

vorausgesetzt.

Natürlich kann man, statt zu maximieren, φ auch minimieren wollen oder auch Nebenbedingungen in anderer Form betrachten, aber alle so entstehenden Formen lassen sich auf die oben definierte *natürliche Form* zurückführen. In Sektion 1.4 wird dieses im Detail durchgeführt.

Oft werden die Daten von LP-, MILP- und ILP-Aufgaben durch die Setzung

$$A := (a_1, \dots, a_m)^T \in \mathbb{R}^{m \times n} \quad \wedge \quad b := (\beta_1, \dots, \beta_m)^T \in \mathbb{R}^m \quad \wedge \quad c := (\gamma_1, \dots, \gamma_n)^T \in \mathbb{R}^n$$

zusammengefasst und die Aufgaben suggestiver geschrieben, etwa

$$\begin{aligned} \max \quad & c^T x \\ Ax \leq \quad & b \end{aligned}$$

bei einer LP-Aufgabe oder

$$\begin{aligned} \max \quad & c^T x \\ Ax \leq \quad & b \\ x \in \quad & \mathbb{Z}^n \end{aligned}$$

bei einer ILP-Aufgabe. Die Ungleichheitszeichen sind hier (und im Folgenden) jeweils komponentenweise zu verstehen.

Die zulässigen Bereiche linearer Optimierungsaufgaben haben eine spezielle Struktur, die in den Kapiteln 4 und 5 detailliert untersucht und zur Entwicklung eines in der Praxis sehr effizienten Algorithmus herangezogen wird. Hier soll aber zumindest bereits der entsprechende zentrale Begriff eingeführt werden.

1.1.6 Definition. (a) Sei $P \subset \mathbb{R}^n$. P heißt **Polyeder**³ [engl.: polyhedron], wenn, und nur wenn, es

$$n, m \in \mathbb{N}_0 \quad \wedge \quad A \in \mathbb{R}^{m \times n} \quad \wedge \quad b \in \mathbb{R}^m$$

gibt mit

$$P = \{x \in \mathbb{R}^n : Ax \leq b\}.$$

P heißt **Polytop** genau dann, wenn P beschränkt und ein Polyeder ist.

(b) Sind P ein Polyeder bzw. spezieller ein Polytop, und $n, m \in \mathbb{N}_0$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ mit $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, so heißt das Tupel (m, n, A, b) **\mathcal{H} -Darstellung**⁴ von P . Man spricht dann von P auch als **\mathcal{H} -Polyeder** bzw. **\mathcal{H} -Polytop**.

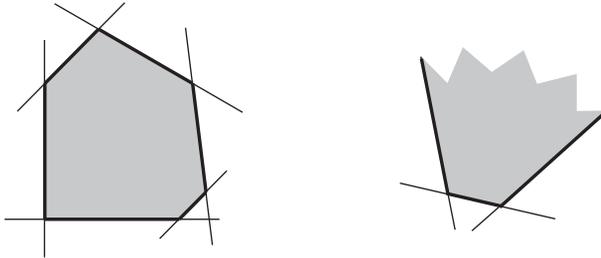
(c) Seien P ein Polyeder, $a \in \mathbb{R}^n$ und $\beta \in \mathbb{R}$. Die Ungleichung $a^T x \leq \beta$ heißt **redundant** für P , wenn

$$P = P \cap \{x \in \mathbb{R}^n : a^T x \leq \beta\}.$$

Seien (m, n, A, b) eine \mathcal{H} -Darstellung von P , $A = (a_1, \dots, a_m)^T$, $b = (\beta_1, \dots, \beta_m)^T$ sowie für $i \in \{1, \dots, m\}$

$$P_i := \{x \in \mathbb{R}^n : (j \in \{1, \dots, m\} \wedge j \neq i) \Rightarrow a_j^T x \leq \beta_j\}.$$

Das Tupel (m, n, A, b) heißt **irredundant**, wenn für kein $i \in \{1, \dots, m\}$ die Ungleichung $a_i^T x \leq \beta_i$ redundant für P_i ist.



1.1 Abbildung. Zwei Polyeder (grau, mit schwarzer Begrenzung) im \mathbb{R}^2 ; das links abgebildete Polyeder ist ein Polytop.

Man beachte, dass zu einer gegebenen \mathcal{H} -Darstellung (m, n, A, b) für jedes $\mu \in]0, \infty[$ auch $(m, n, \mu A, \mu b)$ eine \mathcal{H} -Darstellung von P ist. Ist ferner $a^T x \leq \beta$ eine für P redundante Ungleichung und setzt man

³ Der Begriff kommt aus dem Griechischen, $\pi\omicron\lambda\upsilon\epsilon\delta\rho\omicron$, und bedeutet 'vielflächig'. Laut Duden ist es Neutrum. Manchmal wird es in der Literatur aber auch fälschlich maskulin verwendet.

⁴ Diese Bezeichnung resultiert aus der geometrischen Interpretation der einzelnen, P darstellenden Ungleichungen als *Halbräume* des \mathbb{R}^n ; vgl. Definition 4.2.4.

$$\hat{A} := \begin{pmatrix} A \\ a^T \end{pmatrix} \quad \wedge \quad \hat{b} := \begin{pmatrix} b \\ \beta \end{pmatrix},$$

so ist auch (m, n, \hat{A}, \hat{b}) eine \mathcal{H} -Darstellung von P .

Ein Kernpunkt für spätere Algorithmen wird darin bestehen, geometrische, d.h. darstellungsinvariante Eigenschaften von Polyedern aus gegebenen Darstellungen zu extrahieren.

1.2 Einige Beispiele

Im Folgenden werden verschiedene Beispiele von Optimierungsaufgaben angeführt. Diese sind in einer auf ihre praktische Motivation zielenden Weise formuliert, wobei jedoch oftmals starke Vereinfachungen gegenüber den in der Praxis konkret verwendeten Modellen vorgenommen werden. Das geschieht einerseits, um die Modellierung nicht zu umfangreich werden zu lassen, und andererseits, um relevante Grundtypen von Optimierungsproblemen hervorzuheben.

Produktionsplanung: Ein Hersteller von Umweltlacken hat die technischen Möglichkeiten, n verschiedene Endprodukte herzustellen: Lacke L_1, \dots, L_n . Dafür stehen entsprechende Ressourcen (Rohstoffe, Arbeitsleistungen etc.) R_1, \dots, R_m zur Verfügung. Diese unterliegen Mengenrestriktionen β_1, \dots, β_m , d.h. für $i = 1, \dots, m$ können von Ressource R_i maximal β_i Mengeneinheiten eingesetzt werden. Für $j = 1, \dots, n$ erzielt man bei der Herstellung einer Mengeneinheit von Lack L_j einen Reingewinn von γ_j .⁵ Es soll ein Produktionsprogramm mit maximalem Gesamtgewinn ζ^* aufgestellt werden.

Offenbar besteht ein Produktionsprogramm gerade aus der Festlegung der Mengeneinheiten ξ_1^*, \dots, ξ_n^* , die von den Lacken L_1, \dots, L_n hergestellt werden sollen. Wird angenommen, dass sich der Gesamtgewinn additiv aus den Reingewinnen der einzelnen Lacke bestimmt, so ergibt er sich als

$$\sum_{j=1}^n \gamma_j \xi_j^*.$$

Natürlich müssen die Kapazitätsrestriktionen für die Ressourcen R_1, \dots, R_m eingehalten werden. Dazu benötigt man die genaue quantitative Kenntnis der Ingredienzien der einzelnen Lacke L_1, \dots, L_n .⁶ Wir nehmen an, zur Herstellung von L_j werden $\alpha_{i,j}$ Mengeneinheiten von Ressource R_i benötigt, und das gelte für jedes $i = 1, \dots, m$ und $j = 1, \dots, n$.

Natürlich kann $\alpha_{i,j}$ auch 0 sein, dann nämlich, wenn Lack L_j völlig R_i -frei ist. Die Beschränkungen der Ressourcen ergeben sich dann als

$$\sum_{j=1}^n \alpha_{i,j} \xi_j^* \leq \beta_i \quad (i = 1, \dots, m).$$

Somit erfüllt jede optimale Lösung $x^* := (\xi_1^*, \dots, \xi_n^*)^T$ folgende Bedingungen:

⁵ Die Preise werden hier als konstant angenommen; das ist in vielen Fällen gerechtfertigt, naturgemäß aber nicht in der Nähe der Marktsättigung. Ferner werden die Nettoerlöse als proportional zur hergestellten Menge angenommen. Es werden also insbesondere keine Fixkosten etc. berücksichtigt.

⁶ Üblicherweise werden diese als Betriebsgeheimnis behandelt, wodurch so manches Optimierungspotential ungenutzt bleibt.

$$\begin{aligned}
\sum_{j=1}^n \gamma_j \xi_j^* &= \zeta^* \\
\sum_{j=1}^n \alpha_{1,j} \xi_j^* &\leq \beta_1 \\
&\vdots \\
\sum_{j=1}^n \alpha_{m,j} \xi_j^* &\leq \beta_m.
\end{aligned}$$

Wenn außerdem die Produktionsprozesse als nicht umkehrbar angenommen werden⁷, dürfen wir keinen Lack in negativer Quantität herstellen, obwohl das durchaus manchmal profitabel sein könnte. Es muss also noch

$$\xi_1^*, \dots, \xi_n^* \geq 0$$

gelten. Es ist aber durchaus zugelassen, dass bei der Produktion eines Lacks Zwischenprodukte entstehen, die bei der Herstellung eines anderen Lacks Verwendung finden. Diese werden ebenfalls als Ressourcen behandelt, sind also in der Liste R_1, \dots, R_m aller Ressourcen enthalten. Die zugehörigen Koeffizienten $\alpha_{i,j}$ können somit durchaus negativ sein.

Die Gleichung und die Ungleichungen beschreiben bei bekanntem maximalem Gewinn ζ^* die Menge aller optimalen Produktionsprogramme. Da natürlich ζ^* und optimale Lösungen ξ_1^*, \dots, ξ_n^* im Allgemeinen erst gefunden werden müssen, führen wir stattdessen die Variablen ζ und ξ_1, \dots, ξ_n ein; die Menge aller zulässigen Produktionsprogramme ist dann durch

$$\begin{aligned}
\sum_{j=1}^n \alpha_{i,j} \xi_j &\leq \beta_j & (i = 1, \dots, m) \\
\xi_j &\geq 0 & (j = 1, \dots, n)
\end{aligned}$$

gegeben, und die Aufgabe besteht darin, die durch

$$\varphi(x) := \varphi(\xi_1, \dots, \xi_n) := \sum_{j=1}^n \gamma_j \xi_j$$

für $x \in \mathbb{R}^n$ gegebene Zielfunktion $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ über dem zulässigen Bereich zu maximieren. In Matrixschreibweise lässt sich das Problem formulieren als

$$\begin{aligned}
\max \quad & c^T x \\
Ax &\leq b \\
x &\geq 0,
\end{aligned}$$

wobei

$$c := (\gamma_1, \dots, \gamma_n)^T \quad \wedge \quad A := (\alpha_{i,j})_{\substack{i=1, \dots, m \\ j=1, \dots, n}} \quad \wedge \quad b := (\beta_1, \dots, \beta_m)^T$$

gilt und die Variablen im Vektor

$$x := (\xi_1, \dots, \xi_n)^T$$

zusammengefasst sind.

Formal kann in der Sprache der Produktionstheorie das lineare Produktionsproblem somit wie folgt spezifiziert werden.

⁷ Das ist der Kern der späteren Entsorgungproblematik.

1.2.1 Bezeichnung. Eine (*lineare*) **Produktionsaufgabe** ist spezifiziert durch folgende Daten

$$m, n \in \mathbb{N} \quad \wedge \quad A := (\alpha_{i,j})_{\substack{i=1, \dots, m \\ j=1, \dots, n}} \in \mathbb{R}^{m \times n}$$

$$b := (\beta_1, \dots, \beta_m)^T \in [0, \infty[^m \quad \wedge \quad c := (\gamma_1, \dots, \gamma_n)^T \in [0, \infty[^n.$$

Der Vektor b quantifiziert die **verfügbaren Ressourcen**, c ist der **Preisvektor**, und A heißt **Technologiematrix**.

Ferner sei das Zielfunktional $\varphi : \mathbb{R}^n \rightarrow [0, \infty[$ für jedes $x \in \mathbb{R}^n$ definiert durch $\varphi(x) := c^T x$. Ein Vektor $x \in \mathbb{R}^n$ heißt (**zulässiges**) **Produktionsprogramm**, wenn

$$Ax \leq b \quad \wedge \quad x \geq 0$$

gilt. Seien F die Menge aller zulässigen Produktionsprogramme und $x^* \in F$. Dann heißt x^* **optimales Produktionsprogramm**, wenn

$$x^* \in \operatorname{argmax} \{ \varphi(x) : x \in F \}$$

ist. Ziel der Produktionsaufgabe ist es zu entscheiden, ob ein optimales Produktionsprogramm existiert, und falls dem so ist, eines zu finden.

Die Menge aller Produktionsaufgaben heißt (**lineares**) **Produktionsproblem**.

1.2.2 Beispiel. Die Produktionsaufgabe sei gegeben durch

$$\begin{array}{rcl} \max & \xi_1 + \xi_2 + \xi_3 & \\ \xi_1 + 2\xi_2 + \xi_3 & \leq & 3 \\ -2\xi_1 + \xi_2 & \leq & 0 \\ \xi_1 & \leq & 1 \\ & \xi_2 & \leq 1 \\ & & \xi_3 \leq 1 \\ \xi_1, \xi_2, \xi_3 & \geq & 0. \end{array}$$

Es sollen also drei Endprodukte L_1, L_2, L_3 hergestellt werden, deren Nettoerlöse pro Einheit jeweils 1 sind. Dabei sind fünf verschiedene Ressourcen R_1, \dots, R_5 einzusetzen. Die erste Bedingung besagt, dass bei der Produktion jeweils einer Einheit von L_1, L_2 bzw. L_3 eine, zwei bzw. eine Einheit von R_1 verbraucht werden und insgesamt drei Einheiten von R_1 zur Verfügung stehen.

Die zweite Ungleichung betrifft nur L_1 und L_2 . Bei der Produktion einer Einheit von L_1 bzw. L_2 werden -2 bzw. eine Einheit von R_2 benötigt. Bei der Produktion einer Einheit von L_1 werden also jeweils zwei Einheiten von R_2 mitproduziert. R_2 ist somit ein Zwischenprodukt, das bei der Herstellung von L_1 entsteht, und das als Ressource für die Produktion von L_2 verwendet werden kann. Diese Möglichkeit der Zwischenproduktion von R_2 ist zwingend erforderlich dafür, dass L_2 überhaupt in positiver Menge produziert werden kann. Die rechte Seite 0 der zweiten Nebenbedingung zeigt, dass die Ressource R_2 so knapp ist, dass ihr Verbrauch im Gesamtprozess nicht positiv sein darf. Die Ungleichung erlaubt somit zwar eine überschüssige Produktion von R_2 , nicht aber einen noch so geringen (positiven) Verbrauch.

Die Bedingungen $\xi_1, \xi_2, \xi_3 \leq 1$ besagen, dass die Ressourcen R_3, R_4, R_5 bei der Produktion einer Einheit von L_1, L_2 bzw. L_3 mit jeweils einer Einheit eingehen, insgesamt aber auch nur mit jeweils einer Einheit zur Verfügung stehen.⁸

⁸ Ebenso kann man diese Bedingungen als Produktionsrestriktionen (EU-Obergrenzen, Kartellobergrenzen, Abnahmeobergrenzen, etc.) interpretieren.

Die Nichtnegativitätsbedingungen $0 \leq \xi_1, \xi_2, \xi_3$ modellieren, dass der Produktionsprozess die Lacke L_1, L_2, L_3 in nichtnegativer Menge herstellen soll, es sich also tatsächlich um einen reinen Produktions-, nicht aber um einen (Teil-) Entsorgungsprozess handelt. Da aus

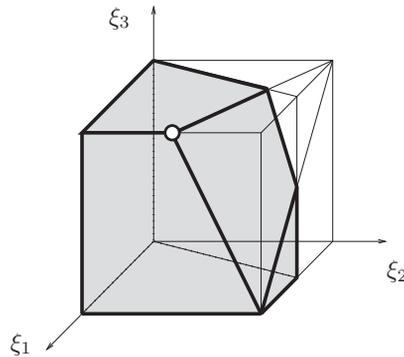
$$-2\xi_1 + \xi_2 \leq 0 \quad \wedge \quad \xi_2 \geq 0$$

direkt

$$\xi_1 \geq \frac{1}{2}\xi_2 \geq 0$$

folgt, ist die erste Nichtnegativitätsbedingung allerdings redundant.

Abbildung 1.2 zeigt die Menge aller zulässigen Produktionsprogramme sowie die (in diesem Beispiel) eindeutig bestimmte optimale Lösung $x^* := (1, \frac{1}{2}, 1)^T$; der Optimalwert der Zielfunktion ist $5/2$.



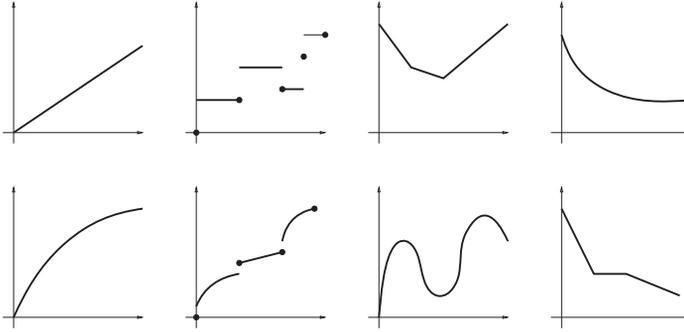
1.2 Abbildung. Geometrische Darstellung des Polytops der zulässigen Produktionsprogramme in Beispiel 1.2.2; der Optimalpunkt x^* ist hervorgehoben.

Man beachte, dass die Restriktionen $\xi_1 + 2\xi_2 + \xi_3 \leq 3$, $\xi_1 \leq 1$ sowie $\xi_3 \leq 1$ im Optimalpunkt mit Gleichheit erfüllt sind, für die übrigen Bedingungen aber noch 'Luft' ist (manchmal auch Schlupf genannt). Das bedeutet, dass die erste, dritte und fünfte Bedingung dafür 'verantwortlich' sind, dass der Produktionsprozess keinen größeren Gewinn ermöglicht. Würde man eine dieser Restriktionen lockern, so könnte der Gewinn gesteigert werden.

In späteren Kapiteln werden wir uns ausführlich mit der Frage beschäftigen, wie man solche linearen Optimierungsaufgaben auch dann lösen kann, wenn sie Hunderttausende oder mehr Variablen und Restriktionen enthalten. Dabei spielen (geometrische) Struktureigenschaften eine zentrale Rolle.

In der Formulierung von Bezeichnung 1.2.1 gehen die Variablen linear ein. Das ist allerdings nur unter bestimmten Bedingungen in der Praxis der Fall. So sinkt etwa in der Nähe der Marktsättigung der pro zusätzlicher Einheit eines Produkts erzielbare Gewinn. In der Ökonomie spricht man von einem *abnehmenden Grenzertrag*.

Umgekehrt kann es auch wichtig sein, *zunehmende Grenzkosten* zu berücksichtigen, wenn durch zunehmende Nachfrage die Preise *knapper Ressourcen* und damit die Kosten der Produktion steigen. Abbildung 1.3 zeigt einige mögliche Typen von (eindimensionalen) Zielfunktionen; aber natürlich können durchaus auch wesentlich kompliziertere Zielfunktionen auftreten.



1.3 Abbildung. Verschiedene mögliche Typen von Zielfunktionen. Die schwarzen Punkte geben den jeweiligen Funktionswert an Sprungstellen an.

Ebenso kann es sein, dass der Einsatz von Ressourcen nichtlinear von der verbrauchten Menge abhängt. Wenn etwa eine große Menge eines Produkts L_1 produziert werden soll, kann es sinnvoll sein, größer ausgelegte Maschinen einzusetzen. Hierdurch kann sich sowohl die erforderliche Maschinenlaufzeit als auch der Bedarf an erforderlicher handwerklicher Arbeitsleistung reduzieren.

Somit können je nach Modellannahmen sowohl in der Zielfunktion als auch in den Nebenbedingungen Funktionen unterschiedlicher Typen auftreten, und wir sind mit nichtlinearen Optimierungsproblemen unterschiedlicher Art konfrontiert.

Ernährungsplanung: Eine Hundezüchterin möchte ihre Bernhardinerwelpen ausgewogen ernähren und ihnen die geeigneten Mengen an Eiweiß, Fett, Kohlenhydraten, Vitaminen, Spurenelementen etc. zuführen. Sie kann verschiedene Futtermittel F_1, \dots, F_n kaufen; der Preis für eine Einheit von F_j ist γ_j für $j = 1, \dots, n$. Die Futtermittel enthalten die erforderlichen Nährstoffe N_1, \dots, N_m in unterschiedlichen Mengen. Für $i = 1, \dots, m$ und $j = 1, \dots, n$ sei etwa $\alpha_{i,j}$ die Menge des Nährstoffs N_i in F_j . Einerseits möchte die Züchterin sparsam wirtschaften, d.h. die Gesamtkosten der Futtermittel minimieren. Andererseits soll für jedes $i = 1, \dots, m$ eine untere Schranke β_i für den Mindestgehalt an Nährstoff N_i eingehalten werden, den das aus F_1, \dots, F_n zusammengestellte Futter enthalten soll.⁹ Insgesamt ergibt sich folgende Aufgabe:

$$\begin{aligned} \min \quad & \sum_{j=1}^n \gamma_j \xi_j \\ \sum_{j=1}^n \alpha_{i,j} \xi_j & \geq \beta_i \quad (i = 1, \dots, m) \\ \xi_j & \geq 0 \quad (j = 1, \dots, n). \end{aligned}$$

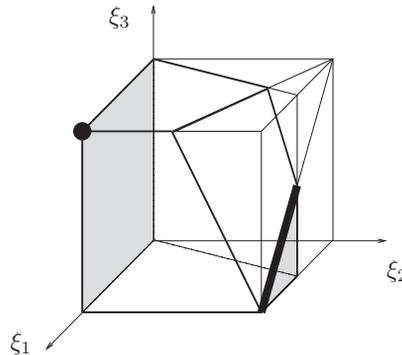
In dieser Formulierung ist vorausgesetzt, dass F_i *homogen* ist, d.h. in beliebigen Mengen eingekauft werden kann. Unter dieser Voraussetzung hat die Aufgabe die gleiche Struktur wie die Produktionsaufgabe des vorherigen Beispiels. Ist dieses jedoch nicht der Fall, weil etwa nur Großpackungen abgegeben werden, gibt es also feste Quantitätsschranken, so gehen diese als weitere Nebenbedingungen an die Mengen ξ_i in die Formulierung der Aufgabe ein. Bezeichnet etwa ξ_2 die Anzahl der zu beschaffenden Großpackungen

⁹ Ggf. können auch obere Schranken ergänzt werden, um etwa Mineralisierungsstörungen durch eine Überversorgung mit Kalzium und Phosphor zu verhindern.

eines (nur im betrachteten Ernährungszeitraum haltbaren) Frischfuttermittels im möglichen Einkauf der Züchterin, so kommt die Bedingung $\xi_2 \in \mathbb{N}_0$ hinzu.¹⁰ Die vorher rein lineare Optimierungsaufgabe wird zu einer *gemischt-ganzzahligen Aufgabe*.

Gemischt-ganzzahlige Optimierungsprobleme sind im Allgemeinen erheblich schwieriger zu lösen als rein lineare, und wir werden verschiedene Methoden besprechen, wie man die Nebenbedingung der Ganzzahligkeit in den Griff bekommen kann.

In dem einfachen einführenden Beispiel 1.2.2 sieht man bereits, welchen Einfluss etwa die Ganzzahligkeitsbedingung für die zweite Variable hat. Die vorher gefundene Lösung $(1, \frac{1}{2}, 1)^T$ der linearen Aufgabe ist nun nicht mehr zulässig. Tatsächlich sind, wenn die zweite Variable ganzzahlig ist, nur noch die beiden in Abbildung 1.4 grau unterlegten Flächen zulässig. Es gibt viele Lösungen mit Zielfunktionswert 2, aber keine besseren.



1.4 Abbildung. Der zulässige Bereich (grau) der Aufgabe aus Beispiel 1.2.2 unter der zusätzlichen Bedingung $\xi_2 \in \mathbb{Z}$. Die Menge der Optimalpunkte besteht aus der schwarz hervorgehobenen Strecke und dem markierten Punkt.

Fordert man, dass *alle* Variablen ganzzahlig sein müssen, so sind

$$(0,0,0)^T, (0,0,1)^T, (1,0,0)^T, (1,1,0)^T, (1,0,1)^T$$

die einzigen zulässigen Lösungen. Optimal sind dann die Punkte $(1,1,0)^T$ und $(1,0,1)^T$.

Standortprobleme: Ein amerikanischer Großhandelskonzern plant den Aufbau eines Vertriebssystems in Europa. Verträge mit Großkunden G_1, \dots, G_n sind bereits geschlossen, mögliche Standorte S_1, \dots, S_m für Versandgroßlager sind in der Beurteilung. Es soll eine optimale Auswahl der Standorte getroffen werden. Dabei treten natürlich Fixkosten κ_j für die Errichtung eines Versandlagers am Standort S_j auf. Andererseits sind für $i = 1, \dots, n$ und $j = 1, \dots, m$ die Transportkosten $\gamma_{i,j}$ bekannt, die auftreten, wenn man den Großkunden G_i vom Standort S_j aus beliefert. Ferner ist die Gesamtlieferungsmenge α_i an G_i vertraglich festgelegt sowie die Kapazität β_j des möglichen Lagers im Standort S_j gegeben.

Wir führen Variablen $\xi_{i,j}$ ein, die angeben sollen, wieviel von Standort S_j aus an den Kunden G_i geliefert werden soll. Da der Bedarf aller Kunden gedeckt werden muss, ergeben sich die Bedingungen

¹⁰ Es kann durchaus erforderlich sein, unterschiedliche Preise bei unterschiedlichen Großpackungsgrößen zu berücksichtigen. Das ist aber innerhalb des angegebenen Modells möglich, indem diese als verschiedene Futtermittel aufgefasst werden, mit gleichen Inhaltsstoffen, aber in unterschiedlichen Packungsgrößen und mit unterschiedlichen Preisen.

$$\begin{aligned} \sum_{j=1}^m \xi_{i,j} &= \alpha_i & (i = 1, \dots, n) \\ \xi_{i,j} &\geq 0 & (i = 1, \dots, n; j = 1, \dots, m). \end{aligned}$$

Ferner modellieren wir die Entscheidung, ein Versandgroßlager am Standort S_j zu errichten, mit Hilfe einer 0-1-Variable η_j , d.h.

$$\eta_j \in \{0,1\} \quad (j = 1, \dots, m).$$

Die Kapazitätsbeschränkungen des möglichen Lagers in S_j führen dann auf die Bedingung

$$\sum_{i=1}^n \xi_{i,j} \leq \beta_j \eta_j \quad (j = 1, \dots, m).$$

Während die Variablen $\xi_{i,j}$ die potentiellen Liefermengen angeben, modellieren die 0-1-Variablen η_j , dass nur dann eine Lieferung von S_j aus erfolgen kann, wenn an diesem Standort tatsächlich ein Lager errichtet wird. Als Zielfunktion wird die Minimierung der Kosten über den Planungshorizont angesetzt, d.h.

$$\min \sum_{i=1}^n \sum_{j=1}^m \gamma_{i,j} \xi_{i,j} + \sum_{j=1}^m \kappa_j \eta_j$$

Es liegt insgesamt also eine gemischt-ganzzahlige lineare Optimierungsaufgabe in den reellen Variablen $\xi_{i,j}$ und den *Entscheidungsvariablen* η_j vor.

Natürlich kann man dieses Modell auf vielfältige Weise modifizieren, um es anderen Gegebenheiten anzupassen. So kann man etwa leicht modellieren, dass die Lieferungen an einen Großkunden G_i nur von einem Lager aus erfolgen sollen, indem man die entsprechenden Variablen $\xi_{i,j}$ den diskreten Bedingungen

$$\xi_{i,j} \in \{0, \alpha_i\} \quad (j = 1, \dots, m)$$

unterzieht. Durch Ersetzung von $\xi_{i,j}$ durch $\alpha_i \cdot \eta_{i,j}$ gelangt man wieder zu den üblichen 0-1-Entscheidungsvariablen.

Das Zuordnungsproblem: In einer Partnervermittlungsgesellschaft stehen vier Personen P_1, P_2, P_3, P_4 vom Geschlecht Γ_1 und drei Personen Q_1, Q_2, Q_3 vom Geschlecht Γ_2 mit $\Gamma_1 \neq \Gamma_2$ zur Bildung von (getrennt geschlechtlichen) Paaren zur Verfügung. Die Agentur hat zu jedem der sieben Kandidaten ein Profil erstellt, auf dessen Grundlage für jedes der zwölf möglichen Paare (P_i, Q_j) mit $i = 1, 2, 3, 4$ und $j = 1, 2, 3$ ein Kompatibilitätskoeffizient $\gamma_{i,j} \in \mathbb{R}$ ermittelt wurde, der angibt, wie gut die entsprechenden Profile zueinander passen. Ziel ist es, eine solche Zuordnung zu finden, die die Summe der entsprechenden Koeffizienten maximiert.

Ähnliche Fragestellungen ergeben sich bei der Vergabe von Aufgaben an Arbeitskräfte, bei der Platzierung von Objekten auf Regalen in Packstationen und in vielen anderen Kontexten (bisweilen unter weiteren zusätzlichen Bedingungen).

Formal ist für $n_1, n_2 \in \mathbb{N}$ eine Teilmenge

$$Z \subset \{1, \dots, n_1\} \times \{1, \dots, n_2\}$$

gesucht, für die einerseits

$$\begin{aligned} (i, j_1) \in Z \wedge (i, j_2) \in Z &\Rightarrow j_1 = j_2 \\ (i_1, j) \in Z \wedge (i_2, j) \in Z &\Rightarrow i_1 = i_2 \end{aligned}$$

gilt, und die andererseits die Summe

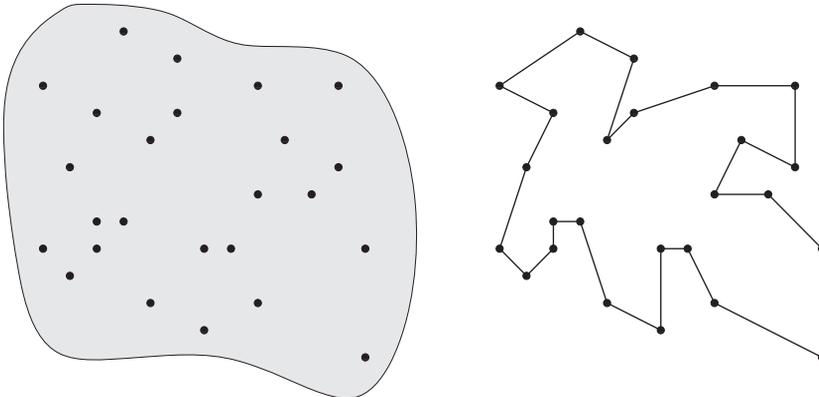
$$\sum_{(i,j) \in Z} \gamma_{i,j}$$

der ‘Einzelnutzen’ $\gamma_{i,j}$ für alle ausgewählten Paare (i,j) maximiert. Mit Hilfe von Variablen $\xi_{i,j} \in \{0,1\}$ kann man beschreiben, ob ein gegebenes Paar (i,j) in einer Zuordnung Z enthalten ist ($\xi_{i,j} = 1$) oder nicht ($\xi_{i,j} = 0$). Die Aufgabe, eine optimale Zuordnung zu finden, lässt sich so mit Hilfe der folgenden LP-Aufgabe beschreiben:

$$\begin{aligned} \max \quad & \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \gamma_{i,j} \xi_{i,j} \\ \sum_{j=1}^{n_2} \xi_{i,j} &\leq 1 && (i = 1, \dots, n_1) \\ \sum_{i=1}^{n_1} \xi_{i,j} &\leq 1 && (j = 1, \dots, n_2) \\ \xi_{i,j} &\in \{0,1\} && (i = 1, \dots, n_1; j = 1, \dots, n_2). \end{aligned}$$

Das Problem des Handlungsreisenden: Das als *Traveling Salesman Problem* oder kürzer als TSP bekannte Rundreiseproblem ist ein Standardproblem der kombinatorischen Optimierung. Hierbei ist eine kürzeste Rundreise (Tour) zu finden, die alle vorgegebenen Städte genau einmal besucht.¹¹

Abbildung 1.5 zeigt eine (fiktive) Landkarte und eine Rundreise durch alle Städte. Dabei wird zwischen zwei Städten jeweils der euklidische Abstand zugrunde gelegt.¹²



1.5 Abbildung. Links: Städte eines (fiktiven) Landes; Rechts: Rundreise durch alle Städte.

¹¹ Diese Fragestellung ist so natürlich, dass – wie vor Kurzem herausgefunden wurde – selbst Hummeln bei ihrem Flug von Blüte zu Blüte um eine ‘effiziente TSP-Logistik’ bemüht sind; vgl. *M. Lihoreau, L. Chittka, N.E. Raine* (Travel optimization by foraging bumblebees through readjustment of traplines after recovery of new feeding locations, *The Amer. Naturalist* 176 (2010), 744 – 757).

¹² Das ist etwa bei Hubschraubern als Verkehrsmittel vernünftig; bei Autofahrten in einem vorgegebenen Straßennetz werden im Allgemeinen jedoch andere als ‘Luftlinienentfernungen’ relevant sein.

Natürlich handelt es sich bei der Frage nach einer kürzesten Rundreise ‘nur’ um eine endliche, diskrete Aufgabe. Man könnte also prinzipiell alle möglichen Touren bestimmen und eine kürzeste auswählen. Die folgende Bemerkung zeigt, dass dieser enumerative Ansatz schon für eine moderate Anzahl von Städten nicht praktikabel ist.

1.2.3 Bemerkung. Für $n \in \mathbb{N} \setminus \{1, 2\}$ sei τ_n die Anzahl aller verschiedenen Rundreisen durch n Städte. Dann gilt

$$\tau_n = \frac{1}{2}(n-1)! > \sqrt{\pi/2}(n-1)^{n-\frac{1}{2}}e^{-n+1}.$$

Beweis: Da die Rundreise geschlossen ist, und es keine Rolle spielt, an welchem Punkt wir starten, beginnen wir bei einer beliebigen Stadt. Nach dem Start stehen als nächste zu besuchende Stadt dann alle $n-1$ anderen zur Verfügung. Wählt man eine aus, so bleiben danach $n-2$ Möglichkeiten fortzufahren. Danach stehen nur noch $(n-3)$ Städte zur Verfügung, und entsprechend geht es weiter. Berücksichtigt man noch, dass durch diese Zählweise jede Tour doppelt gezählt wird, da es keinen Unterschied macht, ob eine Tour ‘vorwärts’ oder ‘rückwärts’ durchlaufen wird, folgt $\tau_n = (n-1)!/2$. Aus der *Stirlingschen*¹³ Formel

$$\sqrt{2\pi n}n^{n+\frac{1}{2}}e^{-n} < n! \leq \sqrt{2\pi n}n^{n+\frac{1}{2}}e^{-n}e^{\frac{1}{12n}} \quad (n \in \mathbb{N})$$

ergibt sich daher die Behauptung. \square

Der Term $(n/e)^n$ in der unteren Schranke von Bemerkung 1.2.3 zeigt, dass τ_n exponentiell in n wächst. Für $n = 21$ ist etwa

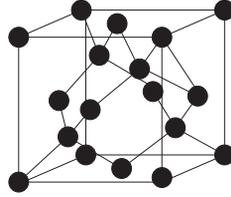
$$\tau_n = (n-1)!/2 = 1.216.451.004.088.320.000 > 1,2 \cdot 10^{18}.$$

Selbst wenn ein Teraflop-Rechner pro Sekunde 10^{12} Touren berechnen könnte, so dauerte ein solches Enumerationsverfahren bereits bei nur 21 Knoten etwa zwei Wochen. Bei 25 Knoten (wie in Abbildung 1.5) brauchte man $21 \cdot 22 \cdot 23 \cdot 24$ mal so lange, also fast 10.000 Jahre. Bei noch größeren, aber immer noch sehr moderaten Knotenzahlen läge selbst dann noch kein Ergebnis vor, wenn sogar auf jedem Elementarteilchen des bekannten Universums ein Teraflop-Rechner seit dem Urknall pausenlos arbeitete. Dieses Phänomen nennt man *kombinatorische Explosion*. Sie ist der Grund dafür, dass ‘endliche Optimierungsprobleme’ im Allgemeinen durchaus nicht trivial sind und man wesentlich ausgeklügeltere Verfahren als bloße Enumeration benötigt.

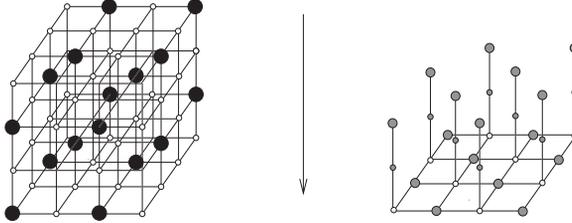
Bemerkung 1.2.3 zeigt insbesondere auch, dass selbst, wenn der zulässige Bereich F einer Optimierungsaufgabe, der beim TSP ja aus allen τ_n Touren besteht, endlich ist, nicht erwartet werden kann, dass alle Elemente *explizit* vorliegen. Tatsächlich besteht ein wesentliches Problem darin, eine geeignete (oft nur approximative) Darstellung von F zu finden, die die ‘Komplexität’ der expliziten Darstellung vermeidet.

Rekonstruktion kristalliner Strukturen: Bei der Herstellung von Silizium-Chips sollen neue bildgebende Verfahren zur Qualitätskontrolle eingesetzt werden. Mit Hilfe moderner Methoden der hochauflösenden Transmissionselektronenmikroskopie (HRTEM) kann man für bestimmte Materialien (wie etwa Silizium; vgl. Abbildung 1.6) in geeigneten Richtungen bestimmen, wieviele Atome in den entsprechenden Atomsäulen enthalten sind.

¹³ James Stirling, 1692 – 1770.

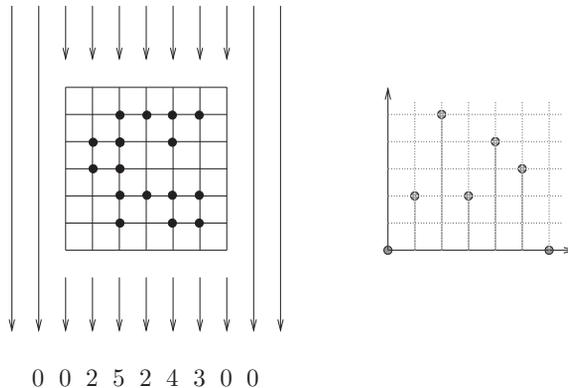


1.6 **Abbildung.** Eine 3-dimensionale ‘Gittermenge’ (Silizium in kristalliner Form).



1.7 **Abbildung.** Links: eine 3-dimensionale Gittermenge; Rechts: Graph der X-ray Messdaten in Richtung der ξ_3 -Koordinatenachse.

Wir nehmen (der Einfachheit der Darstellung halber) an, dass die zugrunde liegende kristalline Gitterstruktur der von \mathbb{Z}^3 entspricht und die zu untersuchende Probe im Bereich $[1, q]^3$ liegt mit $q \in \mathbb{N}$. Ferner seien die Richtungen der Aufnahmen parallel zu den drei Koordinatenachsen im \mathbb{R}^3 . Dann erhält man (bis auf gewisse Messfehler) die Informationen, wieviele Atome jeweils auf jeder Geraden parallel zu einer der Koordinatenachsen liegen. Abbildung 1.7 zeigt eine kristalline Struktur zusammen mit ihrem HRTEM-Bild in Richtung der ξ_3 -Koordinatenachse. Abbildung 1.8 enthält eine analoge Menge im 2-dimensionalen Gitter \mathbb{Z}^2 .



1.8 **Abbildung.** Links: eine 2-dimensionale Gittermenge; Rechts: Graph der X-ray Daten in Richtung der ξ_2 -Koordinatenachse.

Die gegebene Aufgabe der *diskreten Tomographie* besteht in der Rekonstruktion der zugrunde liegenden Gittermenge aus den gegebenen ‘X-ray Daten’. Zur mathematischen

Modellierung führen wir für jeden potentiellen Atommittelpunkt $(i,j,k)^T \in [1,q]^3$, d.h. für jedes der ‘Indextripel’ (i,j,k) ($i,j,k = 1, \dots, q$) eine Variable

$$\xi_{i,j,k} \in \{0,1\}$$

ein, die angibt, ob die betreffende Position durch ein Atom besetzt ist ($\xi_{i,j,k} = 1$) oder nicht ($\xi_{i,j,k} = 0$). Nebenbedingungen erhält man aus den Messwerten. Genauer liegt jeweils für jede Gerade parallel zu $u_1 := (1,0,0)^T$, $u_2 := (0,1,0)^T$ oder $u_3 := (0,0,1)^T$ durch einen Gitterpunkt in $[1,q]^3$ eine Messung vor. Die einzelnen Messdaten werden durch die Zahlen $\alpha_{j,k}, \beta_{i,k}, \gamma_{i,j} \in \mathbb{N}_0$ ($i,j,k = 1, \dots, q$) angegeben. Wenn wir annehmen, dass die Daten messfehlerfrei sind, erhalten wir damit folgende Bedingungen:

$$\begin{aligned} \sum_{i=1}^q \xi_{i,j,k} &= \alpha_{j,k} & (j,k = 1, \dots, q) \\ \sum_{j=1}^q \xi_{i,j,k} &= \beta_{i,k} & (i,k = 1, \dots, q) \\ \sum_{k=1}^q \xi_{i,j,k} &= \gamma_{i,j} & (i,j = 1, \dots, q). \end{aligned}$$

Zusammen mit den 0-1-Bedingungen an die Variablen beschreibt dieses System alle kristallinen Strukturen, die mit den gegebenen Messdaten verträglich sind. Natürlich werden im Allgemeinen drei Messrichtungen nicht ausreichen, um eine eindeutige Rekonstruktion zu erlauben. Die Modellierung von mehr als drei Messrichtungen kann aber ganz analog erfolgen.

Zusätzlich zu den Schwierigkeiten, die die Lösung solcher 0-1-Optimierungsaufgaben macht – in der Praxis hat man etwa 10^9 Variable –, treten noch weitere Komplikationen auf, die damit zusammenhängen, dass sich die Lösungen im allgemeinen unter leichten Änderungen der rechten Seite, d.h. insbesondere unter Messfehlern, nicht stabil verhalten. Es handelt sich um ein sogenanntes *schlecht-gestelltes diskretes inverses Problem*. Verzichtet man auf die Ganzzahligkeit der Daten und fordert stattdessen nur

$$0 \leq \xi_{i,j,k} \leq 1,$$

so liegt eine einfacher zu lösende lineare Optimierungsaufgabe vor. Der Übergang zu reellen Variablen ist etwa dann sinnvoll, wenn es sich bei den einzelnen Punkten nicht um Atome, sondern – auf einer viel größeren Skala – etwa um Gewebedichten einer Region handelt. Tatsächlich sind bei den ersten Ansätzen der Computertomographie genau solche linearen Aufgaben gelöst worden. Hierfür erhielten Hounsfield und Cormack¹⁴ 1979 den Nobelpreis.

Wire Spacing:¹⁵ Ein Halbleiterchip besteht aus einer Grundsicht, auf der Transistoren und andere elektronische Bauteile platziert sind, sowie zusätzlichen Schichten, die Verbindungsdrähte enthalten. Die verschiedenen Schichten sind durch sogenannte Vias verbunden. Die Drähte einer Leitungsschicht verlaufen parallel und senkrecht zu den Nachbarschichten. Bei dem heutzutage erreichten Grad an Miniaturisierung tragen die ‘Streukapazitäten’ zwischen benachbarten Drähten wesentlich zur Hitzeentwicklung

¹⁴ Godfrey Newbold Hounsfield, 1919 – 2004; Allan McLeod Cormack, 1924 – 1998.

¹⁵ Das Anwendungsbeispiel dieses Abschnitts stammt aus *P. Grützmann, M. Ritter, P. Zuber* (Optimal wire ordering and wire spacing in low power semiconductor design, Math. Prog. 121 (2010), 201–220).

in den integrierten Schaltungen bei. Diese treten immer dann auf, wenn Ladungsänderungen in einem der Drähte auftreten, da diese dann wie Kondensatoren wirken. Die ‘Umschalthäufigkeit’ eines Drahtes w wird als eine positive reelle Zahl $\sigma = \sigma(w)$ modelliert. Die Umschaltung erfolgt zwischen der Spannung 0 und der (konstanten) Arbeitsspannung U . Benachbarte Drähte wirken als Kondensator, und zum Aufbau des entsprechenden elektrischen Feldes wird eine Energie benötigt, die direkt proportional zu seiner Kapazität ist. Diese wiederum ist proportional zum Quotienten der Oberfläche und dem Abstand der beiden betroffenen Leitungen. Nimmt man an, dass die Abmessungen der Drähte fest sind, so hängt der ‘Stromverlust’ eines Drahtes w pro Zeiteinheit nur von den Abständen d_{links} und d_{rechts} zu seinen beiden Nachbardrähten ab, und zwar in der Form

$$\sigma(w) \cdot \left(\frac{1}{d_{\text{links}}} + \frac{1}{d_{\text{rechts}}} \right).$$

Als Modellannahme liegt unter anderem die Voraussetzung zugrunde, dass benachbarte Drähte nicht simultan umschalten. Ferner werden die (ohnehin geringen) Kapazitäten zwischen nicht benachbarten Drähten und verschiedenen Schichten vernachlässigt.

Die Aufgabe besteht nun darin, bei bekannten Umschaltaktivitäten $\sigma_1, \dots, \sigma_n$ einer gegebenen Anzahl n paralleler Drähte einer Schicht die Abstände der Drähte so zu wählen, dass die auftretende Gesamtstreukapazität minimiert wird. Gegeben ist ferner ein minimal tolerierbarer Abstand $\delta \in]0, \infty[$ zwischen Drähten und zum Rand des Chips sowie die zur Platzierung der Drähte zur Verfügung stehende Gesamtbreite $\rho \in]0, \infty[$ des Chips. Wir führen für die Abstände der Drähte Variablen ξ_1, \dots, ξ_{n+1} ein. Dabei gibt ξ_1 den Abstand des linken Randes des Chips zu Draht 1, ξ_{n+1} den Abstand von Draht n zum rechten Rand und ξ_i für $i = 2, \dots, n$ den Abstand zwischen Draht $i - 1$ und Draht i an.

Die physikalisch-mathematische Modellierung führt somit (unter den entsprechenden Modellannahmen) zu folgendem Optimierungsproblem.

1.2.4 Bezeichnung. *Eine Wire Spacing Aufgabe ist spezifiziert durch folgende Daten*

$$n \in \mathbb{N} \quad \wedge \quad \sigma_1, \dots, \sigma_n \in [0, \infty[\quad \wedge \quad \rho \in]0, \infty[\quad \wedge \quad \delta \in]0, \infty[.$$

Die Aufgabe besteht darin zu entscheiden, ob

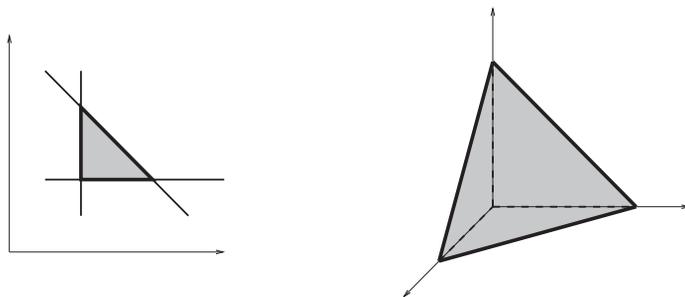
$$\begin{aligned} \min \sum_{i=1}^n \sigma_i \left(\frac{1}{\xi_i} + \frac{1}{\xi_{i+1}} \right) \\ \sum_{i=1}^{n+1} \xi_i &\leq \rho \\ \xi_1, \dots, \xi_{n+1} &\geq \delta \end{aligned}$$

einen Optimalpunkt x^ besitzt und, falls das der Fall ist, einen solchen zu bestimmen.*

*Die Menge aller Wire Spacing Aufgaben heißt **Wire Spacing Problem**.*

Der zulässige Bereich solcher Wire Spacing Aufgaben ist ein sehr einfaches Polytop; vgl. Abbildung 1.9. Die auftretende Zielfunktion ist aber nicht linear. Dennoch lässt sich (wie wir später sehen werden) das Problem effizient lösen.

Eine ‘klassische ganzzahlige Optimierungsaufgabe’: Für $n \in \mathbb{N}$ sei das folgende Optimierungsproblem im \mathbb{R}^3 gegeben:



1.9 Abbildung. Links: Zulässiger Bereich der Wire Spacing Aufgabe für $n = 2$, $\delta = 1$ und $\rho = 3$. Rechts: Der zulässige Bereich für $n = 3$, $\rho = 2$ und (der praktisch nicht zulässigen Wahl) $\delta = 0$.

$$\begin{aligned} \min (\xi_1^n - \xi_2^n + \xi_3^n)^2 \\ \xi_1, \xi_2, \xi_3 &\geq 1 \\ \xi_1, \xi_2, \xi_3 &\in \mathbb{Z}; \end{aligned}$$

$\zeta^*(n)$ bezeichne das Optimum der Zielfunktion. Es gilt $\zeta^*(n) \geq 0$ und $\zeta^*(1) = 0$. Schon vor mehr als 3600 Jahren war bekannt, dass $\zeta^*(2)$ ebenfalls 0 ist; entsprechende Lösungen finden sich auf babylonischen Tontafeln¹⁶. Keine geringeren als *Euler*, *Dirichlet* und *Legendre*¹⁷ zeigten, dass $\zeta^*(n) \geq 1$ ist für $n = 3, 4, 5$.

Offenbar ist $\zeta^*(n) = 0$ genau dann, wenn die Gleichung

$$\xi_1^n + \xi_3^n = \xi_2^n$$

eine positive ganzzahlige Lösung besitzt. Dass dieses für kein $n \geq 3$ der Fall ist, hat im 17. Jahrhundert bereits *Pierre de Fermat*¹⁸ behauptet; der *Satz von Fermat* konnte aber bekanntlich erst 1995 von *Wiles* bzw. *Wiles und Taylor*¹⁹ bewiesen werden.²⁰

Natürlich denkt beim ‘großen Fermat’ eigentlich niemand an Optimierung. Interessant von Seiten der Optimierung ist aber, dass es sich bei der Zielfunktion für jedes n um ein Polynom vom Grad $2n$ handelt, also eine beliebig oft differenzierbare, eigentlich doch sehr einfach strukturierte Funktion. Während sich die vorher vorgestellten Probleme entweder als effizient lösbar erweisen werden oder ihre Schwierigkeit erst in relativ hohen Dimensionen entfalten, führt der Satz von Fermat für festes n auf eine Optimierungsaufgabe der Dimension 3. Es zeigt sich also, dass nichtlineare diskrete Optimierungsaufgaben selbst in kleinen Dimensionen schwierig werden können und keine einfache allgemeine Charakterisierung der Optimalität der Zielfunktion erlauben.

Der ‘volle Fermat’ für alle Dimensionen ist äquivalent zu der Aussage, dass das Minimum des nichtlinearen diskreten Optimierungsproblems

¹⁶ Siehe <http://www.math.ubc.ca/~cass/courses/m446-03/pl322/pl322.html>; später wurden die Lösungen für $n = 2$ *pythagoreische Tripel* genannt.

¹⁷ Leonhard Euler, 1707 – 1783; Peter Gustav Lejeune Dirichlet, 1805 – 1859; Adrien-Marie Legendre, 1752 – 1833.

¹⁸ Pierre de Fermat, 1601 – 1665.

¹⁹ Andrew Wiles, geb. 1954; Richard Taylor, geb. 1962.

²⁰ Selbst die *New York Times* und *The Simpsons* (in der Halloween Episode 1995) haben den Durchbruch (auf durchaus unterschiedliche Weise) gefeiert. Die aufregende Geschichte des Beweises wird im Buch ‘Fermats letzter Satz: Die abenteuerliche Geschichte eines mathematischen Rätsels’ von *Simon Singh* erzählt (dtv, München 2000).

$$\begin{aligned} \min (\xi_1^\eta - \xi_2^\eta + \xi_3^\eta)^2 \\ \xi_1, \quad \xi_2, \quad \xi_3 &\geq 1 \\ &\eta \geq 3 \\ \xi_1, \quad \xi_2, \quad \xi_3, \quad \eta &\in \mathbb{Z} \end{aligned}$$

mindestens 1 ist; es ‘lebt’ in der Dimension 4.

1.3 Standardnotation und Notationsstandards

Wir verwenden im Folgenden nach Möglichkeit die in den einzelnen Teildisziplinen der Optimierung übliche Notation. Da wir eine durchgängig einheitliche Bezeichnung anstreben, die Notation in der Literatur aber nicht immer einheitlich ist und Standards auch noch von Teilgebiet zu Teilgebiet variieren, ist das nur eingeschränkt möglich. Wir geben daher in dieser Sektion kurz die im vorliegenden Text allgemein verwendeten Standardbezeichnungen und Konventionen an. Spezifischere Bezeichnungen, Definitionen und Konzepte werden hingegen in den jeweiligen Kapiteln eingeführt, in denen sie zum ersten mal relevant werden.²¹

Zur Strukturierung der Darstellungen benutzen wir die Begriffe Definition, Bezeichnung, Notation (sowie in dieser Sektion auch Konvention), Satz, Korollar, Lemma, Bemerkung, Wiederholung, Prozedur, Beispiel, Übungsaufgabe und Forschungsproblem. Dabei verwenden wir eine einheitliche, durchgängige Nummerierung aus drei Ziffern, der Nummer des Kapitels, der Nummer der Sektion und einer laufenden Nummer.

Definitionen werden vorrangig zur Einführung von Konzepten, Bezeichnung zur Zuweisung von Benennung und Abkürzungen verwendet.²² Bisweilen werden die eingeführten deutschen Begriffe um die entsprechenden englischen Termini ergänzt. Im Gegensatz zu Definitionen und Bezeichnungen, die global verwendet werden, also im ganzen Buch gelten, sind die unter ‘Notation’ angegebenen Größen nur lokal, d.h. für die spezielle Sektion relevant.

Die Klassifikation eines Results als Satz, Korollar, Lemma oder Bemerkung erfolgt nach den Gesichtspunkten ihrer Bedeutung und Tiefe. Bemerkungen sind einfach zu beweisende, ‘klare’ Aussagen, Lemmata haben oftmals einen vorbereitenden, technischen Charakter, Sätze und Korollare hingegen enthalten die wesentlichen Aussagen, die häufig auch in anderen Kapiteln von Bedeutung sind. Oftmals hängt es von der subjektiven Gewichtung ab, welches Ergebnis ein Satz und welches eine daraus abgeleitete Folgerung, ein Korollar ist.

Forschungsprobleme sind ungelöste Fragen von zentraler Bedeutung. Sie werden angegeben, um den Stand der Forschung zu erläutern.²³

Algorithmen werden in strukturierter Form (durchaus mit unterschiedlichem Detaillierungsgrad), nicht jedoch als implementierbarer Code angegeben; sie werden als Prozedur gekennzeichnet.

Im Folgenden führen wir einige Konventionen ein (von denen manche vielleicht schon in den vorherigen Sektionen aufgefallen sind.)

²¹ Des einfacheren Zugangs wegen werden zusätzlich einige der folgenden Bezeichnungen später noch einmal wiederholt.

²² Der Übergang ist fließend und die Wahl des Begriffs oft durchaus subjektiv.

²³ Dass ein Problem noch offen ist, heißt nicht zwangsläufig, dass es keine elegante, einfache Lösung gibt. Man sollte sich die offenen Fragen also ruhig einmal genauer ansehen.

1.3.1 Konvention. *Definierende Setzungen werden mit $:=$ bzw. $=:$ gekennzeichnet; der Doppelpunkt steht auf der Seite des zu definierenden Objekts. Sind also etwa $\alpha_1, \alpha_2 \in \mathbb{R}$ gegeben, so wird durch die Setzung $a := (\alpha_1, \alpha_2)^T$ der Spaltenvektor dieser Komponenten definiert. Ist umgekehrt der Vektor a gegeben, so bedeutet die Setzung $a =: (\alpha_1, \alpha_2)^T$, dass α_1, α_2 seine erste und zweite Komponente sind. Analog wird die Bezeichnung $:\Leftrightarrow$ verwendet.*

Oftmals werden wir Setzungen sprachlich etwas verkürzt einführen. Statt der korrekten, aber etwas umständlichen Formulierung ‘Seien A eine nichtleere, endliche Teilmenge von \mathbb{R} , $m \in \mathbb{N}$ und $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ mit $A = \{\alpha_1, \dots, \alpha_m\}$ ’ schreiben wir meistens einfacher ‘Sei $A := \{\alpha_1, \dots, \alpha_m\} \subset \mathbb{R}$ ’ oder ‘Sei $A =: \{\alpha_1, \dots, \alpha_m\} \subset \mathbb{R}$ ’, je nachdem, ob wir $\alpha_1, \dots, \alpha_m$ oder A als primär auffassen. In Definitionen oder Bezeichnungen werden wir häufiger – wie allgemein üblich – der leichteren Lesbarkeit wegen die präzise Form ‘wenn, und nur wenn’ bzw. ‘dann, und nur dann’ etc. durch ‘wenn’ bzw. ‘dann’ abkürzen. Da niemals eine andere Interpretation auftritt, sollte es hierdurch zu keinen Missverständnissen kommen.

1.3.2 Bezeichnung. *Die aussagenlogischen Verknüpfungen **Konjunktion**, **Disjunktion**, **Negation**, **Implikation** (von links nach rechts bzw. umgekehrt) und **Äquivalenz** werden mit*

$$\wedge, \vee, \neg, \Rightarrow, \Leftarrow, \Leftrightarrow$$

bezeichnet. Wir verwenden auch die **All-** bzw. **Existenzquantoren**

$$\bigwedge, \bigvee, \forall, \exists.$$

Die ersten beiden treten allerdings nur in den entsprechenden Formeln in Sektion 3.3 auf. Oftmals werden wir eine für mehrere Objekte geltende Aussage durch Nachstellung ihres Gültigkeitsbereichs abgekürzt angeben. Statt etwa $\bigwedge_{i \in I} \xi_i \geq 0$ oder $\forall(i \in I) : \xi_i \geq 0$ bzw. $(i \in I \Rightarrow \xi_i \geq 0)$ schreiben wir dann auch einfacher $\xi_i \geq 0$ ($i \in I$).

1.3.3 Bezeichnung. \mathbb{N} , \mathbb{N}_0 , \mathbb{Z} , \mathbb{Q} und \mathbb{R} bezeichnen die **natürlichen**, **nichtnegativen ganzen**, **ganzen**, **rationalen** und **reellen Zahlen**.²⁴ Für $n \in \mathbb{N}$ ist

$$[n] := \{1, \dots, n\}.$$

Intervalle werden mit eckigen Klammern bezeichnet. Dabei sind für $\alpha, \beta \in \mathbb{R}$

$$\begin{aligned} [\alpha, \beta] &:= \{\xi \in \mathbb{R} : \alpha \leq \xi \leq \beta\} \quad \wedge \quad [\alpha, \beta[:= \{\xi \in \mathbb{R} : \alpha \leq \xi < \beta\} \\]\alpha, \beta] &:= \{\xi \in \mathbb{R} : \alpha < \xi \leq \beta\} \quad \wedge \quad]\alpha, \beta[:= \{\xi \in \mathbb{R} : \alpha < \xi < \beta\}. \end{aligned}$$

Ferner bedeutet die Schreibweise $i = 1, \dots, k$, dass i alle Werte in $[1, k] \cap \mathbb{Z}$ annimmt. Für $k < 1$ ist dabei natürlich $[1, k] = \emptyset$.

Für $p \in \mathbb{N} \setminus \{1\}$ bezeichnen $\equiv (\bmod p)$ die Äquivalenzrelation **modulo** p auf \mathbb{Z} und \mathbb{Z}_p den **Restklassenring modulo** p . Ist M eine endliche Menge, so bezeichnet $|M|$ ihre **Kardinalität**, d.h. die Anzahl der Elemente von M .

Im Folgenden geben wir die verwendete Bezeichnungskonvention an.

²⁴ Man beachte, dass abweichend von einer bestehenden DIN-Norm, aber in Übereinstimmung mit weiten Teilen der Mathematik-Community, die natürlichen Zahlen nicht die 0 enthalten.

1.3.4 Konvention. Skalare werden meistens mit kleinen griechischen Buchstaben gekennzeichnet, für Vektoren werden kleine lateinische Buchstaben verwendet. Ihre Komponenten sind dann passende²⁵ griechische Buchstaben. Wir schreiben also etwa

$$x := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

für den Vektor $x \in \mathbb{R}^n$, der durch Auflistung seiner Komponenten ξ_1, \dots, ξ_n definiert wird. Ohne weitere Kennzeichnung sind Vektoren stets Spaltenvektoren. Zeilenvektoren werden durch Transposition gekennzeichnet. So ist also $x^T = (\xi_1, \dots, \xi_n)$. Die **Dimension** von (Unter-) Vektorräumen U wird mit $\dim(U)$ bezeichnet.

Die verwendeten Standardvektorräume sind \mathbb{R}^n , \mathbb{Q}^n mit $n \in \mathbb{N}$. Daneben tritt auch noch der Modul \mathbb{Z}^n auf.

Das Standardskalarprodukt von zwei Vektoren $a, b \in \mathbb{R}^n$ wird als $a^T b$ geschrieben. Ist X ein linearer Teilraum von \mathbb{R}^n , so bezeichnet X^\perp das **orthogonale Komplement** von X in \mathbb{R}^n .

Mengen und Teilräume werden mit großen lateinischen Buchstaben benannt, ebenso Matrizen.

Die Letter m, n werden meistens für natürliche Zahlen und Dimensionen verwendet; Indizes werden oft mit i, j, k bezeichnet.²⁶

Ungleichungen zwischen Vektoren wie $Ax \leq b$ sind stets komponentenweise zu verstehen.

1.3.5 Bezeichnung. Seien $X, Y \subset \mathbb{R}^n$. Dann heißt die durch

$$X + Y := \{x + y : x \in X \wedge y \in Y\}$$

definierte Menge²⁷ **Minkowski**²⁸-**Summe** von X und Y . Ist X einpunktig, d.h. $X =: \{x\}$, so schreiben wir statt $\{x\} + Y$ auch kürzer $x + Y$. Es handelt sich dann um eine **Translation** von Y um den Vektor x .

Für $\lambda \in \mathbb{R}$ sei

$$\lambda \cdot X := \{\lambda x : x \in X\}.$$

Ferner wird $0 \cdot \emptyset = \{0\}$ gesetzt.²⁹ $\lambda \cdot X$ heißt **Dilat** von X um λ . Allgemeiner ist für $S \subset \mathbb{R}$ mit $S \neq \emptyset$ auch

$$S \cdot X := \{\lambda x : x \in X \wedge \lambda \in S\}.$$

Von besonderer Bedeutung ist der Fall, dass S ein Intervall ist. Speziell für $S := [0, 1]$ und $X =: \{x\}$ ist $[0, 1]X$ oder (kürzer) $[0, 1]x$ die Strecke mit den Endpunkten 0 und x .

Gilt $X = -X$, so wird die Menge X **zentralsymmetrisch** (zum Ursprung) genannt.

²⁵ Passend bezieht sich lediglich auf den 'Phänotyp', d.h. ist mnemotechnisch gemeint, nicht auf ihre 'intrinsische' Beziehung. Bei $x := (\xi_1, \dots, \xi_n)^T$ fällt beides zusammen, bei $v := (\nu_1, \dots, \nu_n)^T$ o.ä. mögen es die Gräzisten (und Linguisten) verzeihen.

²⁶ Wir verzichten hier und an manchen anderen Stellen nicht auf übliche Liebe, aber nicht immer völlig konsistente Gewohnheiten. Natürlich sind natürliche Zahlen auch Skalare; dennoch werden sie hier nicht mit kleinen griechischen Buchstaben notiert.

²⁷ Ganz präzise müsste die Setzung $X + Y := \{z : \exists(x \in X \wedge y \in Y) : z = x + y\}$ lauten, um die Menge von der entsprechenden Familie, in der Elemente mehrfach vorkommen, zu unterscheiden. Wir verzichten der einfacheren Lesbarkeit wegen auf diese Präzisierung. Falls im Folgenden Familien auftreten, so wird dieses explizit angegeben.

²⁸ Hermann Minkowski, 1864 – 1909.

²⁹ Diese Setzung mag zunächst seltsam anmuten, erlaubt aber später eine 'glattere' Darstellung relevanter Ergebnisse.

Die folgende Konvention betrifft Sequenzen und Folgen.

1.3.6 Konvention. Elemente von Sequenzen oder Folgen werden oft mit unten stehenden Indizes bezeichnet, etwa $(\alpha_k)_{k \in N}$ für eine gegebene Indexmenge $N \subset \mathbb{N}_0$. Ist N endlich, so sprechen wir von Sequenzen.

Bisweilen verwenden wir auch obenstehende, eingeklammerte Indizes, etwa $(x^{(k)})_{k \in N}$, insbesondere um bei den Komponenten von Vektoren oder Matrizen Doppel- oder gar Dreifachindizes zu vermeiden. (Diese Art der Bezeichnung tritt häufiger dann auf, wenn die Sequenzen oder Folgen durch (iterative) Algorithmen erzeugt werden und wird dann innerhalb eines Algorithmus einheitlich, d.h. für alle auftretenden Größen, verwendet.)

1.3.7 Bezeichnung. Sei $n \in \mathbb{N}$. Dann bezeichnen u_1, \dots, u_n [engl.: standard unit vectors] die **Standardeinheitsvektoren**³⁰ des \mathbb{R}^n . Unabhängig von der Dimension bezeichnet 0 immer das neutrale Element der Addition, sei es als Skalar, als Nullvektor oder als Nullmatrix. Der Vektor $\mathbb{1} := (1, \dots, 1)^T \in \mathbb{R}^n$ mit Komponenten 1 wird durch den ‘Doppelstrich’ besonders hervorgehoben.

Die Menge der reellen $(m \times n)$ -Matrizen wird mit $\mathbb{R}^{m \times n}$ bezeichnet. Analoge Bezeichnungen finden auch für andere Zahlbereiche Verwendung.

Für $\kappa_1, \dots, \kappa_n \in \mathbb{R}$ bezeichnet $\text{diag}(\kappa_1, \dots, \kappa_n)$ die $(n \times n)$ -**Diagonalmatrix** mit κ_i als Diagonaleintrag in der Position i für $i \in [n]$. Speziell ist $E_n := \text{diag}(1, \dots, 1)$ die $(n \times n)$ -**Einheitsmatrix**.

Seien $A := (a_1, \dots, a_m)^T \in \mathbb{R}^{m \times n}$ und $I \subset [m]$. Dann sei

$$A_I := (a_i : i \in I)^T,$$

d.h. A_I entsteht aus A durch Streichen aller Zeilen, deren Indizes in $[n] \setminus I$ liegen.³¹ Teilmatrizen, die aus A durch Streichen von Spalten einer Indexmenge $[n] \setminus J$ entstehen, werden mit A^J bezeichnet. Es ist also $A^J = ((A^T)_J)^T$.

Der **Rang** bzw. die **Determinante** einer Matrix $A \in \mathbb{R}^{n \times n}$ wird mit $\text{rang}(A)$ bzw. $\det(A)$ bezeichnet. Ist A regulär, d.h. gilt $\text{rang}(A) = n$, so bezeichnet A^{-1} die **Inverse** von A .

Ist $A \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit oder positiv semidefinit, so schreiben wir $A \succ 0$ bzw. $A \succeq 0$.

Die folgenden Bezeichnungen betreffen Rundungen.

1.3.8 Bezeichnung. Seien $\lceil \cdot \rceil : \mathbb{R} \rightarrow \mathbb{Z}$, $\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{Z}$ sowie $\langle \cdot \rangle : \mathbb{R} \rightarrow [0, 1[$ für $\eta \in \mathbb{R}$ definiert durch

$$\lceil \eta \rceil := \min\{\gamma \in \mathbb{Z} : \eta \leq \gamma\} \quad \wedge \quad \lfloor \eta \rfloor := \max\{\gamma \in \mathbb{Z} : \gamma \leq \eta\} \quad \wedge \quad \langle \eta \rangle := \eta - \lfloor \eta \rfloor.$$

Dann heißen $\lceil \cdot \rceil$ bzw. $\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{Z}$ bzw. $\langle \cdot \rangle : \mathbb{R} \rightarrow [0, 1[$ die **obere und untere Gauß**³²-**Klammer** bzw. der **fraktionelle Anteil**. $\lfloor \eta \rfloor$ wird auch **ganzzahliger Anteil** von η genannt.

Analog werden diese Begriffe auch auf Vektoren angewendet; sie sind dann komponentenweise zu verstehen. Für $y := (\eta_1, \dots, \eta_n)^T \in \mathbb{R}^n$ sind also

$$\lfloor y \rfloor := (\lfloor \eta_1 \rfloor, \dots, \lfloor \eta_n \rfloor)^T \quad \wedge \quad \lceil y \rceil := (\lceil \eta_1 \rceil, \dots, \lceil \eta_n \rceil)^T \quad \wedge \quad \langle y \rangle := (\langle \eta_1 \rangle, \dots, \langle \eta_n \rangle)^T.$$

³⁰ Man beachte, dass etwa u_1 stets den ersten Standardeinheitsvektor bezeichnet, unabhängig davon, in welcher Dimension wir uns befinden; u_1 kann also (1) , $(1, 0)^T$, $(1, 0, 0)^T$ oder ein anderer Vektor vom Typ $(1, 0, \dots, 0)^T$ sein.

³¹ Natürlich kann man die obige Bezeichnung auch auf $(m \times 1)$ Matrizen, also auf Vektoren anwenden.

³² Johann Carl Friedrich Gauß, 1777 – 1855.

Die folgenden Begriffe betreffen Teilbarkeiten.

1.3.9 Definition. Seien $\eta, \mu \in \mathbb{Z}$. Dann ist η ein **Teiler** von μ , wenn es ein $\tau \in \mathbb{Z}$ gibt mit $\tau \cdot \eta = \mu$. Ist η ein Teiler von μ , so schreibt man auch $\eta | \mu$. Sind $\eta_1, \dots, \eta_n, \kappa \in \mathbb{Z}$, so heißt κ **größter gemeinsamer Teiler** von η_1, \dots, η_n , wenn gilt

$$(\kappa \geq 0) \quad \wedge \quad (\kappa | \eta_1 \wedge \dots \wedge \kappa | \eta_n) \quad \wedge \quad (\tau \in \mathbb{Z} \wedge \tau | \eta_1 \wedge \dots \wedge \tau | \eta_n \Rightarrow \tau | \kappa).$$

Ist κ größter gemeinsamer Teiler von η_1, \dots, η_n , so schreibt man auch $\kappa = \text{ggT}(\eta_1, \dots, \eta_n)$. Sind $\eta, \mu \in \mathbb{Z}$ und gilt $\text{ggT}(\eta, \mu) = 1$, so werden η und μ **teilerfremd** genannt.

Der größte gemeinsame Teiler $\text{ggT}(\eta_1, \dots, \eta_n)$ existiert stets und ist eindeutig bestimmt. Obwohl jede ganze Zahl 0 teilt, folgt aus der dritten Bedingung in Definition 1.3.9, dass $\text{ggT}(0, \dots, 0) = 0$ gilt. Man beachte ferner, dass für $\mu \in \mathbb{N}$ die Zahlen 0 und μ genau dann teilerfremd sind, wenn $\mu = 1$ gilt.

1.3.10 Bezeichnung. S_n bezeichne die **symmetrische Gruppe**, d.h. die Menge der Permutationen auf $[n]$. Sei $\text{sign}: S_n \rightarrow \{-1, 1\}$ die **Signum-Funktion**, d.h. für $\sigma \in S_n$ ist $\text{sign}(\sigma) = 1$, falls σ durch eine gerade Anzahl von Transpositionen in die Identität überführt werden kann, sonst $\text{sign}(\sigma) = -1$.

Die folgende Bezeichnung fasst einige elementare topologische Begriffe zusammen.

1.3.11 Bezeichnung. (Analysis)

(a) **Normen** im \mathbb{R}^n werden mit $\| \cdot \|$ bezeichnet. Für $p \in [1, \infty]$ ist speziell $\| \cdot \|_{(p)}$ die für $x := (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n$ durch

$$\|x\|_{(p)} := \left(\sum_{j=1}^n |\xi_j|^p \right)^{\frac{1}{p}} \quad (p \in [1, \infty[)$$

$$\|x\|_{(\infty)} := \max\{|\xi_1|, \dots, |\xi_n|\}$$

definierte p -Norm. Insbesondere sind $\| \cdot \|_{(1)}$ die **Betragssummennorm**, $\| \cdot \|_{(2)}$ die **euklidische Norm** und $\| \cdot \|_{(\infty)}$ die **Maximumnorm**.

Die jeweils zugehörige **Einheitskugel** wird mit $\mathbb{B}, \mathbb{B}_{(p)}, \mathbb{B}_{(\infty)}$ oder (zur Betonung der Dimension) mit $\mathbb{B}^n, \mathbb{B}_{(p)}^n$ bzw. $\mathbb{B}_{(\infty)}^n$ bezeichnet. Die entsprechenden **Einheits-sphären** sind dann $\mathbb{S}, \mathbb{S}_{(p)}, \mathbb{S}_{(\infty)}, \mathbb{S}^{n-1}, \mathbb{S}_{(p)}^{n-1}$ bzw. $\mathbb{S}_{(\infty)}^{n-1}$. (Man beachte, dass die Sphäre \mathbb{S}^{n-1} im \mathbb{R}^n liegt; das hochgestellte $n-1$ gibt, wie üblich, ihre Dimension an.)

(b) Sind

$$(p, q) \in]1, \infty[^2 \quad \wedge \quad \frac{1}{p} + \frac{1}{q} = 1$$

oder

$$(p, q) \in \{(1, \infty), (\infty, 1)\},$$

so heißen die beiden Normen $\| \cdot \|_{(p)}$ und $\| \cdot \|_{(q)}$ **konjugiert**.

(c) Seien $X \subset \mathbb{R}^n$ und $\| \cdot \|$ eine Norm des \mathbb{R}^n . Dann bezeichnen $\text{int}(X)$, $\text{cl}(X)$ und $\text{bd}(X)$ das **Innere**, den **Abschluss** bzw. den **Rand** von X (bez. der durch $\| \cdot \|$ erzeugten Topologie).³³

³³ Man beachte, dass alle Normen im \mathbb{R}^n äquivalent sind, so dass man hier nicht nach verschiedenen Normen differenzieren muss.

Ohne es immer explizit zu benennen, werden sich alle verwendeten topologische Begriffe auf den *Minkowski-Raum* $(\mathbb{R}^n, \|\cdot\|)$ beziehen.

1.3.12 Konvention. Funktionen werden im Allgemeinen mit kleinen griechischen Buchstaben bezeichnet. Standardfunktionen bilden hier eine Ausnahme; für sie wird die übliche, meist lateinische Notation verwendet. Insbesondere ist \log der **Logarithmus** zur Basis 2 und \ln der Logarithmus zur Basis e .

Sind $\varphi : X \rightarrow Y$ eine Funktion und $S \subset X$ sowie $T \subset Y$, dann bezeichnet $\varphi|_S$ die **Einschränkung** von φ auf S . Ferner sind $\varphi(S) := \{\varphi(x) : x \in S\}$ und $\varphi^{-1}(T) := \{x \in X : \varphi(x) \in T\}$ das **Bild** von S bzw. das **Urbild** von T unter φ .

Sind $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ und $x := (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n$, so wird bisweilen zur ‘Explizierung’ von $\varphi(x)$ statt $\varphi((\xi_1, \dots, \xi_n)^T)$ auch einfacher $\varphi(\xi_1, \dots, \xi_n)$ geschrieben. Analog wird auch für Funktionen auf kartesischen Produkten verfahren.

Ist die Funktion $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ (hinreichend oft) differenzierbar in einem Punkt $x_0 \in \mathbb{R}^n$, so bezeichnet $\varphi'(x_0)$ ihren (als Spaltenvektor geschriebenen) **Gradienten**³⁴ und $\varphi''(x_0)$ ihre **Hesse**³⁵-**Matrix**, d.h.

$$\varphi'(x_0) := \left(\frac{\partial}{\partial \xi_1} \varphi(x_0), \dots, \frac{\partial}{\partial \xi_n} \varphi(x_0) \right)^T \quad \wedge \quad \varphi''(x) := \left(\frac{\partial^2}{\partial \xi_i \partial \xi_j} \varphi(x_0) \right)_{i,j \in [n]}.$$

1.4 Ergänzung: Verschiedene Formen von LP-Aufgaben

In Definition 1.1.5 wurden lineare Optimierungsaufgaben in natürlicher Form eingeführt. Man kann aber durchaus auch LP-Aufgaben betrachten, die formal anders aussehen, etwa auch Gleichungen oder vorzeichenbeschränkte Variablen enthalten. Wir zeigen jetzt, dass alle diese Formen aufeinander zurückgeführt werden können.

1.4.1 Bezeichnung. (a) Eine allgemeine **lineare Optimierungsaufgabe** \mathcal{I} ist durch folgende Daten spezifiziert:

$$\begin{aligned} & \text{opt} \in \{\min, \max\} \\ & m, n \in \mathbb{N} \quad \wedge \quad a_1, \dots, a_m \in \mathbb{R}^n \setminus \{0\} \quad \wedge \quad \beta_1, \dots, \beta_m \in \mathbb{R} \quad \wedge \quad \gamma_1, \dots, \gamma_n \in \mathbb{R} \\ & \text{Partition } \{M_1, M_2, M_3\} \text{ von } [m] \quad \wedge \quad \text{Partition } \{N_1, N_2, N_3\} \text{ von } [n]. \end{aligned}$$

Ferner sei das lineare Funktional $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ für $x \in \mathbb{R}^n$ durch $\varphi(x) := c^T x$ definiert, und es sei F die Menge aller $x := (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n$ mit

$$\begin{aligned} a_i^T x &= \beta_i & (i \in M_1) & \quad (\text{Gleichungsnebenbedingungen}) \\ a_i^T x &\leq \beta_i & (i \in M_2) & \quad (\text{Ungleichungsnebenbedingungen}) \\ a_i^T x &\geq \beta_i & (i \in M_3) & \\ \xi_j &\geq 0 & (j \in N_1) & \quad (\text{Nichtnegativitätsbedingungen}) \\ \xi_j &\leq 0 & (j \in N_2) & \quad (\text{Nichtpositivitätsbedingungen}). \end{aligned}$$

Ziel ist es, φ über F zu optimieren.

³⁴ Diese Schreibweise der Ableitungen wird gewählt, um die Analogie vieler Aussagen zu dem Fall $n = 1$ zu unterstreichen. Dass der Gradient als Spaltenvektor aufgefasst wird, entspricht unserer generellen Konvention.

³⁵ Ludwig Otto Hesse, 1811 -1874.

(b) Gilt

$$M_2 = M_3 = N_2 = N_3 = \emptyset \quad \wedge \quad \text{opt} = \min,$$

so spricht man von einer Aufgabe in **Standardform**. Sind

$$M_1 = N_2 = N_3 = \emptyset \wedge M_2 = \emptyset \quad \wedge \quad \text{opt} = \min$$

bzw.

$$M_1 = N_2 = N_3 = \emptyset \wedge M_3 = \emptyset \quad \wedge \quad \text{opt} = \max,$$

so liegt die Aufgabe in **kanonischer Form** vor.

(c) Die Menge aller (allgemeinen) linearen Optimierungsaufgaben heißt (**allgemeines**) **lineares Optimierungsproblem**. Analog sind auch LP-Probleme in Standard- bzw. kanonischer Form definiert.³⁶

Lineare Optimierungsaufgaben gemäß Definition 1.1.5 liegen in *natürlicher Form* vor; es gilt also entsprechend $M_1 = M_3 = N_2 = N_3 = \emptyset$ und $\text{opt} = \max$.

Wir zeigen nun wie sich die verschiedenen Formen linearer Optimierungsprobleme ineinander überführen lassen. Wir beginnen mit den einfachsten Fällen.

1.4.2 Bemerkung. (a) Die Ungleichungen ' \leq ' und ' \geq ' lassen sich durch Multiplikation mit -1 ineinander überführen.

(b) Es gilt

$$\max_{x \in F} c^T x = - \min_{x \in F} (-c)^T x \quad \wedge \quad \operatorname{argmax}_{x \in F} c^T x = \operatorname{argmin}_{x \in F} (-c)^T x,$$

d.h. das Maximierungsproblem kann als Minimierungsproblem geschrieben werden und umgekehrt.

(c) Eine Gleichungsbedingung $a^T x = \beta$ kann durch die beiden Ungleichungsbedingungen

$$a^T x \leq \beta \quad \wedge \quad a^T x \geq \beta$$

ersetzt werden.

Man beachte, dass (b) einen etwas anderen Charakter hat als (a) und (c), da die Überführung einer Maximierungs- in eine Minimierungsaufgabe nicht nur eine Änderung des Inputs sondern auch eine Änderung des Outputs, nämlich des Vorzeichens des Zielfunktionswerts erfordert.

Zur Überführung von Ungleichungen in Gleichungen und von nicht vorzeichenbeschränkten Variablen in vorzeichenbeschränkte führt man üblicherweise neue Variablen ein. Die mathematische Grundlage für die erste Reduktion liefert das folgende Lemma.

1.4.3 Lemma. Seien $a \in \mathbb{R}^n \setminus \{0\}$ und $\beta \in \mathbb{R}$,

$$R := \left\{ \begin{pmatrix} x \\ \xi \end{pmatrix} \in \mathbb{R}^{n+1} : a^T x \leq \beta \wedge \xi = 0 \right\},$$

$$S := \left\{ \begin{pmatrix} x \\ \xi \end{pmatrix} \in \mathbb{R}^{n+1} : a^T x + \xi = \beta \wedge \xi \geq 0 \right\},$$

³⁶ Die Bezeichnungen 'Standardform' und 'kanonische Form' sind weder so kanonisch noch ein so fixer Standard, wie man denken würde. Manche Autoren verwenden sie in genau vertauschter Bedeutung.

und $\Psi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}$ sei die durch

$$\Psi \begin{pmatrix} x \\ \xi \end{pmatrix} := \begin{pmatrix} 0 \\ \beta \end{pmatrix} + \begin{pmatrix} x \\ \xi - a^T x \end{pmatrix}$$

definierte Abbildung. Dann ist Ψ bijektiv, und es gilt $\Psi(R) = S$. Eingeschränkt auf die Menge $H := \left\{ \begin{pmatrix} x \\ \xi \end{pmatrix} : a^T x + \xi = \beta \right\}$ ist Ψ^{-1} die Orthogonalprojektion auf $\mathbb{R}^n \times \{0\}$, d.h. auf den Raum der Vektoren der ersten n Koordinaten.

Beweis: Bezeichnet E_n (wie in Bezeichnung 1.3.7) die $(n \times n)$ Einheitsmatrix, so ist der lineare Anteil von Ψ durch die Matrix

$$\begin{pmatrix} E_n & 0 \\ -a^T & 1 \end{pmatrix}$$

gegeben. Sie hat Rang $n + 1$; Ψ ist somit bijektiv. Ferner gilt für $(x^T, \xi)^T \in R$ natürlich $a^T x \leq \beta$ und $\xi = 0$, d.h.

$$\Psi \begin{pmatrix} x \\ \xi \end{pmatrix} = \begin{pmatrix} x \\ \beta - a^T x \end{pmatrix} \in S.$$

Umgekehrt hat jeder Punkt aus S die Gestalt $(x^T, \beta - a^T x)^T$ mit $\beta - a^T x \geq 0$, ist also Bild von $(x^T, 0)^T$ unter Ψ . Hiermit ist auch der zweite Teil der Behauptung bewiesen. Offenbar ist

$$\Psi^{-1} \begin{pmatrix} x \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ -\beta \end{pmatrix} + \begin{pmatrix} E_n & 0 \\ a^T & 1 \end{pmatrix} \begin{pmatrix} x \\ \xi \end{pmatrix},$$

d.h. ein Punkt $(x^T, \xi)^T \in H$ wird abgebildet auf

$$\Psi^{-1} \begin{pmatrix} x \\ \xi \end{pmatrix} = \begin{pmatrix} x \\ -\beta + a^T x + \xi \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix},$$

und es folgt der letzte Teil der Behauptung. \square

Durch das Hinzufügen einer Variablen ξ zusammen mit der Bedingung $\xi = 0$ wird eine LP-Aufgabe lediglich in den durch $\xi = 0$ gegebenen Koordinatenunterraum des \mathbb{R}^{n+1} eingebettet, und das hat auf ihre Lösbarkeit und das Optimum einer zugehörigen LP-Aufgabe ebenso wenig Einfluss wie das anschließende 'Liften' durch Ψ .

1.4.4 Korollar. Seien $a \in \mathbb{R}^n$ und $\beta \in \mathbb{R}$. Die Ungleichung $a^T x \leq \beta$ ist äquivalent zu

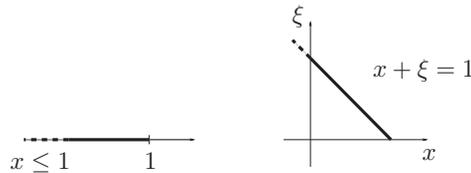
$$a^T x + \xi = \beta \wedge \xi \geq 0,$$

d.h. kann durch eine Gleichung und eine Nichtnegativitätsbedingung ersetzt werden.

Beweis: Die Aussage folgt unmittelbar aus Lemma 1.4.3. \square

1.4.5 Bezeichnung. Die in Korollar 1.4.4 hinzukommende Variable ξ heißt **Schlupfvariable** für die Ungleichung $a^T x \leq \beta$.

Die Überführung einer LP-Aufgabe in eine solche mit Nichtnegativitätsbedingungen für alle Variablen erfordert einen anderen Zugang.



1.10 Abbildung. Geometrische Interpretation der Einführung einer Schlupfvariable.

1.4.6 Bemerkung. Aus einer gegebenen LP-Aufgabe \mathcal{I} entstehe eine neue Aufgabe \mathcal{I}^* durch Ersetzen jeder Variablen ξ_j mit $j \in N_3$ durch die Differenz $\mu_j - \nu_j$ zweier neuer Variablen und Hinzufügen der Nichtnegativitätsbedingungen $\mu_j, \nu_j \geq 0$.

Die Setzung

$$\mu_j := \max\{0, \xi_j\} \quad \wedge \quad \nu_j := -\min\{\xi_j, 0\},$$

der neuen Variablen ordnet jedem zulässigen Punkt x der ursprünglichen Aufgabe \mathcal{I} einen zulässigen Punkt der modifizierten Aufgabe \mathcal{I}^* zu. Umgekehrt überführt die Setzung

$$\xi_j := \mu_j - \nu_j$$

eine zulässige Lösung für \mathcal{I}^* in eine von \mathcal{I} .

Wir werden auf die Geometrie, die Bemerkung 1.4.6 zugrunde liegt, noch einmal in Sektion 4.5 zurückkommen.

Die verwendete Reduktion ermöglichen es, Verfahren zur Lösung einer Form von LP-Problemen auch für andere Formen linearer Optimierungsprobleme einzusetzen. Der Körper der zugrunde liegenden Inputdaten wird dabei nicht verlassen. Liegen diese insbesondere in \mathbb{Q} , so bleibt das auch nach der Transformation der Fall.

Bei der Zurückführung verschiedener Formen von linearen Optimierungsproblemen aufeinander wird zum Teil die Dimension deutlich vergrößert, und die Zuordnung braucht auf den zulässigen Bereichen auch nicht bijektiv zu sein. Was bedeutet dann aber die ‘Äquivalenz’ der verschiedenen Formen von LP-Problemen wirklich, und welche Auswirkungen hat sie auf die Optimierungspraxis? Eine algorithmische Präzisierung dieses Konzepts wird in Satz 3.1.41 gegeben.

1.5 Übungsaufgaben

1.5.1 Übungsaufgabe. In der Verfahrenstechnik stehen zahlreiche chemische, biologische und physikalische Verfahren³⁷ zur Verfügung, um aus einem Rohstoff verschiedene Produkte zu gewinnen. Die entsprechende Ausbeute hängt vom eingesetzten Verfahren (‘Prozess’) ab. Wir nehmen an, dass in einem chemischen Filtrations- und Aufbereitungsverfahren aus einem Rohstoff zunächst Fluide verschiedener Zusammensetzungen, Reinheiten und Konzentrationen ‘hoch’ (H), ‘mittel’ (M) und ‘niedrig’ (N) hergestellt werden, die danach in unterschiedlichen Folgeprozessen weiter verarbeitet werden können.

In der Produktion stehen zwei Verfahren zur Verfügung, die bezogen auf 100 Mengeneinheiten des Rohstoffs die folgenden Ergebnisse liefern und Kosten (K) verursachen:

Prozess A: $H = 15$, $M = N = 25$, $K = 3$; Prozess B: $H = 45$, $M = 20$, $N = 10$, $K = 5$.

Ein aktueller Auftrag erfordert $(H, M, N) = (40, 50, 30)$; die Mengen sollen so kostengünstig wie möglich produziert werden.

(a) Man modelliere die Aufgabe als lineares Programm.

³⁷ Hierzu gehören das Zerkleinern, die Filtration und Destillation, die Oxidation und Polymerisation, die Elektrolyse, die Gärung etc.

(b) Man löse das lineare Programm graphisch.

1.5.2 Übungsaufgabe. Eine für die Herstellung von zwei Produkten P_1 und P_2 benötigte Maschine steht einem Unternehmen pro Tag 16 Stunden zur Verfügung. Um eine Einheit von P_1 zu fertigen, ist eine Maschinenlaufzeit von 6 Minuten erforderlich; für eine Einheit von P_2 werden hingegen 18 Minuten benötigt. Rüst- und Umrüstzeiten fallen nicht an. Die variablen Stückkosten liegen bei 4,- Euro für P_1 bzw. bei 44,- Euro für P_2 . Auf dem Absatzmarkt lassen sich hierfür Stückpreise von 11,- Euro bzw. 49,- Euro erzielen. Allerdings können von P_1 maximal 100 und von P_2 maximal 40 Stück pro Tag verkauft werden.

(a) Man stelle lineare Programme für die beiden Unternehmensziele der Maximierung des Gewinns bzw. der Maximierung des Umsatzes auf.

(b) Man bestimme die jeweiligen Optima.

(c) Da hier Umsatz und Gewinn nicht gleichzeitig maximiert werden können, wird entschieden, eine Produktion zu realisieren, die beide Zielwerte zu einem maximalen gleichen Prozentsatz erreicht. Man formuliere eine entsprechende lineare Optimierungsaufgabe.

1.5.3 Übungsaufgabe. Bei den Produktionsaufgaben gemäß Bezeichnung 1.2.1 geht man davon aus, dass der Erlös linear von der hergestellten Quantität der Produkte abhängt. Oftmals sind die Produzenten aber gezwungen, ihren Kunden Mengenrabatte einzuräumen, so dass die Grenzerträge abnehmen. So ergebe sich etwa aus den getroffenen Liefervereinbarungen, dass bis zur Produktion von 1000 Stück des Produkts P der Gewinn pro Stück 100 ist, für die Stückzahlen 1001 – 10000 sinkt er auf 80, und für jedes weitere Stück bis zur Produktionsgrenze von 100000 wird nur noch ein Gewinn von 50 erzielt. Man modelliere diese Variante als ganzzahlige Optimierungsaufgabe. Kann man das auf eine Weise tun, dass die Variablen dabei nur linear eingehen?

1.5.4 Übungsaufgabe. Im Folgenden sind verschiedene lineare Optimierungsaufgaben

$$\max\{c_i^T x : A_j x \leq b_k \wedge x \geq 0\}$$

gegeben. Dabei seien

$$A_1 := \begin{pmatrix} -1 & 2 \\ 3 & 1 \end{pmatrix} \wedge A_2 := \begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix} \wedge b_1 := \begin{pmatrix} 4 \\ 9 \end{pmatrix} \wedge b_2 := \begin{pmatrix} 4 \\ -5 \end{pmatrix} \\ b_3 := \begin{pmatrix} 4 \\ -2 \end{pmatrix} \wedge c_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix} \wedge c_2 := \begin{pmatrix} 3 \\ 1 \end{pmatrix} \wedge c_3 := \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Man skizziere die zulässigen Bereiche für die folgenden Tripel (A_j, b_k, c_i) und bestimme graphisch den Optimalwert und alle Optimallösungen:

(a) (A_1, b_1, c_1) ; (b) (A_2, b_2, c_1) ; (c) (A_2, b_3, c_1) ; (d) (A_1, b_1, c_2) ; (e) (A_1, b_1, c_3) .

1.5.5 Übungsaufgabe. Seien

$$A := \begin{pmatrix} -1 & 2 \\ 3 & 1 \end{pmatrix} \wedge b := \begin{pmatrix} 4 \\ 9 \end{pmatrix}$$

sowie

$$(a) v_1 := \begin{pmatrix} 2 \\ 3 \end{pmatrix}; (b) v_2 := \begin{pmatrix} 0 \\ 2 \end{pmatrix}; (c) v_3 := \begin{pmatrix} 1 \\ 5/2 \end{pmatrix}; (d) v_4 := \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Für jedes $i \in [4]$ bestimme man die Menge der möglichen Zielfunktionsvektoren c , so dass der Punkt v_i Optimallösung der Aufgabe $\max\{c^T x : Ax \leq b \wedge x \geq 0\}$ ist. Man beweise jeweils die Optimalität durch Herleitung einer geeigneten oberen Schranke für den optimalen Zielfunktionswert.

Hinweis: Schranken können durch positive Skalierung und Addition von Ungleichungen gewonnen werden.

1.5.6 Übungsaufgabe. Seien $P := \{x \in \mathbb{R}^3 : Ax \leq b\}$ mit

$$A^T := \begin{pmatrix} -1 & 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 & 2 \end{pmatrix} \wedge b^T := (0, 0, 0, 6, 4) \wedge c^T := (1, -1, 2).$$

(a) Man skizziere das Polyeder P .

(b) Man bestimme $\max_{x \in P} c^T x$ und $\min_{x \in P} c^T x$.

1.5.7 Übungsaufgabe. In einer Schule haben sich zwei Lehrer um eine Beförderung beworben, mit der die Aufgabe der Koordination der Stundenpläne verbunden ist. Um sich von den Fähigkeiten der Bewerber zu überzeugen, hat ihnen die Direktorin den folgenden kleinen Testdatensatz für zwei Klassen geschickt.

Lehrkraft	Fach	Klasse A	Klasse B
Herr Meyer	Deutsch	3	3
Frau König	Spanisch	1	1
Herr Wenner	Erdkunde	2	2
Herr Bauer	Mathematik	4	0
Frau Weinmann	Mathematik	0	4
Frau Dr. Matthes	Physik	2	2
Herr Krämser	Englisch	2	2
Herr Gahl	Wirtschaft/Recht	0	1
Frau Volkmar	Wirtschaft/Recht	1	0

Für die Stunden stehen die Zeiträume

Montag – Freitag: 8:00 – 8:45 Uhr, 8:50 – 9:35 Uhr, 9:45 – 10:30 Uhr, 10:35 – 11:20 Uhr

zur Verfügung. Die Tabelle zeigt die Lehrkraft und die Anzahl der Wochenstunden, die jeweils von ihr in den beiden Klassen zu unterrichten sind. Dabei soll Spanisch immer in der letzten Stunde eines Tages stattfinden, die erste Stunde am Montag soll frei bleiben. Herr Bauer kann aufgrund anderer Verpflichtungen Montags nicht unterrichten; Frau König steht Mittwochs nicht zur Verfügung. Jedes Fach soll höchstens zwei Stunden pro Tag und Klasse unterrichtet werden. Wird ein Fach zwei Stunden an demselben Tag unterrichtet, so müssen diese hintereinander stattfinden (Doppelstunde). Ziel ist es, im Studienplan möglichst wenig Freistunden zu haben, d.h. Stunden ohne Unterricht, obwohl vorher und nachher an demselben Tag noch Unterricht in der Klasse stattfindet.

Man modelliere die Studienplanung als ganzzahlige lineare Optimierungsaufgabe.

1.5.8 Übungsaufgabe. Aus Eisenstäben der Länge $\beta \in \mathbb{N}$ sollen für $j \in [n]$ jeweils $\tau_j \in \mathbb{N}$ Stäbe der Länge α_j zurecht geschnitten werden. Die Anzahl der angeschnittenen Eisenstäbe sei dabei minimal. Man formuliere diesen Auftrag als ganzzahlige lineare Optimierungsaufgabe.

1.5.9 Übungsaufgabe. Im Sortiment eines Metallhändlers seien m verschiedene Typen von Eisenstäben, die sich jedoch nur in ihren Längen unterscheiden. Für $i \in [m]$ seien $\beta_i \in \mathbb{N}$ die Länge und $\kappa_i \in \mathbb{N}$ die vorhandene Anzahl von Eisenstäben des Typs i .

Ein Kunde möchte für $j \in [n]$ jeweils $\tau_j \in \mathbb{N}$ Stäbe der Länge $\alpha_j \in \mathbb{N}$ zurecht geschnitten bekommen. Der Händler kann nun verschiedene Optimierungsziele verfolgen. Hierzu gehören die Minimierung

- der Summe der Verschnittreste aller angeschnittenen Eisenstäbe;
- der Anzahl der angeschnittenen Stäbe;
- der Anzahl der verschiedenen Typen angeschnittener Stäbe;
- der Summe der Verschnittreste aller Stäbe, deren Restlänge kleiner ist als die Länge der kürzesten Stange im Sortiment;
- der Summe der Verschnittreste aller Stäbe, deren Restlänge nicht in $\{\beta_1, \dots, \beta_m\}$ liegt.

Man formuliere entsprechende Optimierungsaufgaben in ganzzahligen Variablen.

1.5.10 Übungsaufgabe. Seien $m, n \in \mathbb{N}$, $A := (a_1, \dots, a_m)^T \in \mathbb{R}^{m \times n}$ und $b := (\beta_1, \dots, \beta_m)^T \in \mathbb{R}^m$. Im Allgemeinen wird das Gleichungssystem $Ax = b$ nicht lösbar sein. Gesucht ist daher ein Vektor $x \in \mathbb{R}^n$, für den der ‘Worst-case-Fehler’, d.h. das Maximum der Einzelfehler in den m Gleichungen minimal ist. Es liegt also die Aufgabe

$$\min_{x \in \mathbb{R}^n} \max_{i \in [m]} |a_i^T x - \beta_i|$$

vor. Man formuliere diese Aufgabe als lineares Programm.

1.5.11 Übungsaufgabe. Eine Fahrradfabrik stellt zwei verschiedene Spezialräder her, Modell A und Modell B. Pro Stück ist eine manuelle Bearbeitung von 20 Stunden Arbeitszeit für Modell A und 10 Stunden für Modell B erforderlich. Insgesamt stehen (in dem betrachteten Produktionszeitraum) 16.000 Arbeitsstunden zur Verfügung. Für die maschinelle Bearbeitung werden für beide Modelle jeweils 4 Stunden Maschinenzeit pro Stück benötigt. Insgesamt sind 4.000 Maschinenstunden verfügbar.

Der Bedarf an Spezialschrauben beträgt pro Rad 6 Stück für Typ A bzw. 15 Stück für Typ B. Insgesamt sind 9000 dieser Schrauben vorrätig; eine Nachlieferung kann innerhalb des betrachteten Produktionszeitraums nicht erfolgen. Alle anderen erforderlichen Teile und Instrumente stehen in ausreichender Zahl zur Verfügung. Beim Verkauf ergibt sich (nach Abzug aller Kosten) ein Reingewinn von 160 Euro pro Fahrrad des Modells A und 320 Euro pro Rad des Modells B. Der Gesamtgewinn soll maximiert werden.

- Man modelliere die beschriebene Produktion als ganzzahlige lineare Optimierungsaufgabe.
- Man stelle den zulässigen Bereich graphisch dar und ermittle alle Optimallösungen.

1.5.12 Übungsaufgabe. Die Rekonstruktion kristalliner Strukturen gemäß Sektion 1.2 soll auf den Polyatombfall erweitert werden. Hierbei treten mehrere verschiedene Atomsorten auf, für die jeweils getrennte Messdaten in den Richtungen u_1, u_2, u_3 vorliegen. Wie zuvor wissen wir daher, wieviele Atome des jeweiligen Typs auf den einzelnen Gittergeraden liegen. Natürlich kann sich auf jedem Gitterpunkt höchstens ein Atom befinden. Man modelliere die Rekonstruktion als ganzzahlige Zulässigkeitsaufgabe.

1.5.13 Übungsaufgabe. Durch Hinzunahme von Ganzzahligkeitsbedingungen für alle Variablen entsteht aus einem linearen Programm LP eine zugehörige ganzzahlige Optimierungsaufgabe $G(LP)$. Man konstruiere eine Folge $(LP_n)_{n \in \mathbb{N}}$ linearer Programme LP_n mit den folgenden Eigenschaften:

- LP_n hat n Variable.
- Alle bei der Beschreibung von LP_n auftretenden Daten sind aus $\{0,1\}$.
- Der zulässige Bereich von LP_n enthält die Punkte $0, u_1, \dots, u_n$.
- Für jedes $k \in \mathbb{N}$ gibt es ein $n(k) \in \mathbb{N}$, so dass für alle $n \geq n(k)$ die Differenz der Optimalwerte von LP_n und $G(LP_n)$ mindestens k beträgt.

1.5.14 Übungsaufgabe. Der Betreiber eines Senders möchte die Sendeleistung σ optimieren. Sie soll einerseits möglichst klein sein, andererseits aber einen störungsfreien Empfang in seinem Sendegebiet gewährleisten. Es liegen die folgenden Fakten zugrunde. Das mit Funkempfang zu versorgende Gebiet ist durch die Bedingungen

$$\xi_1 - \xi_2 \geq -1 \quad \wedge \quad \xi_1 + \xi_2 \leq 2 \quad \wedge \quad 3\xi_1 - \xi_2 \leq 3 \quad \wedge \quad 2\xi_1 + \xi_2 \geq -2$$

gegeben; der Sender steht im Ursprung und strahlt in alle Richtungen gleichmäßig ab. Die in einem Punkt $x \in \mathbb{R}^2$ ankommende Signalstärke beträgt $\sigma / (1 + \|x\|_{(2)}^2)$. Für einen klaren Empfang ist eine Signalstärke von mindestens 1 erforderlich.

- Man modelliere die Angelegenheit als Optimierungsaufgabe.
- Man löse die Aufgabe graphisch.
- Man konstruiere ein lineares Programm, mit dessen Hilfe man beweisen kann, dass die in (b) gefundene Lösung tatsächlich optimal ist.

1.5.15 Übungsaufgabe. In einem Bieterverfahren, an dem sich Fluggesellschaften beteiligen können, sollen die Rechte vergeben werden, zwischen 20 verschiedenen Städten direkte Flugverbindungen zu betreiben. Diese sind jeweils in beide Richtungen benutzbar; keine zwei von ihnen sollen allerdings dieselben beiden Städte verbinden.

- Angenommen, die Fluggesellschaft wäre völlig frei in der Wahl der Verbindungen, was wäre die Mindestzahl von Flugverbindungen, mit denen man garantieren kann, dass man von jeder Stadt in jede andere fliegen kann, ohne dabei mehr als einmal umsteigen zu müssen. Wie verändert sich diese Zahl, wenn man zweimaliges oder dreimaliges Umsteigen akzeptiert?
- Angenommen es bestünden Verpflichtungen, bestimmte vorgegebene Direktverbindungen zu bedienen. Man bestimme die Höchstzahl solcher vorgeschriebener Flugverbindungen, mit denen noch nicht garantiert ist, dass man von jeder Stadt in jede andere fliegen kann, ohne dabei mehr als einmal umsteigen zu müssen.

1.5.16 Übungsaufgabe. Im euklidischen Traveling Salesman Problem (E-TSP) ist die Distanz von je zwei Punkten durch ihren euklidische Abstand gegeben. Man beweise die folgenden Aussagen:

- Es existiert eine Konstante α , so dass für jedes $n \in \mathbb{N}$ jede Aufgabe von E-TSP mit n Knoten in $[0,1]^2$ eine Tour T besitzt, dessen Länge $\zeta(T)$ die Ungleichung $\zeta(T) \leq \alpha\sqrt{n}$ erfüllt.
- Es existiert eine positive Konstante β , so dass für jedes $n \in \mathbb{N}$ eine Aufgabe von E-TSP mit n Knoten in $[0,1]^2$ existiert, so dass für jede Tour T die Ungleichung $\zeta(T) \geq \beta\sqrt{n}$ gilt.

Hinweis: Man verwende geeignete Zerlegungen von $[0,1]^2$ in Streifen bzw. Quadrate gleicher Größe.

1.5.17 Übungsaufgabe. Gegeben sei eine endliche Punktmenge $F \subset \mathbb{R}^3$. Die Menge F ist nicht direkt zugänglich; es ist aber für jede Gerade parallel zur ξ_1 - oder zur ξ_2 -Achse bekannt, wieviele Punkte aus F sie trifft. (Natürlich sind alle bis auf endlich viele dieser Geraden zu F disjunkt.) Ziel ist es, F zu rekonstruieren. Man zeige zunächst, dass F im Allgemeinen durch diese Daten nicht eindeutig bestimmt ist. Um eine möglichst gute Lösung zu erhalten, soll die Rekonstruktion von F sukzessive in Schichten parallel zur (ξ_1, ξ_2) -Ebene in der Reihenfolge zunehmender ξ_3 -Koordinaten erfolgen, um bei der Rekonstruktion einer Schicht Informationen aus der Vorgängerschicht zu verwenden. Die Experten kennen den Schnitt von F mit der ersten Schicht exakt und gehen davon aus, dass unter allen möglichen Lösungen in der $(k+1)$ -ten Schicht, die mit den Messdaten übereinstimmen, solche am wahrscheinlichsten sind, die an den wenigsten Stellen von der gefundenen Lösung in der k -ten Schicht abweichen. Man gebe einen Algorithmus an, der ein solches Verfahren zur Bestimmung einer 'möglichst plausiblen' Lösung umsetzt. Als Subroutine kann dabei eine Methode zur Lösung ganzzahliger Optimierungsaufgaben verwendet werden.

1.5.18 Übungsaufgabe. Ein Lastwagen soll so beladen werden, dass der Gesamtwert der Ladung maximiert wird, ohne jedoch seine Ladekapazität zu überschreiten. Zur Beladung stehen n Güter G_1, \dots, G_n bereit; für $i \in [n]$ habe G_i das Gewicht $\gamma_i \in]0, \infty[$ und den Wert $\omega_i \in]0, \infty[$. Das zulässige Ladegewicht des Wagens betrage $\beta \in]0, \infty[$. Die Auswahl der zu verladenden Güter erfolgt durch den Betreiber nach der folgenden Systematik \mathcal{A}_1 . Die Güter werden zunächst dem Wert nach sortiert, d.h. so umnummeriert, dass $\omega_1 \geq \dots \geq \omega_n$ gilt. Dann wird der Reihe nach jedes Gut geladen, solange hierdurch die Ladekapazität nicht überschritten wird.

(a) Man zeige, dass für jedes $\epsilon \in]0, \infty[$ eine Instanz existiert, so dass für das Optimum ζ^* und den Zielfunktionswert $\zeta(\mathcal{A}_1)$ der durch \mathcal{A}_1 gefundenen Lösung

$$\zeta(\mathcal{A}_1) \leq \epsilon \cdot \zeta^*$$

gilt, das Verfahren also beliebig schlecht sein kann.

Ein Problem von \mathcal{A}_1 liegt offenbar darin, dass die Reihenfolge der Verladung ausschließlich vom Wert, nicht aber vom Gewicht der Güter bestimmt wird. Man kann daher hoffen, dass der analoge Algorithmus \mathcal{A}_2 , der die Güter in der Reihenfolge fallender Quotienten ω_i/γ_i sortiert, bessere Ergebnisse liefert.

(b) Stimmt das? (Beweis oder Gegenbeispiel)

1.5.19 Übungsaufgabe. Die Siegerin einer Quizshow darf sich Preise aus n verschiedenen Kategorien K_1, \dots, K_n auswählen. Aus jeder Kategorie darf sie dabei so viele Preise nehmen, wie sie möchte; allerdings darf das Gesamtgewicht aller gewählten Gegenstände eine (durch den Punktestand des vorherigen Spielverlaufs festgelegte) Schranke $\beta \in \mathbb{N}$ nicht überschreiten. Alle Gewinne einer Kategorie haben jeweils das gleiche Gewicht und den gleichen Wert. Für $i \in [n]$ sei $\gamma_i \in \mathbb{N}$ das Gewicht und $\omega_i \in \mathbb{N}$ der Wert jedes Preises aus K_i . Diese Größen sind der Spielerin bekannt. Da sie nicht viel Zeit hat zu überlegen, wie sie eine optimale Auswahl bestimmen kann, verfährt die Spielerin nach dem folgenden Verfahren \mathcal{A} :

```

BEGIN   I ← [n]; ρ ← β
        WHILE I ≠ ∅ DO
          BEGIN
            Wähle  $i^* \in \operatorname{argmax} \left\{ \frac{\omega_i}{\gamma_i} : i \in I \right\}$ ;  $I \leftarrow I \setminus \{i^*\}$ 
            IF  $\gamma_{i^*} \leq \rho$  THEN wähle  $\lfloor \frac{\rho}{\gamma_{i^*}} \rfloor$  Preise aus  $K_{i^*}$ ;  $\rho \leftarrow \rho - \lfloor \frac{\rho}{\gamma_{i^*}} \rfloor \gamma_{i^*}$ 
          END
        END

```

Seien ζ^* der maximal mögliche Wert des Gewinns und $\zeta(\mathcal{A})$ der Wert der durch \mathcal{A} gefundenen Lösung. Gibt es eine Konstante $\kappa \in]0, \infty[$ mit

$$\zeta(\mathcal{A}) \geq \kappa \cdot \zeta^*?$$

Falls ja, gebe man ein möglichst großes solches κ an. Falls nein, konstruiere man für jedes positive κ eine Aufgabe, für die $\zeta(\mathcal{A}) \leq \kappa \cdot \zeta^*$ gilt. Man vergleiche die Ergebnisse mit denen von Übungsaufgabe 1.5.19.

1.5.20 Übungsaufgabe. Man konstruieren einen Algorithmus, der für $n \in \mathbb{N}$, paarweise orthogonale Vektoren $a_1, \dots, a_n \in \mathbb{R}^n$ und $t \in \mathbb{R}^n$

$$\max \left\{ \left\| t + \sum_{i=1}^n \lambda_i a_i \right\|_{(2)}^2 : \lambda_1, \dots, \lambda_n \in [-1, 1] \right\}$$

mit maximal $n + 1$ Skalarproduktberechnungen und n Ergebnisvergleichen berechnet und beweise seine Korrektheit.

1.5.21 Übungsaufgabe. Gegeben seien die linearen Optimierungsaufgaben

$$\begin{array}{ll} \max (-2\xi_1 + \xi_2) & \max (-2\xi_1 + \xi_2^+ - \xi_2^-) \\ \text{(I) } \begin{array}{l} \xi_1 - \xi_2 \geq -1 \\ \xi_1 \geq 0 \end{array} & \text{(II) } \begin{array}{l} \xi_1 - \xi_2^+ + \xi_2^- \geq -1 \\ \xi_1, \xi_2^+, \xi_2^- \geq 0. \end{array} \end{array}$$

Der Zusammenhang zwischen (I) und (II) ist vermöge $(\xi_2 = \xi_2^+ - \xi_2^- \wedge \xi_2^+, \xi_2^- \geq 0)$ gegeben. Durch diese Einbettung in den \mathbb{R}^3 unterliegen dann alle Variablen Nichtnegativitätsbedingungen.

- Man bestimme die Optimallösungen von (I) und (II).
- Der Zielfunktionsvektor $c := (-2, 1)^T$ von (I) sei nicht exakt, sondern fehlerbehaftet in der Form $(-2, 1 + \epsilon)^T$ mit $\epsilon \in [0, \infty[$ gegeben. Wie groß darf ϵ maximal werden, wenn die ursprüngliche Optimallösung erhalten bleiben soll?
- Aufgrund fehlerhafter Kodierung wird der Zielfunktionsvektor von (II) als $(-2, 1 + 10^{-17}, -1)^T$ gespeichert. Wie lautet dann die Lösung von (II)? (Beweis)