

KONVEXITÄT

Skript zur Vorlesung an der Universität Konstanz
im Wintersemester 2011/2012

Daniel Plaumann

VORWORT

Die Vorlesung, der dieses Skript entstammt, wurde im Wintersemester 2011/2012 an der Universität Konstanz gehalten. Ziel war eine Einführung in die Theorie konvexer Mengen, basierend auf dem Buch von Barvinok [1], gefolgt von einem vertiefenden Teil, in dem die Bezüge zur reellen algebraischen Geometrie und zur polynomialen Optimierung betont wurden. Die Vorlesung war zweistündig, begleitet von einer zweistündigen Übung. Die Teilnehmer waren etwa zwölf Studentinnen und Studenten in unterschiedlichen Phasen (Bachelor bis Promotion). Das Skript zur Vorlesung wurde in jeder Sitzung zu Beginn zum Mitlesen zur Verfügung gestellt. Im wesentlichen entspricht jeder Abschnitt genau einer Doppelstunde (90 Minuten), wobei manchmal ein kleinerer Teil in die Übung verlagert wurde. In der Übung wurden jede Woche nur zwei vorbereitete Aufgaben besprochen, die übrige Zeit wurde mit dem gemeinsamen Lösen von Präsenzübungsaufgaben in der Gruppe und manchmal mit Ergänzungen zur Vorlesung verbracht. Die Übungsaufgaben sind in den Text eingebunden.

Dies ist ein Vorlesungsskript, kein Lehrbuch. Das bedeutet insbesondere:

- Ergänzende Erläuterungen und veranschaulichende Zeichnungen fehlen manchmal, da sie zum Teil nur während der Vorlesung an der Tafel gegeben wurden.
- Es wird kein Anspruch auf Originalität erhoben. Die ersten zwei Drittel der Vorlesung orientieren sich stark am Buch von Barvinok. Vor allem die Beweise und ein großer Teil der Übungsaufgaben sind oft sinngemäß übernommen, ohne dass dies im Einzelnen kenntlich gemacht ist. Abschnitt 12 ist stark an die Darstellung im Buch von Forst und Hoffmann [3] angelehnt. Die späteren Abschnitte basieren zum Teil auf wissenschaftlichen Arbeiten, die dann auch im Anhang zitiert sind.
- Die Liste der zitierten Literatur stellt keine umfassende Übersicht dar, sondern enthält nur die Referenzen, auf die im Text direkt Bezug genommen wird.

Sollte ich die Vorlesung noch einmal halten, würde ich natürlich das eine oder andere ändern: Die Sätze von Radon und Helly stehen etwas isoliert und kommen später nicht wieder vor. Der Isolationssatz für algebraisch-offene Mengen in Abschnitt 2 wird in dieser Form nicht weiter verwendet, weil die unendlich-dimensionale Theorie schließlich ganz außen vor bleibt. Die erste Definition von Spektraeder in Abschnitt 7 ist etwas ungenügend und wird später in Abschnitt 13 leicht verallgemeinert. Die Liste ließe sich fortsetzen, und auch an der Notation gäbe es einiges zu vereinheitlichen. Da ich dieses Skript weiter pflegen möchte, bin ich für Anregungen und Korrekturhinweise dankbar.

Konstanz, 8. Februar 2012
Daniel Plaumann

Inhalt

Vorwort	3
1. Einführung	7
1.1. Grundlegende Definitionen	7
1.2. Der Satz von Carathéodory	9
1.3. Die Sätze von Radon und Helly	10
2. Trennende Hyperebenen	11
2.1. Hyperebenen	11
2.2. Dimension und Topologie im euklidischen Raum	12
2.3. Der Isolationssatz	13
2.4. Exkurs: Der Satz von Kirchberger	15
3. Seiten und Extrempunkte	17
4. Polyeder	21
4.1. Die Seitenstruktur von Polyedern	21
4.2. Das Birkhoff-Polytop	23
5. Trennungssätze und Kegel	25
5.1. Weitere Trennungssätze	25
5.2. Kegel	26
6. Positiv semidefinite Matrizen	29
6.1. Grundlagen	29
6.2. Der psd-Kegel	31
7. Positiv semidefinite Matrizen II	33
7.1. Spektraeder und lineare Gleichungen in Matrizen	33
7.2. Anwendung: Quadratische Konvexität	35
8. Dualität	37
8.1. Der Bidualsatz	37
8.2. Dualität von Polytopen	39
9. Kegelprogrammierung	41
10. Lineare Programmierung	45
Anwendung: Das Transport-Problem	47
11. Semidefinite Programmierung	49
11.1. Dualität für semidefinite Programme	49
11.2. Anwendung: Cliques- und Färbungszahl eines Graphen	51
12. Innere-Punkte-Verfahren	53
12.1. Das Primär-Duale System	53
12.2. Barrierefunktionen und der zentrale Pfad	54
12.3. Ein Beispiel mit Matlab/SeDuMi/Yalmip	57

13. Spektraeder und starre Konvexität	59
13.1. Spektraeder als Punktmenen	59
13.2. Starre Konvexität	61
14. Spektraeder II	63
14.1. Starre Konvexität (Fortsetzung)	63
14.2. Spektrale Schatten	65
15. Die Lasserre-Relaxierung	67
Zitierte Literatur	71

1. EINFÜHRUNG

1.1. GRUNDLEGENDE DEFINITIONEN

Es sei V ein \mathbb{R} -Vektorraum. Häufig nennen wir die Elemente eines Vektorraums *Punkte*. Seien $v, w \in V$ zwei Punkte, dann ist

$$[v, w] = \{ \alpha v + (1 - \alpha)w \mid \alpha \in \mathbb{R} \text{ mit } 0 \leq \alpha \leq 1 \}$$

das *Geradensegment zwischen v und w* . Eine Teilmenge C von V heißt *konvex*, wenn $[v, w] \subset C$ für alle $v, w \in C$ gilt. Für jede Teilmenge S von V heißt

$$\text{conv}(S) = \left\{ \sum_{i=1}^k \alpha_i v_i \mid v_1, \dots, v_k \in S, \alpha_1, \dots, \alpha_k \in \mathbb{R}_{\geq 0} \text{ mit } \sum_{i=1}^k \alpha_i = 1, k \in \mathbb{N} \right\}$$

die *konvexe Hülle* von S . Die Elemente von $\text{conv}(S)$ heißen *Konvexkombinationen* der Elemente von S . Offenbar gilt $[v, w] = \text{conv}(\{v, w\})$.

Übung 1.1. Zeige, dass $\text{conv}(S)$ konvex ist.

Übung 1.2. Es sei $S \subset \mathbb{R}^d$ mit $\text{conv}(S) \neq \mathbb{R}^d$ und seien $u, v \in \mathbb{R}^d \setminus \text{conv}(S)$. Zeige: Falls $u \in \text{conv}(S \cup \{v\})$ und $v \in \text{conv}(S \cup \{u\})$, dann ist $u = v$.

Beispiele 1.1.

- (1) Ein paar kleine Bilder an der Tafel.
- (2) Ein *Polytop* ist die konvexe Hülle von endlich vielen Punkten. Ein wichtiges Beispiel für ein Polytop ist der *d -dimensionale Standard-Simplex* in \mathbb{R}^{d+1} :

$$\Delta_d = \left\{ (\alpha_1, \dots, \alpha_{d+1})^T \in \mathbb{R}^{d+1} \mid \sum_{i=1}^{d+1} \alpha_i = 1 \text{ und } \alpha_i \geq 0 \text{ für } i = 1, \dots, d+1 \right\}$$

Das ist also gerade die konvexe Hülle der $d+1$ Standardbasisvektoren in \mathbb{R}^{d+1} .

- (3) Ein (*konvexes*) *Polyeder* ist der Durchschnitt von endlich vielen Halbräumen, also eine Menge, die durch endlich viele lineare Ungleichungen beschrieben ist. In Formeln ausgedrückt ist ein Polyeder P also gegeben durch Linearformen $\ell_1, \dots, \ell_k : \mathbb{R}^d \rightarrow \mathbb{R}$ und reelle Zahlen $\beta_1, \dots, \beta_k \in \mathbb{R}$ als die Menge

$$P = \left\{ x \in \mathbb{R}^d \mid \ell_1(x) \leq \beta_1, \dots, \ell_k(x) \leq \beta_k \right\}.$$

Später beweisen wir den *Satz von Weyl-Minkowski*: Jedes Polytop ist ein kompaktes Polyeder und umgekehrt. Man mache sich diese Aussage an ein paar Beispielen in der Ebene klar. Wie könnte ein Beweis aussehen?

- (4) Eine konvexe Menge $C \subset \mathbb{R}^d$ mit der Eigenschaft $\lambda x \in C$ für alle $x \in C$ und $\lambda \in \mathbb{R}_{\geq 0}$ heißt ein *konvexer Kegel*. Der *positive Orthant*

$$\mathbb{R}_{\geq 0}^d = \{(x_1, \dots, x_d)^T \in \mathbb{R}^d \mid x_i \geq 0 \text{ für } i = 1, \dots, d\}$$

ist ein Beispiel für einen konvexen Kegel, der auch ein Polyeder ist.

- (5) Ein für uns sehr wichtiges Beispiel ist der *psd-Kegel*. Zur Erinnerung: Eine symmetrische $d \times d$ -Matrix A mit Einträgen in \mathbb{R} wird *positiv semidefinit* (psd) genannt, falls $v^T A v \geq 0$ für alle (Spaltenvektoren) $v \in \mathbb{R}^d$ gilt. Sind A, B psd, so ist offensichtlich auch $\alpha A + (1 - \alpha)B$ psd, für alle $0 \leq \alpha \leq 1$. Die Menge der psd 2×2 -Matrizen kann man als Teilmenge von \mathbb{R}^3 zeichnen und sieht einfach eine Eistüte (warum?). Aber für $d \geq 3$ ist es mit dieser Einfachheit schnell vorbei.

Proposition 1.2. Sei V ein \mathbb{R} -Vektorraum und sei $S \subset V$. Dann ist $\text{conv}(S)$ eine konvexe Teilmenge von V . Außerdem ist $\text{conv}(S)$ die kleinste konvexe Teilmenge von V , die S enthält, d.h. ist $C \subset V$ konvex und gilt $S \subset C$, so folgt $\text{conv}(S) \subset C$.

Beweis. Der erste Teil war schon als Übung dran. Für den zweiten Teil, sei C eine konvexe Teilmenge von V mit $S \subset C$. Sei $v \in \text{conv}(S)$, d.h. v ist eine Konvexkombination

$$v = \alpha_1 v_1 + \dots + \alpha_k v_k$$

von Punkten v_1, \dots, v_k in S . Wir müssen $v \in C$ zeigen. Wenn eins der $\alpha_i = 0$ ist, dann lassen wir das entsprechende v_i einfach weg. Wir dürfen also ohne Einschränkung annehmen, dass $\alpha_i > 0$ für alle $i = 1, \dots, k$ gilt. Wir beweisen die Behauptung durch Induktion nach k : Der Fall $k = 1$ ist klar, wegen $S \subset C$. Sei $k > 1$. Dann ist also $\alpha_k < 1$, und es gilt

$$v = (1 - \alpha_k)w + \alpha_k v_k \quad \text{wobei} \quad w = \frac{\alpha_1}{1 - \alpha_k} v_1 + \dots + \frac{\alpha_{k-1}}{1 - \alpha_k} v_{k-1}.$$

Nun ist w eine Konvexkombination von v_1, \dots, v_{k-1} , denn es gilt

$$\sum_{i=1}^{k-1} \frac{\alpha_i}{1 - \alpha_k} = \frac{1}{1 - \alpha_k} \sum_{i=1}^{k-1} \alpha_i = \frac{1 - \alpha_k}{1 - \alpha_k} = 1.$$

Nach Induktionsannahme ist $w \in C$, und da C konvex ist, folgt auch $v \in [w, v_k] \subset C$. \square

Übung 1.3. Es sei V ein reeller Vektorraum und seien A, B zwei konvexe Teilmengen von V .

- (1) Welche der Mengen $A \cup B, A \cap B, A \setminus B, A \times B$ (in $V \times V$) sind notwendig konvex?
- (2) Sei W ein weiterer reeller Vektorraum und $T: V \rightarrow W$ linear. Zeige, dass $T(A)$ konvex ist.

Übung 1.4. Die punktweise Summe

$$A + B = \{x + y \mid x \in A, y \in B\}$$

heißt die *Minkowski-Summe* von A und B .

- (1) Zeige, dass $A + B$ konvex ist.
- (2) Zeichne $A + B$ exemplarisch für einige einfache Figuren in der Ebene (Punkte, Geraden-segmente, Dreiecke, Kreise...).
- (3) Zeige, dass $(\alpha + \beta)A = \alpha A + \beta A$ für alle $\alpha, \beta \geq 0$ gilt. Gib ein Beispiel dafür, dass dies nicht der Fall sein muss, wenn α oder β negativ sind.

Übung 1.5. Beweise den Satz von Gauß-Lucas: Es sei $f \in \mathbb{C}[z]$ ein komplexes Polynom in einer Veränderlichen. Die konvexe Hülle der Nullstellen von f in der komplexen Ebene (aufgefasst als \mathbb{R}^2) enthält alle Nullstellen der Ableitung von f nach z .

Hinweis. Sei $f = (z - \alpha_1) \cdots (z - \alpha_n)$ und sei $\beta \in \mathbb{C}$ mit $f'(\beta) = 0$. Multipliziere die Gleichung $\overline{f'(\beta)} = 0$ mit $\prod_{i=1}^n (\beta - \alpha_i)$.

1.2. DER SATZ VON CARATHÉODORY

Satz 1.3 (Carathéodory). *Es sei $S \subset \mathbb{R}^d$ eine Teilmenge. Jeder Punkt der konvexen Hülle von S ist eine Konvexkombination von höchstens $d + 1$ Punkten aus S .*

Beweis. Sei $x = \alpha_1 y_1 + \dots + \alpha_m y_m$ eine Konvexkombination von Punkten $y_1, \dots, y_m \in S$ und angenommen $m > d + 1$. Wir müssen zeigen, dass wir x auch durch eine kürzere Konvexkombination erhalten können.

Betrachte das homogene lineare Gleichungssystem

$$\lambda_1 y_1 + \dots + \lambda_m y_m = 0 \quad \text{und} \quad \lambda_1 + \dots + \lambda_m = 0$$

in m reellen Unbestimmten $\lambda_1, \dots, \lambda_m$. Da die y_i aus \mathbb{R}^d sind, hat dieses System $d + 1$ Gleichungen und somit nach Voraussetzung weniger als die Anzahl der Unbestimmten. Es gibt also eine nicht-triviale Lösung $\lambda_1, \dots, \lambda_m$. Da sich die λ_i zu 0 summieren aber nicht alle 0 sind, gibt es mindestens ein i mit $\lambda_i > 0$. Sei

$$\tau = \min \left\{ \frac{\alpha_i}{\lambda_i} \mid \lambda_i > 0 \right\}$$

und setze

$$\beta_i = \alpha_i - \tau \lambda_i \quad \text{für } i = 1, \dots, m.$$

Es gelten nun $\beta_i \geq \alpha_i - \frac{\alpha_i}{\lambda_i} \lambda_i = 0$, $\sum_{i=1}^m \beta_i = \sum_{i=1}^m \alpha_i - \tau \sum_{i=1}^m \lambda_i = 1$ und

$$\beta_1 y_1 + \dots + \beta_m y_m = \alpha_1 y_1 + \dots + \alpha_m y_m - \tau (\lambda_1 y_1 + \dots + \lambda_m y_m) = x.$$

Wir haben die Darstellung dabei tatsächlich verkürzt: Denn sei j ein Index mit $\tau = \frac{\alpha_j}{\lambda_j}$, dann ist nach Konstruktion $\beta_j = 0$, so dass wir y_j weglassen können. \square

Die $d + 1$ Punkte im Satz von Carathéodory hängen natürlich von x ab. Sonst würde der Satz ja zeigen, dass jede konvexe Menge ein Polytop ist. Aber auch für Polytope können die Punkte im allgemeinen nicht unabhängig von x gewählt werden. Das zeigt schon das Beispiel eines Rechtecks in der Ebene als konvexe Hülle seiner vier Eckpunkte.

Übung 1.6. Die Schranke im Satz von Carathéodory ist die bestmögliche: Für jedes $d > 0$ gibt es eine Teilmenge S von \mathbb{R}^d und einen Punkt $x \in \text{conv}(S)$, der nicht als Konvexkombination von d Punkten aus S geschrieben werden kann.

Übung 1.7. Eine Teilmenge $S \subset \mathbb{R}^d$ heißt *wegzusammenhängend* wenn es zu je zwei Punkten $x, y \in S$ eine stetige Abbildung $\varphi: [0, 1] \rightarrow \mathbb{R}^d$ gibt derart, dass $\varphi(0) = x$, $\varphi(1) = y$ und $\varphi(t) \in S$ für alle $t \in [0, 1]$. Zeige: Ist S die Vereinigung von höchstens d wegzusammenhängenden Teilmengen von \mathbb{R}^d , dann ist jeder Punkt in $\text{conv}(S)$ eine Konvexkombination von höchstens d Punkten aus S .

Korollar 1.4. *Die konvexe Hülle einer kompakten Teilmenge von \mathbb{R}^d ist kompakt.*

Beweis. Sei $S \subset \mathbb{R}^d$ eine kompakte Teilmenge und sei $\Delta_d \subset \mathbb{R}^{d+1}$ der d -dimensionale Standard-Simplex, wie oben definiert. Dieser ist kompakt und damit auch das kartesische Produkt $S^{d+1} \times \Delta_d$. Definiere nun die Abbildung $\Phi: S^{d+1} \times \Delta_d \rightarrow \mathbb{R}^d$ durch

$$\Phi(u_1, \dots, u_d; \alpha_1, \dots, \alpha_{d+1}) = \alpha_1 u_1 + \dots + \alpha_{d+1} u_{d+1}.$$

Das Bild von Φ ist offensichtlich in $\text{conv}(S)$ enthalten und der Satz von Carathéodory sagt gerade, dass Φ surjektiv auf $\text{conv}(S)$ abbildet. Da Φ stetig ist und stetige Bilder kompakter Mengen kompakt sind, folgt die Behauptung. \square

Übung 1.8. Gib ein Beispiel einer abgeschlossenen Teilmenge von \mathbb{R}^2 , deren konvexe Hülle nicht abgeschlossen ist.

1.3. DIE SÄTZE VON RADON UND HELLY

Mit einer ähnlichen Idee wie im Beweis des Satzes von Carathéodory kann man noch einen anderen schönen Satz zeigen:

Satz 1.5 (Radon). *Es sei $S \subset \mathbb{R}^d$ eine Menge von mindestens $d + 2$ Punkten. Dann gibt es zwei Teilmengen $R \subset S$ („rote Punkte“) und $B \subset S$ („blaue Punkte“) derart, dass*

$$R \cap B = \emptyset \quad \text{und} \quad \text{conv}(R) \cap \text{conv}(B) \neq \emptyset.$$

Beweis. Seien $y_1, \dots, y_m \in S$ verschiedene Punkte, $m \geq d + 2$. Betrachte wieder das homogene lineare Gleichungssystem

$$\lambda_1 y_1 + \dots + \lambda_m y_m = 0 \quad \text{und} \quad \lambda_1 + \dots + \lambda_m = 0,$$

das wegen $m \geq d + 2$ eine nicht-triviale Lösung $\lambda_1, \dots, \lambda_m$ besitzt. Setze

$$I_R = \{i \mid \lambda_i > 0\} \quad \text{und} \quad I_B = \{i \mid \lambda_i < 0\}$$

und

$$R = \{y_i \mid i \in I_R\} \quad \text{und} \quad B = \{y_i \mid i \in I_B\}.$$

Natürlich gilt $R \cap B = \emptyset$. Sei $\beta = \sum_{i \in I_R} \lambda_i$, dann folgt $\sum_{i \in I_B} \lambda_i = -\beta$, da sich die λ_i zu 0 summieren. Da außerdem $\lambda_1 y_1 + \dots + \lambda_m y_m = 0$ gilt, folgt

$$\sum_{i \in I_R} \lambda_i y_i = \sum_{i \in I_B} (-\lambda_i) y_i \quad \text{und damit} \quad \sum_{i \in I_R} \frac{\lambda_i}{\beta} y_i = \sum_{i \in I_B} \frac{-\lambda_i}{\beta} y_i.$$

Also ist $\text{conv}(R) \cap \text{conv}(B) \neq \emptyset$. □

Beispiel 1.6. Für vier Punkte in der Ebene gibt der Satz von Radon die folgende Dichotomie: Entweder einer von ihnen liegt in der konvexen Hülle der anderen drei, oder sie lassen sich in zwei Paare aufteilen, deren Verbindungsstrecken sich schneiden.

Satz 1.7 (Helly). *Es seien C_1, \dots, C_m konvexe Teilmengen von \mathbb{R}^d , $m \geq d + 1$. Angenommen jede Wahl von $d + 1$ dieser Mengen hat einen nicht-leeren Schnitt, d.h. es gelte*

$$C_{i_1} \cap \dots \cap C_{i_{d+1}} \neq \emptyset \quad \text{für alle verschiedenen } i_1, \dots, i_{d+1} \in \{1, \dots, m\}.$$

Dann gilt auch

$$C_1 \cap \dots \cap C_m \neq \emptyset.$$

Beweis. Der Beweis benutzt den Satz von Radon und Induktion nach m , beginnend mit $m = d + 1$. Für $m = d + 1$ ist die Folgerung gleich der Voraussetzung und die Aussage damit wahr. Sei $m > d + 1$. Für jedes $i = 1, \dots, m$ können wir einen Punkt p_i aus $\bigcap_{j \neq i} C_j$ wählen, denn dieser Durchschnitt ist nach Induktionsvoraussetzung nicht leer. Falls es $j \neq i$ mit $p_i = p_j$ gibt, so gilt $p_i \in \bigcap_{i=1}^m C_i$ und wir sind fertig. Andernfalls gibt es nach dem Satz von Radon eine disjunkte Unterteilung $\{p_1, \dots, p_m\} = R \cup B$ und einen Punkt

$$p \in \text{conv}(R) \cap \text{conv}(B).$$

Wir zeigen, dass dieser magische Punkt p in allen C_i enthalten ist: Es gilt $R \subset C_i$ für alle i mit $p_i \in R$ und analog $B \subset C_i$ für alle i mit $p_i \in B$. Da die Mengen C_i konvex sind, folgt

$$p \in \text{conv}(R) \subset \bigcap_{i: p_i \in R} C_i \quad \text{und} \quad p \in \text{conv}(B) \subset \bigcap_{i: p_i \in B} C_i.$$

Wegen $R \cap B = \emptyset$ folgt daraus die Behauptung. □

Übung 1.9. Finde ein Beispiel einer unendlichen Familie C_1, C_2, \dots von konvexen Mengen derart, dass je $d + 1$ von ihnen einen nicht-leeren Schnitt haben, aber trotzdem $\bigcap_{i=1}^{\infty} C_i = \emptyset$ gilt.

2. TRENNENDE HYPEREBENEN

2.1. HYPEREBENEN

Definition 2.1. Es sei V ein reeller Vektorraum. Eine lineare Abbildung $\ell: V \rightarrow \mathbb{R}$ heißt *lineares Funktional auf V* . Der Vektorraum $\text{Hom}(V, \mathbb{R})$ aller linearen Funktionalen auf V heißt der Dualraum von V und wird mit V^\vee bezeichnet. Eine Teilmenge H von V heißt *Hyperebene in V* , falls es ein lineares Funktional $\ell \neq 0$ auf V gibt und ein $\alpha \in \mathbb{R}$ mit

$$H = \{v \in V \mid \ell(v) = \alpha\}.$$

Der Vorteil dieser abstrakten Definition von Hyperebene ist, dass sie auch für unendlich-dimensionale Vektorräume sinnvoll ist. Im endlich-dimensionalen Fall gibt es bekanntlich eine ganze Reihe äquivalenter Beschreibungen:

- (1) Eine Hyperebene H in \mathbb{R}^d ist ein affin-linearer Unterraum der Dimension $d - 1$. D.h., es gibt einen linearen Unterraum U der Dimension $d - 1$ und ein $b \in \mathbb{R}^d$ mit

$$H = U + b = \{u + b \mid u \in U\}.$$

Es gilt dann $H = U + b'$ für jedes $b' \in H$, die Wahl des Basispunkts b spielt also keine Rolle. Ist $H = \{v \in V \mid \ell(v) = \alpha\}$ wie oben, so gilt $U = \ker(\ell)$ und $\alpha = \ell(b)$.

- (2) Die Hyperebene H ist auch die Lösungsmenge einer einzelnen linearen Gleichung $c_1x_1 + \dots + c_dx_d = \alpha$ mit Koeffizientenvektor $c \in \mathbb{R}^d \setminus \{0\}$ und Variablenvektor x . Dann ist U der Lösungsraum des homogenen Systems (also mit 0 auf der rechten Seite) und b eine spezielle Lösung.
- (3) Wir können auch $c_1x_1 + \dots + c_dx_d = \langle c, x \rangle$ als Skalarprodukt schreiben. Weil U die Dimension $d - 1$ hat, ist das orthogonale Komplement eindimensional, d.h. es gilt

$$U^\perp = \{x \in \mathbb{R}^d \mid \langle x, u \rangle = 0 \text{ für alle } u \in U\} = \text{span}(c)$$

und damit $U = \text{span}(c)^\perp$ und

$$H = \{x \in \mathbb{R}^d \mid \langle c, x \rangle = \langle c, b \rangle\}.$$

Das Funktional ℓ wie in der Definition bekommen wir dann als $\ell(x) = \langle x, c \rangle$ zurück, und es gilt $\alpha = \langle c, b \rangle$.

Es sei $H = \{v \in V \mid \ell(v) = \alpha\}$ mit $\ell \neq 0$ eine Hyperebene in einem reellen Vektorraum V . Dann definieren wir

$$H_+ = \{v \in V \mid \ell(v) > \alpha\} \quad \text{und} \quad H_- = \{v \in V \mid \ell(v) < \alpha\}$$

und nennen H_+ und H_- die durch H bestimmten *offenen Halbräume*. Die Notation ist etwas missbräuchlich, denn wenn man ℓ durch $-\ell$ und α durch $-\alpha$ ersetzt, dann ändert sich H nicht, aber H_+ und H_- werden vertauscht. Die Notation ist also nur gerechtfertigt,

nachdem ℓ festgelegt wurde. Wir verwenden sie aber auch, wenn das Vorzeichen einfach keine Rolle spielt. Weiter definieren wir auch noch

$$\overline{H_+} = \{v \in V \mid \ell(v) \geq \alpha\} \quad \text{und} \quad \overline{H_-} = \{v \in V \mid \ell(v) \leq \alpha\},$$

die *abgeschlossenen Halbräume*. Und noch ein paar nachlässige Kurzschreibweisen:

$$H = \{\ell = \alpha\} \quad H_+ = \{\ell > \alpha\} \quad \text{usw.}$$

Regeln 2.2. Es gelten

$$V = H \uplus H_+ \uplus H_- \quad \text{und} \quad H = \overline{H_+} \cap \overline{H_-}.$$

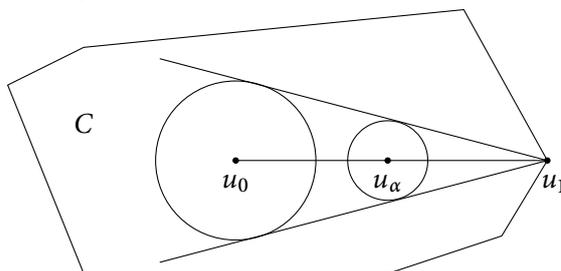
Für $V = \mathbb{R}^d$ ist außerdem $\overline{H_+}$ der (topologische) Abschluss von H_+ , wie durch die Balkennotation suggeriert wird, und H der Rand von $\overline{H_+}$; entsprechend für H_- . (An diese topologischen Begriffe wird im nächsten Abschnitt erinnert).

2.2. DIMENSION UND TOPOLOGIE IM EUKLIDISCHEN RAUM

Es sei $S \subset \mathbb{R}^d$ eine Teilmenge. Erinnerung: Das *Innere* von S besteht aus allen Punkten $x \in S$ für die es $\varepsilon > 0$ mit $B_\varepsilon(x) \subset S$ gibt und wird mit $\text{int}(S)$ bezeichnet. Dabei ist $B_\varepsilon(x) = \{y \in \mathbb{R}^d \mid \|y - x\| < \varepsilon\}$ die offene Kugel mit Radius ε um x . Der *Abschluss* $\text{clos}(S)$ von S ist das Komplement von $\text{int}(\mathbb{R}^d \setminus S)$ und der *Rand* von S ist $\partial S = \text{clos}(S) \setminus \text{int}(S)$.

Übung 2.1. Sei $C \subset \mathbb{R}^d$ eine konvexe Teilmenge. Zeige:

- (1) Sei $u_0 \in \text{int}(C)$. Für jedes $u_1 \in C$ und jedes $0 \leq \alpha < 1$ gilt $u_\alpha := (1 - \alpha)u_0 + \alpha u_1 \in \text{int}(C)$.
Hinweis. Betrachte das folgende Bild:



- (2) Die Menge $\text{int}(C)$ ist konvex.
 (3) Ist $\text{int}(C) \neq \emptyset$, so gilt $C \subset \text{clos}(\text{int}(C))$.

Die *Dimension* $\dim(C)$ einer konvexen Teilmenge $C = \text{conv}(S)$ von \mathbb{R}^d ist die Dimension des kleinsten affinen Unterraums von \mathbb{R}^d , der C enthält. Dieser ist genau

$$\text{aff}(S) = \left\{ \sum_{i=1}^k \alpha_i v_i \mid v_1, \dots, v_k \in S, \alpha_1, \dots, \alpha_k \in \mathbb{R} \text{ mit } \sum_{i=1}^k \alpha_i = 1, k \in \mathbb{N} \right\},$$

die *affine Hülle* von S . Ihre Elemente heißen die *Affinkombinationen* von Punkten von S . Der Unterschied zu den Konvexkombinationen besteht darin, dass die α_i hier auch negativ sein dürfen. Die affine Hülle ist ein affiner Unterraum, also ein verschobener linearer Unterraum, wie oben. Ihre Dimension ist per Definition genau jene des Unterraums. Details, die man sich hierzu überlegen muss, stecken in der folgenden Übung:

Übung 2.2. Es sei $S \subset \mathbb{R}^d$ eine Teilmenge. Für jedes $u \in \text{aff}(S)$ ist dann $L := \text{aff}(S) - u := \{v - u \mid v \in \text{aff}(S)\}$ ein linearer Unterraum, der nicht von der Wahl von u abhängt. Die Dimension von $\text{aff}(S)$ ist per Definition die Dimension von L . Es gilt dann also $\dim(\text{conv}(S)) = \dim(\text{aff}(S)) = \dim(L)$.

Wir brauchen den folgenden, sehr anschaulichen, Zusammenhang zwischen Dimension und Topologie:

Proposition 2.3. Eine konvexe Teilmenge von \mathbb{R}^d hat genau dann die Dimension d , wenn ihr Inneres nicht leer ist.

Beweis. Sei $C \subset \mathbb{R}^d$ konvex. Falls $\text{int}(C) \neq \emptyset$, dann können wir ohne Einschränkung $0 \in \text{int}(C)$ annehmen. Es gibt dann $\varepsilon > 0$ mit $B_\varepsilon(0) \subset C$, also ist $\text{span}(C) \supset \text{span}(B_\varepsilon(0)) = \mathbb{R}^d$, so dass C die Dimension d hat. Für die umgekehrte Richtung gelte $\dim(C) = d$. Wieder können wir $0 \in C$ annehmen, also $\text{aff}(C) = \text{span}(C)$ und $\dim(\text{span}(C)) = d$. Dann gibt es d linear unabhängige Punkte $x_1, \dots, x_d \in C$, und es genügt zu zeigen, dass das Polytop $P = \text{conv}(0, x_1, \dots, x_d)$ nicht-leeres Inneres besitzt. Es gilt $P = \{\sum \alpha_i x_i \mid \sum_{i=1}^d \alpha_i \leq 1 \text{ und } \alpha_i \geq 0 \text{ für alle } i = 1, \dots, d\}$. Da die x_1, \dots, x_d eine Basis von \mathbb{R}^d sind, hat jedes $v \in \mathbb{R}^d$ eine eindeutige Darstellung $v = \sum_{i=1}^d \alpha_i x_i$ und die α_i hängen stetig von v ab. Daraus folgt $\text{int}(P) = \{\sum \alpha_i x_i \mid \sum_{i=1}^d \alpha_i < 1 \text{ und } \alpha_i > 0 \text{ für alle } i = 1, \dots, d\} \neq \emptyset$. \square

Definition 2.4. Es sei $C \subset \mathbb{R}^d$ eine konvexe Menge und $L = \text{aff}(C)$. Das *relative Innere* $\text{relint}(C)$ von C ist das Innere von C als Teilmenge von $L \cong \mathbb{R}^{\dim(L)}$. Explizit heißt das

$$\text{relint}(C) = \{x \in C \mid \exists \varepsilon > 0: B_\varepsilon(x) \cap L \subset C\}.$$

Damit ist das Folgende eine direkte Konsequenz von Prop. 2.3 und Übung 2.1:

Korollar 2.5. Das relative Innere einer konvexen nicht-leeren Teilmenge von \mathbb{R}^d ist ebenfalls konvex und nicht leer. \square

Definition 2.6. Es sei V ein reeller Vektorraum. Eine Teilmenge A von V heißt *algebraisch offen*, wenn $A \cap L$ für jede Gerade L in V eine offene Teilmenge von L ist. Ist $C \subset V$ konvex, so heißt dies explizit das Folgende: Für jede parametrisierte Gerade

$$L = \{v + \tau u \mid \tau \in \mathbb{R}\}, \quad u, v \in \mathbb{R}^d, \quad u \neq 0$$

in V gebe es $\alpha, \beta \in \mathbb{R} \cup \{\infty, -\infty\}$ mit $\alpha \leq \beta$ derart, dass

$$C \cap L = \{v + \tau u \mid \alpha < \tau < \beta\}.$$

(Dabei ist $\alpha \geq \beta$ auch erlaubt und entspricht dem Fall $C \cap L = \emptyset$).

Übung 2.3.

- (1) Zeige, dass eine konvexe Teilmenge von \mathbb{R}^d genau dann offen ist, wenn sie algebraisch offen ist.
- (2) Gib ein Beispiel einer nicht-konvexen Teilmenge von \mathbb{R}^d die algebraisch offen ist, aber nicht offen.

2.3. DER ISOLATIONSSATZ

Definition 2.7. Es seien $A, B \subset V$ zwei Teilmengen und $H = \{\ell = \alpha\}$ eine Hyperebene. Wir sagen,

- H trennt A und B , falls

$$(A \subset \overline{H_+} \text{ und } B \subset \overline{H_-}) \quad \text{oder} \quad (A \subset \overline{H_-} \text{ und } B \subset \overline{H_+}).$$

- H trennt A und B strikt, falls

$$(A \subset H_+ \text{ und } B \subset H_-) \quad \text{oder} \quad (A \subset H_- \text{ und } B \subset H_+).$$

- H isoliert A , falls A und \emptyset von H getrennt werden, also $A \subset \overline{H_+}$ oder $A \subset \overline{H_-}$.
- H isoliert A strikt, falls $A \subset H_+$ oder $A \subset H_-$.

Satz 2.8 (Isolationssatz). *Sei V ein reeller Vektorraum und $C \subset V$ eine algebraisch offene konvexe Teilmenge. Zu jedem Punkt $u \in V \setminus C$ existiert eine Hyperebene durch u , die C strikt isoliert.*

Beweis. Ohne Einschränkung können wir $u = 0$ annehmen. Wir nehmen weiter $\dim(V) \geq 2$ und $C \neq \emptyset$ an, denn in den übrigen Fällen ist die Aussage trivial.

Wir behandeln zunächst ziemlich explizit den Fall $V = \mathbb{R}^2$. Sei $S = \{x \in \mathbb{R}^2 \mid \|x\| = 1\}$ der Einheitskreis und sei $\varphi: \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow S$ die radiale Projektion $x \mapsto x/\|x\|$. Da C offen und zusammenhängend ist, ist $\varphi(C)$ ein offener Kreisbogen

$$\varphi(C) = \{(\cos \vartheta, \sin \vartheta) \mid \alpha < \vartheta < \beta\}$$

für $\alpha, \beta \in [0, 2\pi)$. (Bild!) Da C konvex ist und $0 \notin C$, muss die Bogenlänge von $\varphi(C)$ kleiner als π sein, zu jedem $v \in \varphi(C)$ liegt also der Antipodenpunkt $-v$ nicht in $\varphi(C)$. Ist nun $v \in \partial\varphi(C)$ und H die Gerade durch $u = 0$ und v , so wird C durch H strikt isoliert.

Sei nun $\dim(V) \geq 2$ und P eine 2-dimensionale Ebene durch 0 in V . Der Durchschnitt $P \cap C$ ist eine konvexe, algebraisch offene Teilmenge von $P \cong \mathbb{R}^2$, die 0 nicht enthält. Wie bereits gezeigt gibt es also eine Gerade $L \subset P$ mit $0 \in L$ und $L \cap C = \emptyset$. Wähle nun einen linearen Unterraum H von V , der maximal ist bzgl. der Eigenschaft $H \cap C = \emptyset$. D.h. wenn H' ein linearer Unterraum von V ist mit $H \subset H'$ und $H' \cap C = \emptyset$, dann gilt $H' = H$. Die Menge aller solcher Unterräume ist nicht-leer, denn sie enthält die Gerade L . Die Existenz eines maximalen solchen Unterräume folgt damit aus dem Zornschen Lemma.

Wir müssen zeigen, dass H eine Hyperebene ist. Wegen $C \neq \emptyset$ ist $H \neq V$. Sei $\rho: V \rightarrow V/H$ die Abbildung auf den Quotientenvektorraum. Angenommen H ist keine Hyperebene, dann gilt $\dim(V/H) \geq 2$, und man überzeugt sich, dass $\rho(C)$ eine konvexe, algebraisch offene Teilmenge von V/H ist. Wir können also wieder unsere Konstruktion oben anwenden, d.h. es gibt eine Gerade $L \subset V/H$ mit $0 \in L$ und $L \cap \rho(C) = \emptyset$. Dann ist das Urbild $G = \rho^{-1}(L)$ ein linearer Unterraum von V mit $G \cap C = \emptyset$ und $H \subsetneq G$, ein Widerspruch. \square

Übung 2.4. Zeige: Jede Hyperebene, die eine algebraisch offene Menge isoliert, isoliert diese strikt.

Der Isolationssatz ist der Ausgangspunkt für viele Trennungssätze, die es in zahlreichen Varianten gibt. Die Frage, unter welchen Voraussetzungen zwei konvexe Mengen getrennt oder strikt getrennt werden können, kann subtiler sein, als man zunächst glauben möchte, vor allem in unendlicher Dimension. Siehe dazu etwa folgendes Beispiel:

Übung 2.5. Sei $V = \mathbb{R}[t]$ der Vektorraum der Polynome in einer Variablen t und sei C die Teilmenge aller Polynome mit strikt positivem Leitkoeffizient. Zeige:

- (1) Die Teilmenge C ist konvex und nicht algebraisch offen.
- (2) Es gibt keine Hyperebene H durch 0 , die C isoliert.

Definition 2.9. Es sei $C \subset \mathbb{R}^d$ eine konvexe Teilmenge und sei $u \in \partial C$. Eine Hyperebene durch u , die C isoliert, wird *Stützhyperebene* in u an C genannt.

Korollar 2.10. *Es sei $C \subset \mathbb{R}^d$ eine konvexe Teilmenge. Zu jedem $u \in \partial C$ existiert eine Stützhyperebene in u an C .*

Beweis. Wir können $\text{int}(C) \neq \emptyset$ annehmen. Andernfalls ist C nach Satz 2.3 in einem affinen Unterraum U enthalten, und jede Hyperebene, die U enthält, isoliert C . Nach Übung 2.1 ist $\text{int}(C)$ eine konvexe Menge und nach Übung 2.3 auch algebraisch offen. Deshalb können wir den Isolationssatz anwenden und erhalten zu $u \in \partial C$ eine Hyperebene H durch u , die $\text{int}(C)$ strikt isoliert, etwa $\text{int}(C) \subset H_+$. Es folgt dann $C \subset \text{clos}(\text{int}(C)) \subset \text{clos}(H_+) = \overline{H_+}$ (siehe Übung 2.1(3)), also wird C durch H isoliert. \square

2.4. EXKURS: DER SATZ VON KIRCHBERGER

Im Zusammenhang mit trennenden Hyperebenen diskutieren wir noch kurz eine Anwendung des Satzes von Helly (1.7) aus der ersten Woche.

Satz 2.11 (Kirchberger). *Es seien $R, B \subset \mathbb{R}^d$ zwei endliche Teilmengen. Angenommen, die folgende Bedingung ist erfüllt: Für jede Teilmenge S von $R \cup B$ mit höchstens $d + 2$ Punkten gibt es eine Hyperebene H in \mathbb{R}^d , die $S \cap R$ und $S \cap B$ strikt trennt. Dann gibt es auch eine Hyperebene H die R und B strikt trennt.*

Beweis. Für jeden Punkt $r \in R$, definiere

$$A_r = \{(c, \alpha) \in \mathbb{R}^{d+1} \mid \langle c, r \rangle < \alpha\}$$

und analog für $b \in B$:

$$B_b = \{(c, \alpha) \in \mathbb{R}^{d+1} \mid \langle c, b \rangle > \alpha\}.$$

Diese Mengen codieren genau die Hyperebenen, die uns interessieren: Einem Paar (c, α) mit $c \neq 0$ ordnen wir die Hyperebene $\{x \in \mathbb{R}^d \mid \langle c, x \rangle = \alpha\}$ zu.

Die Mengen A_r und B_b sind konvex, für alle $r \in R, b \in B$. Der Satz von Helly sagt, dass der Durchschnitt

$$\left(\bigcap_{r \in R} A_r\right) \cap \left(\bigcap_{b \in B} B_b\right)$$

genau dann nicht leer ist, wenn

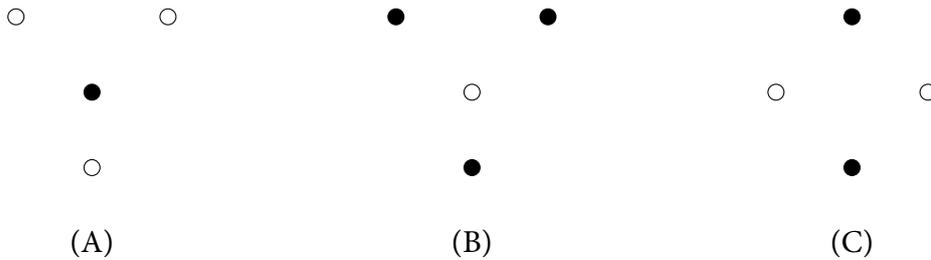
$$\left(\bigcap_{r \in R \cap S} A_r\right) \cap \left(\bigcap_{b \in B \cap S} B_b\right)$$

für jede Teilmenge $S \subset R \cup B$ aus höchstens $d + 2$ Punkten nicht leer ist. Die Teilmengen A_r und B_b sind offen. Endliche Durchschnitte dieser Mengen sind daher auch offen. Es folgt, dass der Durchschnitt

$$\left(\bigcap_{r \in R \cap S} A_r\right) \cap \left(\bigcap_{b \in B \cap S} B_b\right)$$

genau dann nicht leer ist, wenn er einen Punkt (c, α) mit $c \neq 0$ enthält. Ein solcher Punkt entspricht genau der Hyperebene $H = \{x \in \mathbb{R}^d \mid \langle c, x \rangle = \alpha\}$, welche $R \cap S$ und $B \cap S$ strikt trennt. \square

Beispiel 2.12. Seien $R, B \subset \mathbb{R}^2$ zwei Teilmengen der Ebene. Nach dem Satz von Kirchberger können R und B genau dann *nicht* durch eine Gerade getrennt werden, wenn es vier Punkte $S \subset R \cup B$ gibt, die sich bereits nicht trennen lassen. Man kann sich überlegen, dass das genau dann der Fall ist, wenn die vier Punkte qualitativ eine der folgenden drei Konfigurationen A,B,C bilden:



3. SEITEN UND EXTREMALPUNKTE

Es sei immer V ein reeller Vektorraum.

Definition 3.1. Sei $C \subset V$ konvex.

- (1) Eine konvexe Teilmenge F von C heißt *Seite* von C , falls folgendes gilt: Sind $u, v \in C$ und $\lambda \in (0, 1)$ mit $\lambda u + (1 - \lambda)v \in F$, dann gilt $u \in F$ und $v \in F$. Eine Seite F heißt *echt*, falls $\emptyset \subsetneq F \subsetneq C$ gilt.
- (2) Ist $u \in C$ ein Punkt derart, dass $\{u\}$ eine Seite von C ist, so heißt u ein *Extremalpunkt* von C . Die Extremalpunkte sind also genau die nulldimensionalen Seiten. Die Menge aller Extremalpunkte von C wird mit $\text{ex}(C)$ bezeichnet.

Beispiel 3.2. Sei $\Delta = \text{conv}(e_1, \dots, e_{d+1})$ der d -dimensionale Standardsimplex in \mathbb{R}^{d+1} . Die Seiten von Δ sind genau die k -dimensionalen Simplizes

$$\Delta_I = \text{conv}(e_{i_1}, \dots, e_{i_{k+1}}) \quad \text{für} \quad I = \{i_1, \dots, i_{k+1}\} \subset \{1, \dots, d+1\}, \quad |I| = k+1, \quad k \leq d$$

Dabei entspricht der Fall $k = d - 1$ genau den *Facetten*, also dem, woran man bei dem Wort Seite zuerst denkt. Die Extremalpunkte sind gerade die *Ecken* e_i ($i = 1, \dots, d + 1$).

Übung 3.1.

- (1) Zeige, dass man in der Definition von Seite immer $\lambda = \frac{1}{2}$ nehmen kann. D.h. F ist genau dann eine Seite, wenn gilt: Für alle $u, v \in C$ mit $\frac{1}{2}(u + v) \in F$ gilt $u \in F$ und $v \in F$.
- (2) Sei $S \subset V$ eine Teilmenge. Zeige, dass $u \in V$ genau dann ein Extremalpunkt von $\text{conv}(S)$ ist, wenn $u \in S$ und $u \notin \text{conv}(S \setminus \{u\})$.

Übung 3.2. Jede Seite einer Seite von C ist eine Seite von C .

Übung 3.3. Es sei $C \subset V$ konvex.

- (1) Für jede Seite F von C gilt $F = \text{aff}(F) \cap C$.
- (2) Sei $V = \mathbb{R}^d$. Sind F, F' zwei Seiten von C mit $F' \subsetneq F$, so folgt $\dim(F') < \dim(F)$.
- (3) Sei $V = \mathbb{R}^d$. Jede echte Seite von C ist in ∂C enthalten.

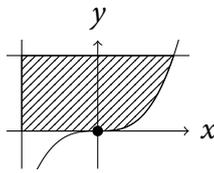
Proposition 3.3. Sei $C \subset V$ konvex und H eine Hyperebene in V , die C isoliert. Dann ist $H \cap C$ eine Seite von C .

Beweis. Es sei $F = H \cap C$, wobei $H = \{\ell = \alpha\}$ mit $\ell \in V^\vee$ und $\alpha \in \mathbb{R}$. Da C durch H isoliert wird, gelte o.E. $C \subset \overline{H_-}$, also $\ell(v) \leq \alpha$ für alle $v \in C$. Sind $u, v \in C$ mit $\frac{1}{2}(u + v) \in F$, so folgt $\alpha = \ell(\frac{1}{2}(u + v)) = \frac{1}{2}(\ell(u) + \ell(v))$. Wegen $\ell(u), \ell(v) \leq \alpha$ folgt daraus $\ell(u) = \ell(v) = \alpha$. (Dies heißt gerade, dass α ein Extremalpunkt des Intervalls $(-\infty, \alpha]$ ist.) Also gilt $u, v \in F$. \square

Definition 3.4. Es sei $C \subset V$ konvex. Eine Seite F von C heißt *exponiert*, wenn es eine Hyperebene H in V gibt, die C isoliert und $H \cap C = F$ erfüllt. Die trivialen Seiten \emptyset und C definiert man ebenfalls als exponiert.

Warnung 3.5. Die Terminologie in der Literatur ist an dieser Stelle nicht einheitlich. Viele Quellen, darunter auch das Buch von Barvinok, bezeichnen mit Seite („face“) das, was wir exponierte Seite nennen. Außerdem ist es üblich, in der Definition von Seite vorauszusetzen, dass C abgeschlossen ist. Das tun wir nur deshalb nicht, um uns nicht schneller als nötig auf endlich-dimensionale Räume festzulegen.

Beispiel 3.6. Es gibt auch nicht exponierte Seiten, sogar nicht-exponierte Extrempunkte:



Die konvexe Menge im Bild wird rechts durch die Kurve $y = x^3$ begrenzt und unten durch die x -Achse. Der Nullpunkt ist ein nicht-exponierter Extrempunkt, denn die x -Achse ist die einzige Stützgerade durch den Nullpunkt. Sie schneidet aber eine größere Seite aus.

Als Folgerung aus dem Isolationssatz 2.8 ergibt sich aber immerhin noch Folgendes:

Korollar 3.7. Sei $C \subset \mathbb{R}^d$ eine abgeschlossene konvexe Teilmenge mit nicht-leerem Inneren und sei $x \in \partial C$ ein Randpunkt. Dann ist x in einer echten exponierten Seite von C enthalten.

Beweis. Nach Kor. 2.10 gibt es eine Stützhyperebene H an C in x . Dann ist $F = C \cap H$ eine exponierte Seite von C . Wegen $\text{int}(C) \neq \emptyset$ und $\text{int}(C) \cap H = \emptyset$ ist F echt. \square

Verzichtet man auf die Forderung nach Exponiertheit der Seite, so gilt sogar:

Proposition 3.8. Sei $C \subset \mathbb{R}^d$ eine abgeschlossene konvexe Teilmenge und $x \in C$ ein Punkt. Dann gibt es eine eindeutig bestimmte Seite von C , die x in ihrem relativen Inneren enthält. Diese ist genau die kleinste Seite von C , die x enthält.

Beweis. Ist \mathcal{F} die Menge aller Seiten von C , die x enthalten, so überzeugt man sich, dass $F = \bigcap \mathcal{F}$ wieder eine Seite ist. Offenbar ist F die kleinste Seite von C , die x enthält. Wegen $C \in \mathcal{F}$ ist $F \neq \emptyset$. Es sei $U = \text{aff}(F)$ die affine Hülle von F . Nach Definition gilt $x \in \text{relint}(F)$ genau dann, wenn x im Inneren von F in $U \cong \mathbb{R}^{\dim(F)}$ liegt. Andernfalls ist x ein Randpunkt von F in U . Nach Kor. 3.7 ist x also in einer echten (exponierten) Seite von F in U enthalten. Diese ist nach Übung 3.2 auch eine Seite von C , im Widerspruch zur Minimalität von F . Die Eindeutigkeit von F sieht man wie folgt: Ist F' eine weitere Seite von C mit $x \in F'$, so folgt $F \subset F'$ nach Definition von F . Ist $F' \neq F$, so ist F eine echte Seite von F' , so dass x nicht in $\text{relint}(F')$ liegen kann (Übung 3.3(3)). \square

Proposition 3.9. Jede echte Seite einer abgeschlossenen konvexen Teilmenge von \mathbb{R}^d ist in einer echten exponierten Seite enthalten.

Beweis. Sei $F \subset C$ eine echte Seite und sei $x \in \text{relint}(F)$ (beachte Kor. 2.5). Nach Kor. 3.7 gibt es eine echte exponierte Seite F' von C , die x enthält. Dann gilt $F \subset F'$, denn andernfalls wäre $F \cap F'$ eine echte Seite von F , die x enthält, im Widerspruch zu $x \in \text{relint}(F)$. \square

Korollar 3.10. Sei $C \subset \mathbb{R}^d$ konvex und abgeschlossen. Jede Seite der Dimension $\dim(C) - 1$ von C ist exponiert.

Beweis. Sei $F \subset C$ eine Seite mit $\dim(F) = \dim(C) - 1$. Nach Aufgabe 3.3(2) ist F die einzige echte Seite von C , die F enthält und damit nach Prop. 3.9 exponiert. \square

Korollar 3.11. Jede Seite einer abgeschlossenen konvexen Teilmenge von \mathbb{R}^d ist abgeschlossen.

Beweis. Sei $C \subset \mathbb{R}^d$ konvex und abgeschlossen und $F \subset C$ eine Seite. Wir führen Induktion nach $n = \dim(C) - \dim(F)$. Für $n = 0$ ist $C = F$ nach Aufgabe 3.3(2). Für $n \geq 1$ ist F nach Prop. 3.9 in einer echten exponierten Seite F' von C enthalten. Da F' echt ist, gilt $\dim(F') - \dim(F) < n$. Nun ist F' abgeschlossen, also auch F nach Induktionsvoraussetzung. \square

Nicht jede abgeschlossene konvexe Menge besitzt einen Extrempunkt, wie man schon am Beispiel eines linearen Unterraums sieht. Der folgende Satz sagt jedoch, dass dies in einem gewissen Sinn die einzige Art von Beispiel ist:

Proposition 3.12. *Sei $C \subset \mathbb{R}^d$ eine nicht-leere abgeschlossene konvexe Teilmenge. Genau dann besitzt C einen Extrempunkt, wenn C keine Gerade enthält.*

Beweis. Angenommen C enthält keine Gerade. Wir benutzen Induktion nach d . Für $d = 0$ ist nichts zu zeigen, sei also $d > 0$. Wegen Prop. 2.3 können wir annehmen, dass C nicht-leeres Inneres besitzt; andernfalls können wir \mathbb{R}^d durch $\text{aff}(C)$ ersetzen und die Induktionsvoraussetzung anwenden. Sei $x \in C$ und sei L eine beliebige Gerade durch x . Nach Voraussetzung gilt $\emptyset \neq L \cap C \subsetneq L$. Sei $x \in L$ ein Randpunkt des Intervalls $C \cap L$. Offenbar gilt $x \in \partial C$. Nach Kor. 3.7 gibt es eine echte Seite F von C , die x enthält. Nun ist F eine abgeschlossene konvexe Teilmenge, die keine Geraden enthält mit $\dim(F) < \dim(C) = d$. Nach Induktionsvoraussetzung hat F einen Extrempunkt x , der nach Übung 3.2 auch ein Extrempunkt von C ist.

Umgekehrt enthalte C die Gerade $L = \{u + \tau v \mid \tau \in \mathbb{R}\}$. Es sei $x \in C$, dann zeigen wir, dass x kein Extrempunkt von C ist. Es gilt $\lambda(u + \tau v) + (1 - \lambda)x = \lambda u + (1 - \lambda)x + \lambda \tau v \in C$ für alle $\lambda \in (0, 1)$ und alle $\tau \in \mathbb{R}$. Mit $\tau' = \lambda \tau$ gilt damit auch

$$\lambda u + (1 - \lambda)x + \tau' v \in C$$

für alle $\tau' \in \mathbb{R}$. Da C abgeschlossen ist, folgt $x + \tau' v \in C$ für alle $\tau' \in \mathbb{R}$. Insbesondere ist x kein Extrempunkt. \square

Übung 3.4.

- (1) Zeige, dass die Menge der Extrempunkte einer abgeschlossenen konvexen Teilmenge von \mathbb{R}^2 abgeschlossen ist.
- (2) Überlege, wie ein Gegenbeispiel für die analoge Aussage in \mathbb{R}^3 aussehen könnte. (Später werden wir ein exakt durch Ungleichungen beschriebenes Gegenbeispiel sehen.)

Zum Abschluss noch der folgende zentrale Satz im endlich-dimensionalen Fall:

Satz 3.13 (Krein-Milman). *Eine kompakte und konvexe Teilmenge von \mathbb{R}^d ist die konvexe Hülle ihrer Extrempunkte.*

Beweis. Sei $K \subset \mathbb{R}^d$ kompakt und konvex. Wir beweisen die Aussage mit Induktion nach d . Für $d = 0$ ist es klar, es sei also $d > 0$. Falls $\text{int}(K) = \emptyset$, dann ist K nach Prop. 2.3 in einem echten affinen Unterraum enthalten und die Aussage folgt nach Induktionsvoraussetzung. Wir können also $\text{int}(K) \neq \emptyset$ annehmen. Es sei $u \in K$. Wir müssen zeigen, dass u eine Konvexkombination von Extrempunkten ist. Falls $u \in \partial K$, dann gibt es nach Kor. 2.10 eine echte Seite F von K mit $u \in F$. Es gilt $\dim(F) < \dim(K)$, deshalb folgt die Behauptung wieder nach Induktionsannahme.

Es bleibt der Fall $u \in \text{int}(K)$. Sei $L \subset \mathbb{R}^d$ eine Gerade mit $u \in L$. Dann ist $L \cap K$ ein Intervall in L , etwa $L \cap K = [a, b]$ mit $a, b \in \partial K$. (Hier brauchen wir, dass K kompakt ist.) Wieder sind a, b in echten Seiten enthalten, also $a, b \in \text{conv}(\text{ex}(K))$ nach Induktionsvoraussetzung. Da u eine Konvexkombination von a und b ist, folgt die Behauptung. \square

Der Beweis ist deshalb so einfach, weil wir uns auf den Fall $V = \mathbb{R}^d$ beschränkt haben. Er illustriert aber sehr gut die Nützlichkeit von Seiten (und damit letztlich des Isolationsatzes) in Induktionsbeweisen. Der volle Satz von Krein-Milman gilt auch in unendlich-dimensionalen Räumen (mit einem geeigneten Begriff von Kompaktheit) und ist deutlich schwieriger zu beweisen.

4. POLYEDER

4.1. DIE SEITENSTRUKTUR VON POLYEDERN

Wie aus der Einleitung bekannt, ist ein Polyeder eine konvexe Teilmenge von \mathbb{R}^d , die durch endlich viele lineare Ungleichungen beschrieben ist.

Definition 4.1. Die Extrempunkte eines Polyeders heißen *Ecken* und die eindimensionalen Seiten *Kanten*. Die Seiten von Codimension 1 heißen *Facetten*. (Dabei bedeutet Codimension 1, dass $\dim(F) = \dim(P) - 1$ ist, wobei P das Polyeder und F die Facette ist.)

Satz 4.2 (Seitenstruktur von Polyedern). *Es sei*

$$P = \left\{ x \in \mathbb{R}^d \mid \langle c_1, x \rangle \leq \beta_1, \dots, \langle c_k, x \rangle \leq \beta_k \right\},$$

ein Polyeder in \mathbb{R}^d , wobei $c_1, \dots, c_k \in \mathbb{R}^d$ und $\beta_1, \dots, \beta_k \in \mathbb{R}$. Für jedes $x \in \mathbb{R}^d$ sei

$$I(x) = \{ i \in \{1, \dots, k\} \mid \langle c_i, x \rangle = \beta_i \}$$

die Indexmenge der in x aktiven (d.h. also verschwindenden) Ungleichungen.

(1) Sei $u \in P$. Dann ist

$$F(u) = \{ x \in P \mid I(u) \subset I(x) \}$$

eine exponierte Seite von P mit $\text{relint}(F) = \{ x \in P \mid I(u) = I(x) \}$.

(2) Das Polyeder P hat nur endlich viele Seiten, und diese sind alle exponiert.

(3) Sei $u \in P$ und $L(u) = \text{span}\{c_i \mid i \in I(u)\}$. Es gilt $\text{aff}(F(u)) - u = L(u)^\perp$ und damit

$$\dim(F(u)) = d - \dim(L(u)).$$

Insbesondere ist u genau dann eine Ecke, wenn $L(u) = \mathbb{R}^d$ gilt.

(4) Sind $F'' \subset F$ zwei nicht-leere Seiten von P mit $\dim(F'') \leq \dim(F) - 2$, so gibt es eine Seite F' mit $\dim(F'') < \dim(F') < \dim(F)$ und $F'' \subset F' \subset F$.

Genauer gilt: Sei $e = \dim(P) \geq 2$ und seien h, k ganze Zahlen mit $0 \leq h < k < e$. Ist F eine h -dimensionale Seite von P , so ist F der Durchschnitt aller k -dimensionalen Seiten von P , die F enthalten, und es gibt mindestens $k - h + 1$ solche Seiten.

Beweis. (1) Es sei $c = \sum_{i \in I(u)} c_i$, $\beta = \sum_{i \in I(u)} \beta_i$ und $H = \{x \in \mathbb{R}^d \mid \langle c, x \rangle = \beta\}$. Nach Definition gelten $P \subset \overline{H}$ und $F(u) = H \cap P$, also ist $F(u)$ eine exponierte Seite von P . Ist $x \in F$ mit $I(x) = I(u)$, so folgt $\langle c_i, u \rangle < \beta_i$ für alle $i \notin I(u)$. Also gibt es $\varepsilon > 0$ derart, dass $B_\varepsilon(u) \cap H \subset F(u)$ gilt, also $x \in \text{relint}(F(u))$.

(2) Sei F eine Seite von P . Für $F = \emptyset$ ist nichts zu zeigen. Andernfalls, sei $u \in \text{relint}(F)$. Dann ist F nach Prop. 3.8 die eindeutig bestimmte Seite von P , die u im relativen Inneren enthält. Also ist $F = F(u)$ und damit exponiert nach (1). Ferner hängt F nur von $I(u)$ ab, und für $I(u)$ gibt es abhängig von u nur endlich-viele Möglichkeiten.

(3) Sei $x \in \text{aff}(F(u))$, etwa $x = \sum \alpha_j x_j$ mit $x_j \in F(u)$ und $\sum \alpha_j = 1$. Dann gilt für alle $i \in I(u)$: $\langle c_i, x \rangle = \sum \alpha_j \langle c_i, x_j \rangle = \sum \alpha_j \beta_i = \beta_i = \langle c_i, u \rangle$. Also $\langle c_i, x - u \rangle = 0$ für alle $i \in I(u)$ und damit $x - u \in L(u)^\perp$. Damit ist $\text{aff}(F(u)) - u \subset L(u)^\perp$ gezeigt. Umgekehrt, sei $x \in L(u)^\perp$. Für $i = 1, \dots, k$ und jedes $\varepsilon \in \mathbb{R}$ gilt dann

$$\langle c_i, u \pm \varepsilon x \rangle = \langle c_i, u \rangle \pm \varepsilon \langle c_i, x \rangle \begin{cases} = \beta_i & \text{falls } i \in I(u) \\ < \beta_i + \varepsilon \langle c_i, x \rangle & \text{falls } i \notin I(u). \end{cases}$$

Daraus folgt $u \pm \varepsilon x \in F(u)$ für hinreichend kleines $\varepsilon > 0$ und damit auch $x \in \text{aff}(F(u)) - u$.

(4) Indem wir wie üblich \mathbb{R}^d durch $\text{aff}(P)$ ersetzen, können wir $e = d$ und damit $\text{int}(P) \neq \emptyset$ annehmen.

Es sei F eine Seite von P mit $\dim(F) \leq \dim(P) - 2$. Wir können $0 \in F$ annehmen und setzen $U = \text{aff}(F) = \text{span}(F)$ und $h = \dim(F) = \dim(U)$. Es sei H eine Hyperebene in \mathbb{R}^d mit $F = H \cap P$. Durch einen geeigneten Koordinatenwechsel auf \mathbb{R}^d können wir erreichen, dass $H = \{x \in \mathbb{R}^d \mid \langle e_1, x \rangle = 0\} = \{x \in \mathbb{R}^d \mid x_1 = 0\}$, $U = \{x \in \mathbb{R}^d \mid x_1 = \dots = x_{d-h} = 0\}$ und $P \subset \overline{H_+}$ gelten. Für jedes $\varphi \in [0, \pi)$, setze nun

$$v_\varphi = (\sin(\varphi), \cos(\varphi), 0, \dots, 0)^T \in \mathbb{R}^d$$

und $H_\varphi = \{x \in \mathbb{R}^d \mid \langle x, v_\varphi \rangle = 0\}$. Dann ist $H_{\pi/2} = H$ und $\bigcup_{\varphi \in [0, \pi)} H_\varphi = \mathbb{R}^d$. Setze $B = \{\varphi \in [0, \pi) \mid P \subset (H_\varphi)_+\}$. Indem wir gegebenenfalls H durch H_φ für geeignetes φ ersetzen, können wir außerdem annehmen, dass $H_0 \cap \text{int}(P) \neq \emptyset$ gilt. Die Menge B ist abgeschlossen in $[0, \pi)$. Setze $\alpha = \min(B)$ und $\beta = \max(B)$. Nach Konstruktion gilt dann $0 < \alpha < \pi/2 < \beta < \pi$. Nun zeigt man, dass $F_\alpha = H_\alpha \cap P$ und $F_\beta = H_\beta \cap P$ zwei echte Seiten von P sind, die F strikt enthalten (und damit strikt größere Dimension als F haben). Außerdem gilt $F_\alpha \cap F_\beta = F$.

Wir zeigen nun die genauere Aussage im Satz durch Induktion nach $d - h$, beginnend mit $d - h = 2$. In diesem Fall ist die Aussage gerade, dass F der Durchschnitt von zwei Facetten ist, was wir gerade gezeigt haben. Es sei also $d - h > 2$. Wir betrachten dann zunächst den Fall $d - h \geq 2$. Indem wir das obige Argument anwenden, sehen wir, dass es jedenfalls zwei verschiedene echte Seiten G_1 und G_2 von P mit $F \subsetneq G_i$, $i = 1, 2$ gibt. Durch induktive Anwendung auf G_1 und G_2 sehen wir: Es gibt solche G_i mit $\dim(G_i) = d - 1$, d.h. F ist im Durchschnitt von zwei Facetten enthalten. Sei \mathcal{F}_i die Menge aller k -dimensionalen Seiten von G_i , die F enthalten. (Beachte, dass $k < \dim(G_i) = d - 1$ gilt.) Nach Induktionsvoraussetzung gelten $F = \bigcap \mathcal{F}_i$ und $|\mathcal{F}_i| \geq k - h$ für $i = 1, 2$. Da G_1 und G_2 verschieden sind, gilt außerdem $\mathcal{F}_1 \neq \mathcal{F}_2$. (Dazu muss man überlegen, dass $G_i = \text{aff}(\mathcal{F}_i)$ ist.) Da die Elemente von $\mathcal{F}_1 \cup \mathcal{F}_2$ alle k -dimensionale Seiten von P sind und $|\mathcal{F}_1 \cup \mathcal{F}_2| \geq k - h + 1$ gilt, ist die Behauptung bewiesen. Es bleibt noch der Fall $k = d - 1$: Sei dann \mathcal{G} die Menge der $d - 2$ -dimensionalen Seiten, die F enthalten. Wir haben schon gezeigt, dass $F = \bigcap \mathcal{G}$ und $|\mathcal{G}| = d - h - 1$ gelten. Nach dem Fall $h = d - 2$ oben, ist außerdem jedes $G \in \mathcal{G}$ der Durchschnitt von zwei Facetten. Somit ist auch F der Durchschnitt aller dieser Facetten. Da F die Dimension h , muss es außerdem insgesamt mindestens $d - h = k - h + 1$ solche Facetten geben. Damit ist alles gezeigt. \square

Korollar 4.3. *Es sei $P = \{x \in \mathbb{R}^d \mid \langle c_1, x \rangle \leq \beta_1, \dots, \langle c_k, x \rangle \leq \beta_k\}$ ein nicht-leeres Polyeder in \mathbb{R}^d , wobei $c_1, \dots, c_k \in \mathbb{R}^d$ und $\beta_1, \dots, \beta_k \in \mathbb{R}$. Die folgenden Aussagen sind äquivalent.*

- (1) P enthält keine Gerade.
- (2) P besitzt eine Ecke.
- (3) $\text{span}(c_1, \dots, c_k) = \mathbb{R}^d$.

Beweis. (1) \Leftrightarrow (2): gilt ganz allgemein nach Prop. 3.12; (2) \Rightarrow (3) folgt aus (3) im vorangehenden Theorem; (3) \Rightarrow (1): Es sei $L \subset P$ eine Gerade, etwa $L = \{u + \tau v \mid \tau \in \mathbb{R}\}$ für $v \in \mathbb{R}^d \setminus \{0\}$. Dann gilt also $\beta_i \geq \langle c_i, u + \tau v \rangle = \langle c_i, u \rangle + \tau \langle c_i, v \rangle$ für alle $i = 1, \dots, k$ und $\tau \in \mathbb{R}$. Daraus folgt $\langle c_i, v \rangle = 0$ für alle $i = 1, \dots, k$. Also sind c_1, \dots, c_k linear abhängig. \square

Korollar 4.4. *Jedes beschränkte Polyeder ist ein Polytop, also die konvexe Hülle von endlich vielen Punkten.*

Beweis. Das folgt sofort aus Aussage (2) im vorangehenden Satz und dem Satz von Krein-Milman 3.13. \square

Die Umkehrung (Satz von Weyl-Minkowski) beweisen wir später.

Übung 4.1. Zeige, dass ein Polytop nur endlich-viele Seiten hat. (Mit Weyl-Minkowski folgt das natürlich aus der entsprechenden Aussage für Polyeder; es soll hier aber direkt gezeigt werden.)

Übung 4.2. Zeige, dass ein Polyeder in \mathbb{R}^d , das durch k Ungleichungen beschrieben ist, höchstens $\binom{k}{d}$ Ecken hat.

Übung 4.3. Es sei $C \subset \mathbb{R}^d$ abgeschlossen und konvex. Zeige, dass C genau dann ein Polyeder ist, wenn C nur endlich viele Seiten besitzt.

4.2. DAS BIRKHOFF-POLYTOP

Zur Illustration des Hauptsatzes beweisen wir einen Satz von Birkhoff-von Neumann über stochastische Matrizen.

Definition 4.5. Eine quadratische Matrix heißt *doppelt stochastisch*, wenn sie nicht-negative Einträge hat und alle Zeilen und Spalten sich zu 1 summieren. In Symbolen: Eine $d \times d$ -Matrix $A = (a_{ij})$ ist doppelt stochastisch, falls

$$a_{ij} \geq 0, \quad \sum_{k=1}^d a_{ik} = 1 \quad \text{und} \quad \sum_{k=1}^d a_{kj} = 1$$

für alle $i, j = 1, \dots, d$ erfüllt sind.

Die Menge B_d aller doppelt-stochastischen $d \times d$ -Matrizen heißt das *Birkhoff-Polytop*; denn per Definition ist B_d ein beschränktes Polyeder und damit nach Kor. 4.4 ein Polytop.

Übung 4.4. Es sei U eine unitäre $d \times d$ -Matrix (d.h. U hat Einträge in \mathbb{C} und es gilt $U\overline{U^T} = 1$) und sei B die Matrix mit Einträgen $b_{ij} = |u_{ij}|^2$. Zeige, dass B doppelt stochastisch ist.

Ist $\sigma \in S_d$ eine Permutation, so bestimmt σ die Permutationsmatrix $X_\sigma = (x_{ij})$ mit Einträgen $x_{ij} = \delta_{i\sigma(i)}$, also $x_{i\sigma(i)} = 1$ und $x_{ij} = 0$ für alle $j \neq \sigma(i)$. Offenbar enthält B_d alle Permutationsmatrizen der Größe d .

Übung 4.5. Zeige, dass die Permutationsmatrizen der Größe $d \times d$ Extrempunkte von B_d sind.

Satz 4.6 (Birkhoff-von Neumann). *Die Ecken des Birkhoff-Polytops sind genau die Permutationsmatrizen.*

Beweis. Nach der vorangehenden Übung sind alle Permutationsmatrizen Extrempunkte von B_d . Die Umkehrung beweisen wir mit Induktion nach d . Für $d = 1$ ist nichts zu zeigen, denn $B_d = \{1\}$. Sei also $d > 1$ und sei L der Unterraum von $\text{Mat}(d \times d, \mathbb{R}) \cong \mathbb{R}^{d^2}$, der durch

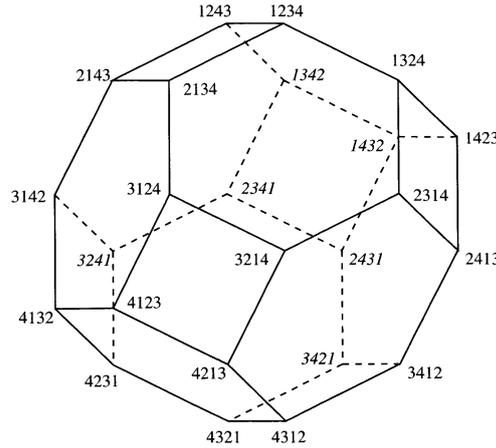


Bild des Permutations-Polytops $P(a)$ für $a = (1, 2, 3, 4)^T \in \mathbb{R}^4$.
 Quelle: Barvinok, *A course in Convexity* [1], Seite 59.

die Gleichungen der Zeilen- und Spaltensummen definiert wird, also

$$L = \left\{ X = (x_{ij}) \in \mathbb{R}^{d^2} \mid \sum_{k=1}^d x_{ik} = \sum_{k=1}^d x_{kj} = 1 \text{ für alle } i, j = 1, \dots, d \right\}.$$

Es gilt $\dim(L) = (d-1)^2$, da wir etwa die $(d-1)^2$ Einträge im linken oberen $(d-1) \times (d-1)$ -Block frei wählen und dann die Einträge in der letzten Zeile und der letzten Spalte daraus berechnen können: Sind x_{ij} für alle $i, j \leq d-1$ gegeben, so erhalten wir $x_{d j_0} = 1 - \sum_{i=1}^{d-1} x_{i j_0}$ und $x_{i_0 d} = 1 - \sum_{j=1}^{d-1} x_{i_0 j}$ für $i_0, j_0 = 1, \dots, d-1$; schließlich noch $x_{dd} = (2-d) + \sum_{i,j}^{d-1} x_{ij}$. Weiter gilt nach Definition $B_d \subset L$, und innerhalb von $L \cong \mathbb{R}^{(d-1)^2}$ ist B_d durch die d^2 Ungleichungen $x_{ij} \geq 0$ beschrieben. Ist $X = (x_{ij})$ eine Ecke von B_d , so müssen nach Thm. 4.2(3) mindestens $(d-1)^2$ dieser Ungleichungen in X aktiv sein (also $|I(X)| \geq (d-1)^2$). Das heißt, es gilt $x_{ij} = 0$ für mindestens $(d-1)^2$ verschiedene Einträge von X . Es gibt also höchstens $d^2 - (d-1)^2 = 2d-1$ von Null verschiedene Einträge in X . Es folgt, dass es eine Zeile von X geben muss, mit nur einem von 0 verschiedenen Eintrag. Sei etwa $x_{i_0 j_0} = 1$ und $x_{i_0 k} = 0$ für alle $k \neq j_0$. Wegen $\sum_{k=1}^d x_{i_0 k} = 1$, folgt $x_{i_0 j_0} = 1$. Dann folgt aus $\sum_{k=1}^d x_{k j_0} = 1$ aber auch $x_{k j_0} = 0$ für alle $k \neq i_0$. Durch Streichen der i_0 -ten Zeile und j_0 -ten Spalte erhalten wir demnach eine doppelt-stochastische $(d-1) \times (d-1)$ -Matrix X' , die ein Extrempunkt von B_{d-1} sein muss. Nach Induktionsvoraussetzung ist X' eine Permutationsmatrix und damit ist es auch X . \square

Übung 4.6. Zeige, dass B_d in L (wie im Beweis) nicht-leeres Inneres hat und damit $\dim(B_d) = (d-1)^2$ gilt.

Es sei $a \in \mathbb{R}^d$ ein Punkt. Wir betrachten die lineare Abbildung $T_a: \text{Mat}(d \times d, \mathbb{R}) \rightarrow \mathbb{R}^d$, $T_a(X) = Xa$ und setzen $P(a) = T_a(B_d)$. Dann ist $P(a)$ ein Polytop, genannt das *Permutations-Polytop*. Aus dem Satz von Birkhoff-von Neumann folgt, dass $P(a)$ die konvexe Hülle aller Punkte ist, die durch Permutation der Koordinaten von a entstehen. Offenbar ist $P(a)$ in der Hyperebene $H = \{x \in \mathbb{R}^d \mid \sum_{i=1}^d x_i = \sum_{i=1}^d a_i\}$ enthalten. Falls a nicht lauter gleiche Einträge hat, so hat $P(a)$ in H nicht-leeres Inneres, also die Dimension $d-1$. Setze zur Illustration $d=4$ und wähle $a = (1, 2, 3, 4) \in \mathbb{R}^4$. Dann ist $P(a)$ ein 3-dimensionales Bild des 9-dimensionalen Birkhoff-Polytops B_4 (siehe Abbildung).

5. TRENNUNGSSÄTZE UND KEGEL

5.1. WEITERE TRENNUNGSSÄTZE

In diesem kurzen Abschnitt diskutieren wir eine Reihe von Trennungssätzen im euklidischen Raum, die wir im wesentlichen aus dem Isolationssatz bekommen.

Satz 5.1 (Isolationssatz). *Sei $A \subset \mathbb{R}^d$ eine offene konvexe Teilmenge. Zu jedem Punkt $u \in \mathbb{R}^d \setminus A$ existiert eine Hyperebene durch u , die A strikt isoliert.*

Beweis. Siehe 2.8. □

Übung 5.1. Sei $A \subset \mathbb{R}^d$ eine konvexe Teilmenge mit $\text{int}(A) \neq \emptyset$. Zeige: Jede Hyperebene, die $\text{int}(A)$ isoliert, isoliert A .

Korollar 5.2. *Es sei $A \subset \mathbb{R}^d$ eine konvexe Teilmenge. Dann gibt es zu jedem $u \in \mathbb{R}^d \setminus A$ eine Hyperebene durch u , die A isoliert und das relative Innere von A strikt isoliert.*

Beweis. Es sei $U = \text{aff}(A)$. Falls $u \notin U$, dann können wir jede Hyperebene H mit $u \in H$ und $H \cap U = \emptyset$ nehmen. Es gelte also $u \in U$. Die Menge $\text{relint}(A)$ ist nicht-leer und offen in U (nach 2.3). Nach dem Isolationssatz und der vorangehenden Übungsaufgabe erhalten wir eine Hyperebene H' in U (also einen affinen Unterraum $H' \subset U \subset \mathbb{R}^d$ mit $\dim(H') = \dim(U) - 1$), die $\text{relint}(A)$ strikt isoliert und damit A isoliert. Wähle nun eine Hyperebene H in \mathbb{R}^d mit $H \cap U = H'$, dann hat H die gewünschte Eigenschaft. □

Korollar 5.3. *Es seien $A, B \subset \mathbb{R}^d$ zwei konvexe Teilmengen mit $A \cap B = \emptyset$. Dann gibt es eine Hyperebene, die A und B trennt.*

Beweis. Es sei $C = A - B = \{a - b \mid a \in A, b \in B\}$. Die Menge C ist konvex (siehe Aufgabe 1.4), und aus $A \cap B = \emptyset$ folgt $0 \notin C$. Nach dem vorangehenden Korollar, gibt es eine Hyperebene H durch 0 , die C isoliert. Mit anderen Worten, es gibt ein lineares Funktional $\ell: \mathbb{R}^d \rightarrow \mathbb{R}$ mit $\ell(a - b) \leq 0$ und damit $\ell(a) \leq \ell(b)$ für alle $a \in A, b \in B$. Wähle $\alpha \in \mathbb{R}$ mit $\ell(a) \leq \alpha$ und $\alpha \leq \ell(b)$ für alle $a \in A, b \in B$, dann ist $\{x \in \mathbb{R}^d \mid \ell(x) = \alpha\}$ die gesuchte Hyperebene. □

Satz 5.4. *Es sei $A \subset \mathbb{R}^d$ eine abgeschlossene konvexe Menge. Dann gibt es zu jedem $u \in \mathbb{R}^d \setminus A$ eine Hyperebene, die A und u strikt trennt.*

Beweis. Wähle $\varepsilon > 0$ mit $B_\varepsilon(u) \cap A = \emptyset$ und setze $B = B_{\varepsilon/2}(0)$. Wir behaupten, dass die Minkowski-Summen $A + B$ und $u + B$ offene disjunkte Teilmengen von \mathbb{R}^d sind: Wegen

$$A + B = \bigcup_{x \in A} B_{\varepsilon/2}(x)$$

ist $A + B$ eine Vereinigung von offenen Mengen und damit offen. Ebenso ist $u + B$, die Translation von B um u , offen. Angenommen es gäbe einen Punkt $x \in (A + B) \cap (u + B)$.

Dann wäre $\|x - u\| < \varepsilon/2$ und es gäbe es $a \in A$ mit $\|x - a\| < \varepsilon/2$. Es würde $\|a - u\| < \varepsilon$ folgen, ein Widerspruch.

Nach Kor. 5.3 gibt es also eine Hyperebene die $A + B$ und $u + B$ trennt. Da $A + B$ und $u + B$ offen sind, trennt H die beiden strikt (siehe Übung 2.4). \square

Korollar 5.5. *Es sei $A \subset \mathbb{R}^d$ eine abgeschlossene konvexe Menge. Dann gibt es zu jedem $u \in \mathbb{R}^d \setminus A$ eine Hyperebene durch u , die A strikt isoliert.*

Beweis. Nach dem vorangehenden Satz, gibt es eine Hyperebene H , die A und u strikt trennt. Sei etwa $H = \{x \in \mathbb{R}^d \mid \ell(x) = \alpha\}$ für ein lineares Funktional $\ell: \mathbb{R}^d \rightarrow \mathbb{R}$. Dann hat $H' = \{x \in \mathbb{R}^d \mid \ell(x) = \ell(u)\}$ die gewünschte Eigenschaft. \square

Man beachte, dass wir aus der Aussage des Korollars nicht genauso problemlos durch Verschieben auf die Aussage des Satzes zurückschließen können: Sei etwa $A = \{(x, y) \in \mathbb{R}^2 \mid xy \geq 1 \text{ und } x \geq 0\}$ und $u = (0, 0)$. Dann ist die x -Achse eine Gerade durch u , die A strikt isoliert. Sie lässt sich aber nicht so verschieben, dass sie A und u strikt trennt.

Übung 5.2. Gib ein Beispiel für zwei disjunkte abgeschlossene konvexe Teilmengen von \mathbb{R}^2 , die nicht strikt durch eine Gerade getrennt werden können.

Übung 5.3.

- (1) Es seien $A, B \subset \mathbb{R}^d$ zwei abgeschlossene konvexe Teilmengen mit $A \cap B = \emptyset$ und B kompakt. Zeige mit Hilfe von Satz 5.4, dass es eine Hyperebene gibt, die A und B strikt trennt.
- (2) Zeige dieselbe Aussage direkt, ohne Verwendung eines bekannten Trennungssatzes.
Hinweis: Zeige zunächst, dass es Punkte $a \in A, b \in B$ gibt derart, dass $\|a - b\| \leq \|x - y\|$ für alle $x \in A, y \in B$ gilt.

5.2. KEGEL

Eine Teilmenge C eines reellen Vektorraums V heißt ein *Kegel*, wenn C konvex ist und $\lambda x \in C$ für alle $\lambda \in \mathbb{R}_{\geq 0}$ und $x \in C$ gilt.¹ Ein Kegel heißt *stumpf*, wenn er eine Gerade enthält (nicht notwendig durch den Ursprung), andernfalls *spitz*. Ein Kegel $C \subset V$ heißt *voll* (oder *erzeugend*), wenn jedes $v \in V$ eine Darstellung $v = v_1 - v_2$ mit $v_1, v_2 \in C$ besitzt. Ein Kegel, der gleichzeitig ein Polyeder ist, heißt ein *polyedrischer Kegel*.

Beispiele 5.6.

- (1) Ein schöner, aber etwas langweiliger Kegel ist die *Eistüte* $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 \leq z^2 \text{ und } z \geq 0\}$.
- (2) Der positive Orthant $\mathbb{R}_{\geq 0}^d = \{(x_1, \dots, x_d) \in \mathbb{R}^d \mid x_i \geq 0\}$ ist ein polyedrischer Kegel.
- (3) Der Kegel der positiv semidefiniten symmetrischen $d \times d$ -Matrizen ist nicht polyedrisch. Mit diesem Kegel werden wir uns noch ausgiebig beschäftigen.
- (4) Jeder lineare Unterraum ist ein Kegel.

Übung 5.4. Genau dann ist $C \subset V$ ein Kegel, wenn $\lambda \cdot C \subset C$ für alle $\lambda \geq 0$ und $C + C \subset C$ gelten.

Übung 5.5. Sei $C \neq \emptyset$ ein Kegel in \mathbb{R}^d . Zeige:

- (1) $\text{clos}(C)$ und $\text{int}(C) \cup \{0\}$ sind ebenfalls Kegel.
- (2) Es gilt $C - C = \text{span}(C) = \text{aff}(C)$.
- (3) Genau dann ist C voll, wenn $\text{int}(C) \neq \emptyset$ gilt.

¹Konvexität ist in der Regel nicht Teil der Definition von *Kegel*, sondern nur die Abgeschlossenheit unter positiven Streckungen. Ich kann aber nicht-konvexen Kegeln nichts abgewinnen und finde es seltsam, zum Beispiel einen Haufen Ursprungsgeraden in der Ebene als Kegel zu bezeichnen.

Übung 5.6. Ist $C \subset V$ ein Kegel und $v \in C$ ein Extrempunkt von C , so ist v der Nullpunkt in V .

Lemma 5.7. Es sei $C \subset \mathbb{R}^d$ ein nicht-leerer abgeschlossener Kegel. Äquivalent sind:

- (1) C ist spitz.
- (2) Der Nullpunkt in \mathbb{R}^d ist ein Extrempunkt von C .
- (3) Es gilt $C \cap (-C) = \{0\}$.

Beweis. (1) \Leftrightarrow (2): Nach Prop. 3.12 enthält C genau dann keine Gerade, wenn C einen Extrempunkt besitzt; nach Übung 5.6 kann nur der Nullpunkt Extrempunkt sein. (2) \Rightarrow (3): Ist $x \in C \cap (-C)$, $x \neq 0$, so ist $0 = \frac{1}{2}(x + (-x))$ kein Extrempunkt. (3) \Rightarrow (2): Ist umgekehrt der Nullpunkt kein Extrempunkt, etwa $0 = \frac{1}{2}(x + y)$ mit $x, y \in C$, $x, y \neq 0$, so folgt $x = -y$ und damit $x \in C \cap (-C)$. \square

Übung 5.7. Überlege, welche Implikationen im vorangehenden Lemma auch für nicht-abgeschlossene Kegel in \mathbb{R}^d gelten und finde in den übrigen Fällen ein Gegenbeispiel.

Sei $S \subset V$ eine Teilmenge. Dann ist

$$\text{cone}(S) = \left\{ \sum_{i=1}^k \alpha_i x_i \mid x_i \in S, \alpha_i \geq 0 \text{ für alle } i = 1, \dots, k (k \in \mathbb{N}) \right\}$$

der von S erzeugte Kegel. Äquivalent ist S der Durchschnitt aller Kegel, die S enthalten. Die Elemente von $\text{cone}(S)$ heißen die *Kegelkombinationen* von Elementen aus S . Der Satz von Carathéodory überträgt sich von $\text{conv}(S)$ auf $\text{cone}(S)$, mit um 1 verbesserter Schranke:

Korollar 5.8 (zum Satz von Carathéodory). Es sei $S \subset \mathbb{R}^d$ eine Teilmenge. Dann ist jeder Punkt in $\text{cone}(S)$ eine Kegelkombination von höchstens d Punkten aus S .

Beweis. Setze $C = \text{cone}(S)$ und sei $x \in C$, etwa $x = \sum_{i=1}^k \alpha_i x_i$ mit $\alpha_i \geq 0$ und $x_i \in S$. Setze $\alpha = \sum_{i=1}^k \alpha_i$. Wir können $x \neq 0$ und damit $\alpha \neq 0$ annehmen. Dann gilt $\alpha^{-1}x = \sum_{i=1}^k \alpha^{-1}\alpha_i x_i \in \text{conv}(x_1, \dots, x_k)$. Nach Carathéodory (Satz 1.3) ist $\alpha^{-1}x$ also eine Konvexkombination von höchstens $d+1$ der x_1, \dots, x_k , nach Umnummerieren etwa $\alpha^{-1}x = \sum_{i=1}^{d+1} \beta_i x_i$ mit $\sum_{i=1}^{d+1} \beta_i = 1$ und $\beta_i \geq 0$ für $i = 1, \dots, d+1$. Sei P das Polytop $\text{conv}(x_1, \dots, x_{d+1}) \subset C$ und sei L_+ die Halbgerade $\{\lambda x \mid \lambda > 0\} \subset C$. Da P kompakt ist mit $L_+ \cap P \neq \emptyset$, L_+ aber unbeschränkt ist, gibt es $y \in L_+ \cap \partial P$, etwa $y = \mu x$ mit $\mu > 0$. Dann ist μx Konvexkombination von d der Punkte x_1, \dots, x_{d+1} , ohne Einschränkung etwa $\mu x = \sum_{i=1}^d \gamma_i x_i$, also $x = \sum_{i=1}^d \mu^{-1} \gamma_i x_i \in \text{cone}(x_1, \dots, x_d)$. \square

Bemerkung 5.9. Wie im vorangehenden Beweis überzeugt man sich leicht, dass generell $\text{cone}(S) = \bigcup_{\lambda \geq 0} (\lambda \cdot \text{conv}(S))$ gilt.

Die um Eins bessere Schranke im Satz von Carathéodory für Kegel illustriert ein generelles Phänomen: Durch die Invarianz unter Streckungen geht Kegeln gegenüber allgemeinen konvexen Mengen gewissermaßen eine Dimension verloren. Zum Beispiel sind zweidimensionale Kegel ähnlich trivial wie eindimensionale konvexe Mengen. Der Versuch, sich diese Tatsache systematisch zu Nutze zu machen, führt zum Begriff der Kegelbasis.

Definition 5.10. Es sei $C \subset V$ ein Kegel. Ein Durchschnitt von C mit einem affin-linearen Unterraum von V heißt eine *Kegelbasis*, wenn $0 \notin B$ und $\text{cone}(B) = C$ gelten.²

Übung 5.8. Zeige: Ist B eine Kegelbasis eines Kegels C , so besitzt jedes $v \in C \setminus \{0\}$ eine eindeutige Darstellung $v = \lambda w$ mit $\lambda > 0$ und $w \in B$.

²Auch hier entspricht unsere Terminologie nicht ganz der üblichen.

Beispiele 5.11. Das Segment $\{(a, 1-a) \in \mathbb{R}^2 \mid a \in [0, 1]\}$ ist eine Kegelbasis des positiven Quadranten. Dagegen besitzt die abgeschlossene obere Halbebene keine Kegelbasis.

Proposition 5.12. Es sei $C \subset \mathbb{R}^d$ ein Kegel, $C \neq \emptyset, \{0\}$.

- (1) Genau dann ist C abgeschlossen, wenn die Menge $\{x \in C \mid \|x\| = 1\}$ kompakt ist.
- (2) Genau dann ist C spitz und abgeschlossen, wenn C eine kompakte Kegelbasis besitzt.

Beweis. (1) Es sei $E = \{x \in \mathbb{R}^d \mid \|x\| = 1\}$. Wenn C abgeschlossen ist, dann ist natürlich $E \cap C$ kompakt. Sei umgekehrt $A = E \cap C$ kompakt. Offenbar gilt $C = \bigcup_{\lambda > 0} \lambda A$. Ist $u \in \mathbb{R}^d$ mit $u \notin C$, dann müssen wir zeigen, dass es $\varepsilon > 0$ mit $B_\varepsilon(u) \cap C = \emptyset$ gibt. Setze $\lambda_0 = \|u\| + 1$. Dann gilt sicher $B_1(u) \cap \lambda A = \emptyset$ für alle $\lambda > \lambda_0$. Andererseits ist $C_{\lambda_0} = \bigcup_{\lambda \in [0, \lambda_0]} \lambda A$ kompakt (denn es ist das Bild der kompakten Menge $[0, \lambda_0] \times A$ unter der stetigen Abbildung $(\lambda, x) \mapsto \lambda x$). Wegen $u \notin C_{\lambda_0}$ gibt es also $1 > \varepsilon > 0$ mit $B_\varepsilon(u) \cap C_{\lambda_0} = \emptyset$, und für dieses ε gilt dann auch $B_\varepsilon(u) \cap C = \emptyset$.

(2) Sei C spitz und abgeschlossen und sei $A = E \cap C$ wie im Beweis von (1). Dann ist $\text{conv}(A)$ eine kompakte (Kor. 1.4) konvexe Menge. Da C spitz ist, gilt $0 \notin \text{conv}(A)$. Nach Satz 5.4 gibt es eine Hyperebene H , die 0 und $\text{conv}(A)$ strikt trennt. Behaupte, dass $B = H \cap C$ eine Kegelbasis von C ist. Denn zu jedem $x \in C \setminus \{0\}$ existieren $\lambda > 0$ und $y \in A$ mit $x = \lambda y$. Andererseits liegen y und 0 auf verschiedenen Seiten von H , also gibt es $z \in H$ und $\mu > 0$ mit $y = \mu z$ und damit $x = \lambda \mu z$. Daraus folgt $\text{cone}(B) = C$. Außerdem ist B beschränkt (nämlich im Inneren der Einheitskugel enthalten) und damit kompakt.

Sei umgekehrt B eine kompakte Kegelbasis von C . Dann ist $A = \{x \in C \mid \|x\| = 1\}$ ebenfalls kompakt, denn die Abbildung $B \rightarrow A, x \mapsto (1/\|x\|)x$ ist ein Homöomorphismus (siehe Aufgabe 5.8). Also ist C abgeschlossen. Wäre C stumpf, so gäbe es nach Lemma 5.7 ein $x \in C \cap (-C)$. Nach Übung 5.8 gäbe es $y, y' \in B$ und $\lambda, \lambda' > 0$ mit $x = \lambda y, -x = \lambda' y'$. Es würde folgen $0 \in \text{conv}(y, y') \subset B$, ein Widerspruch. \square

Übung 5.9.

- (1) Gib ein Beispiel für einen stumpfen Kegel, der dennoch eine Kegelbasis besitzt.
- (2) Gib ein Beispiel einer kompakten Menge $S \subset \mathbb{R}^2$, für die $\text{cone}(S)$ nicht abgeschlossen ist.

Übung 5.10. Jede Stützhyperebene an einen Kegel C (also eine Hyperebene, die C isoliert, aber nicht strikt isoliert) enthält den Nullpunkt.

Übung 5.11. Ist $C \neq \emptyset$ ein spitzer, abgeschlossener Kegel in \mathbb{R}^d , so ist 0 ein exponierter Extrempunkt von C .

Definition 5.13. Eine eindimensionale Seite eines Kegels heißt ein *Extremalstrahl*.

Extremalstrahlen spielen für Kegel im wesentlichen die Rolle, die im allgemeinen die Extrempunkte spielen, was sich in der folgenden Aufgabe zeigt:

Übung 5.12. Es sei C ein Kegel in V . Zeige:

- (1) Jeder Extremalstrahl von C hat die Form $\text{cone}(v)$ mit $v \in C \setminus \{0\}$.
- (2) Sei $V = \mathbb{R}^d$ und C besitze eine kompakte Kegelbasis B . Dann ist $\bigcup_{\lambda > 0} (\lambda \cdot \text{ex}(B))$ genau die Vereinigung aller Extremalstrahlen von C .
- (3) Sei $V = \mathbb{R}^d$ und C besitze eine kompakte Kegelbasis. Dann hat jedes $v \in C$ eine Darstellung

$$v = \sum_{i=1}^m \lambda_i u_i \quad \text{mit } \lambda_i \geq 0, u_i \in C, m \in \mathbb{N},$$

wobei $\text{cone}(u_i)$ Extremalstrahlen von C sind. (*Krein-Milman für Kegel*)

6. POSITIV SEMIDEFINITE MATRIZEN

6.1. GRUNDLAGEN

Definite und semidefinite Matrizen sind natürlich aus der linearen Algebra bekannt. Da sie aber eine wichtige Rolle für einen großen Teil der Vorlesung spielen werden, stellen wir hier noch einmal alles zusammen, was man darüber wissen sollte.

Im Folgenden bedeutet „Matrix“ immer „reelle Matrix“.

Definition 6.1. Eine symmetrische $n \times n$ -Matrix A heißt *positiv semidefinit*, wenn $x^T A x \geq 0$ für alle $x \in \mathbb{R}^n$ gilt. Sie heißt *positiv definit*, falls $x^T A x > 0$ für alle $x \neq 0$.

Notation. Wir schreiben

- Mat_n für den n^2 -dimensionalen Vektorraum aller $n \times n$ -Matrizen.
- Sym_n für den $\frac{(n+1)n}{2}$ -dimensionalen Unterraum der symmetrischen Matrizen.
- S_n^+ bzw. S_n^{++} für die Menge der positiv semidefiniten bzw. definiten Matrizen.
- $A \geq 0$ bzw. $A > 0$ falls die Matrix A positiv semidefinit bzw. definit ist.
- $\text{Diag}(\lambda_1, \dots, \lambda_n)$ für die Diagonalmatrix mit Diagonaleinträgen $\lambda_1, \dots, \lambda_n$.
- I_n für die Einheitsmatrix $\text{Diag}(1, \dots, 1)$ in Mat_n .
- $O_n = \{U \in \text{Mat}_n \mid U^T U = I_n\}$ für die Gruppe der orthogonalen Matrizen.

Wir erinnern an die folgenden Tatsachen:

- (1) Alle Eigenwerte einer symmetrischen Matrix sind reell.¹
- (2) Der Satz über die Hauptachsentransformation sagt, dass zu einer symmetrischen Matrix eine Orthogonalbasis aus Eigenvektoren existiert. Mit anderen Worten, zu jeder symmetrischen Matrix A gibt es eine orthogonale Matrix $U \in O_n$ derart, dass $U^T A U = \text{Diag}(\lambda_1, \dots, \lambda_n)$ diagonal ist. Es folgt, dass die Matrix A genau dann positiv semidefinit ist, wenn keiner der Eigenwerte $\lambda_1, \dots, \lambda_n$ negativ ist. Sie ist genau dann positiv definit, wenn alle Eigenwerte positiv sind.
- (3) Die Spur $\text{tr}(A)$ einer Matrix A ist die Summe ihrer Diagonaleinträge. Es gelten $\text{tr}(A) = \text{tr}(A^T)$ und $\text{tr}(\lambda A) = \lambda \text{tr}(A)$ für alle $A \in \text{Mat}_n$ und $\lambda \in \mathbb{R}$. Die Spur ist nicht multiplikativ, aber für jede $n \times m$ -Matrix A und jede $m \times n$ -Matrix B gilt

$$\text{tr}(AB) = \text{tr}(BA).$$

(Beachte, dass $AB \in \text{Mat}_n$ aber $BA \in \text{Mat}_m$ ist.) Insbesondere folgt $\text{tr}(U^{-1}AU) = \text{tr}(A)$ für alle $A, U \in \text{Mat}_n$ mit U invertierbar. Es folgt, dass die Spur einer Matrix mit der Summe aller (komplexen) Eigenwerte übereinstimmt.

¹Der folgende kurze Beweis wirkt ein bißchen wie ein Zaubertrick: Ist A reell und symmetrisch und $\lambda \in \mathbb{C}$ ein Eigenwert von A mit zugehörigem komplexem Eigenvektor $x \neq 0$, dann gilt $\lambda \|x\|^2 = \lambda x^T \bar{x} = (Ax)^T \bar{x} = x^T A \bar{x} = x^T \overline{Ax} = \overline{\lambda x^T \bar{x}} = \overline{\lambda} \|x\|^2$, also $\lambda = \overline{\lambda}$.

- (4) Identifiziert man Mat_n mit \mathbb{R}^{n^2} , dann erhält man das kanonische (euklidische) Skalarprodukt $\langle A, B \rangle = \sum_{i,j=1}^n a_{ij}b_{ij}$. Sind A und B symmetrisch, so kann man das auch mit Hilfe der Spur hinschreiben:

$$\langle A, B \rangle = \text{tr}(AB) = \text{tr}(BA).$$

Denn es gilt $(AB)_{ij} = \sum_{k=1}^n a_{ik}b_{kj} = \sum_{k=1}^n a_{ki}b_{kj}$, also $\text{tr}(AB) = \sum_{i=1}^n (AB)_{ii} = \sum_{i,k=1}^n a_{ik}b_{ik} = \langle A, B \rangle$.

- (5) Auf den symmetrischen Matrizen operiert die orthogonale Gruppe durch Konjugation $O_n \times \text{Sym}_n \ni (U, A) \mapsto U^{-1}AU$; wegen $U^{-1} = U^T$ ist dabei $U^{-1}AU$ wieder symmetrisch. Das Skalarprodukt ist invariant unter dieser Operation, d.h. für jede orthogonale $n \times n$ -Matrix U gilt

$$\langle U^{-1}AU, U^{-1}BU \rangle = \langle A, B \rangle,$$

wegen $\langle U^{-1}AU, U^{-1}BU \rangle = \text{tr}(U^{-1}AUU^{-1}BU) = \text{tr}(U^{-1}(AB)U) = \text{tr}(AB) = \langle A, B \rangle$.

Proposition 6.2. *Es sei A eine symmetrische $n \times n$ -Matrix vom Rang r .*

- (1) *Genau dann ist A positiv semidefinit, wenn es eine $n \times r$ -Matrix B mit $A = BB^T$ gibt.*
- (2) *Genau dann ist A positiv semidefinit, wenn es $P \in S_n^+$ mit $A = P^2$ gibt.²*

Beweis. Wegen $x^T BB^T x = (B^T x)^T B^T x = \|B^T x\|^2 \geq 0$ für alle $x \in \mathbb{R}^n$, ist jede Matrix der Form BB^T positiv semidefinit. Umgekehrt gibt es nach Hauptachsentransformation eine orthogonale Matrix U mit $U^T AU = \text{Diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$, wobei $\lambda_1, \dots, \lambda_r > 0$. Setze $D = \text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}, 0, \dots, 0)$. Dann gilt $A = UDDU^T = (UD)(UD)^T$. Dabei sind in der Matrix UD alle Einträge in den hinteren $n - r$ Spalten 0. Man erhält B mit der gewünschten Eigenschaft, indem man diese $n - r$ Nullspalten aus UD streicht. Das beweist (1). Für (2) setze $P = UDU^T$. Dann gilt $P \in S_n^+$ und $P^2 = UDU^T UDU = UD^2U^T = A$. \square

Übung 6.1.

- (1) Für jedes $x \in \mathbb{R}^n \setminus \{0\}$ ist die symmetrische $n \times n$ -Matrix xx^T positiv semidefinit vom Rang 1.
- (2) Sei umgekehrt A eine positiv semidefinite Matrix. Zeige: Der Rang von A ist genau die kleinste Zahl r derart, dass A eine Darstellung $A = \sum_{i=1}^r x_i x_i^T$ mit $x_1, \dots, x_r \in \mathbb{R}^n$ besitzt.

Übung 6.2.

- (1) Zeige: Sind $A, B, C \in \text{Sym}_n$, so gilt $\text{tr}(ABC) = \text{tr}(ACB)$.
- (2) Finde ein Beispiel für $A, B, C \in \text{Mat}_2$ mit $\text{tr}(ABC) \neq \text{tr}(ACB)$.

Übung 6.3. Es sei $A \in S_n^+$ und $x \in \mathbb{R}^n$. Zeige: Aus $x^T Ax = 0$ folgt $Ax = 0$.

Übung 6.4.

- (1) Zeige, dass eine symmetrische 2×2 -Matrix $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ genau dann positiv semidefinit ist, wenn $a, c \geq 0$ und $ac \geq b^2$ gelten.
- (2) Für jede positiv semidefinite $n \times n$ -Matrix $A = (a_{ij})$ gilt $a_{ii}a_{jj} \geq a_{ij}^2$ für alle $i, j = 1, \dots, n$.

²Man kann zeigen, dass P durch A eindeutig bestimmt ist. Es ist üblich, P als die Quadratwurzel von A zu bezeichnen und \sqrt{A} zu schreiben.

6.2. DER PSD-KEGEL

Wir nennen S_n^+ kurz den *psd-Kegel* und geben in diesem Abschnitt eine explizite Beschreibung seiner Seitenstruktur, in ähnlicher Weise wie bei Polyedern.

Proposition 6.3.

- (1) Der psd-Kegel S_n^+ ist spitz und abgeschlossen.
- (2) Es gilt $\text{int}(S_n^+) = S_n^{++}$ innerhalb von Sym_n .

Beweis. Es ist klar, dass S_n^+ abgeschlossen unter Addition und unter Multiplikation mit nicht-negativen Skalaren ist, also überhaupt ein Kegel ist. Ist $A \neq 0$ positiv semidefinit, dann ist $-A$ nicht positiv semidefinit, also gilt $S_n^+ \cap (-S_n^+) = \{0\}$. Betrachte die Abbildung

$$\mu: \begin{cases} \text{Sym}_n & \rightarrow & \mathbb{R} \\ A & \mapsto & \min\{x^T A x \mid x \in \mathbb{R}^n \text{ mit } \|x\| = 1\}. \end{cases}$$

Die Abbildung μ ist stetig, und es gilt $S_n^+ = \mu^{-1}([0, \infty))$. Also ist S_n^+ abgeschlossen. Ferner gilt $S_n^{++} = \mu^{-1}((0, \infty))$, also ist S_n^{++} offen. Damit gilt $\partial S_n^+ \subset S_n^+ \setminus S_n^{++}$. Es bleibt die umgekehrte Inklusion zu zeigen. Sei also $A \in S_n^+ \setminus S_n^{++}$, eine positiv semidefinite aber nicht definite Matrix. Sei $x \neq 0$ ein Eigenvektor von A zum Eigenwert 0 und setze $B = xx^T$. Für jedes $\varepsilon > 0$ ist dann $A - \varepsilon B$ nicht positiv semidefinit und damit $A \in \partial S_n^+$, denn es gilt

$$x^T (A - \varepsilon B)x = x^T A x - \varepsilon x^T x x^T x = -\varepsilon \|x\|^4 < 0. \quad \square$$

Satz 6.4 (Seitenstruktur des psd-Kegels). *Für jeden linearen Unterraum L von \mathbb{R}^n sei*

$$F_L = \{X \in S_n^+ \mid L \subset \ker(X)\}.$$

- (1) Für jedes $A \in S_n^+$ ist $F_{\ker(A)}$ die eindeutig bestimmte Seite von S_n^+ , die A in ihrem relativen Inneren $\{X \in S_n^+ \mid \ker(X) = \ker(A)\}$ enthält.
- (2) Alle Seiten von S_n^+ sind exponiert.
- (3) Für jeden Unterraum L hat F_L die Dimension $\frac{r(r+1)}{2}$ mit $r = n - \dim(L)$, und es existiert eine orthogonale Matrix $U \in O_n$ derart, dass

$$U^T F_L U = \left\{ \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} \in \text{Sym}_n \mid B \in S_r^+ \right\}.$$

- (4) Die Zuordnung $L \mapsto F_L$ ist eine Bijektion zwischen den linearen Unterräumen von \mathbb{R}^n und den Seiten von S_n^+ . Dabei gilt für alle Unterräume L_1 und L_2 von \mathbb{R}^n :
 - (a) $F_{L_1+L_2} = F_{L_1} \cap F_{L_2}$
 - (b) $F_{L_1} \subset F_{L_2}$ genau dann, wenn $L_2 \subset L_1$.
 - (c) $F_{L_1 \cap L_2}$ ist die kleinste Seite von S_n^+ , die F_{L_1} und F_{L_2} enthält.

Die Eigenschaften (a)-(c) sagen gerade, dass $L \mapsto F_L$ ein Anti-Isomorphismus des Unterraumverbands von \mathbb{R}^n mit dem Seitenverband von S_n^+ ist.³

Beweis. Zunächst wählen wir die Koordinaten günstig, so dass wir alles explizit überprüfen können. Es sei $0 \leq r \leq n$, und wir setzen $L = \{x \in \mathbb{R}^n \mid x_1 = \dots = x_r = 0\}$. Dann ist

$$F_L = \left\{ \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} \in \text{Sym}_n \mid B \in S_r^+ \right\}$$

die Menge aus (3). Sei nun $r < n$ und setze $C = \text{Diag}(0, \dots, 0, 1, \dots, 1) \neq 0$, wobei die ersten r Einträge 0 und die letzten $n - r$ Einträge 1 sind. Dann ist $H = \{X \in \text{Sym}_n \mid \langle C, X \rangle = 0\}$

³Ein Verband ist eine algebraische Struktur mit zwei Verknüpfungen \wedge und \vee („meet“ und „join“), die in etwa dieselben formalen Eigenschaften wie Vereinigung und Durchschnitt haben. Ein Anti-Isomorphismus ist eine Bijektion, die \wedge und \vee respektiert, aber miteinander vertauscht.

eine Hyperebene. Dabei ist $\langle C, X \rangle = \sum_{i=n-r+1}^n X_{ii} \geq 0$ für alle $X \in S_n^+$, so dass S_n^+ durch H isoliert wird. Behaupte, dass

$$F_L = H \cap S_n^+$$

gilt. Für $X \in F_L$ gilt schon $CX = 0$ und damit erst recht $\langle C, X \rangle = \text{tr}(CX) = 0$. Umgekehrt, sei $X \in H \cap S_n^+$. Wir schreiben X als Blockmatrix aus:

$$X = \begin{pmatrix} B_1 & B_2 \\ B_2 & B_3 \end{pmatrix}, \text{ wobei } B_1 \in \text{Mat}_r, B_2 \in \text{Mat}_{r \times (n-r)}, B_3 \in \text{Mat}_{n-r}.$$

Dann ist B_3 psd mit $\text{tr}(B_3) = \langle C, X \rangle = 0$, also ist $B_3 = 0$. Ist nun $i \neq j$ mit $i > r$, dann folgt aus $X_{ii}X_{jj} - X_{ij}^2 \geq 0$ (nach Aufgabe 6.4) wegen $X_{ii} = 0$ auch $X_{ij} = 0$. Also folgt $B_2 = 0$ und $X \in F_L$. Insgesamt haben wir gezeigt, dass F_L eine exponierte Seite von S_n^+ ist.

(1) Es sei nun $A \in S_n^+$ und $r = \text{rk}(A)$. Falls $r = n$, dann ist A positiv definit und $F_{\ker(A)} = F_{\{0\}} = S_n^+$ ist die eindeutig bestimmte Seite, die A in ihrem Inneren enthält. Es gelte also $0 \leq r < n$. Seien $\lambda_1, \dots, \lambda_r > 0$ die Eigenwerte ungleich Null von A , sei $D = \text{Diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$ und wähle $U \in O_n$ mit $U^T A U = D$ (Hauptachsentransformation). Die Abbildung $X \mapsto U^T X U$ ist eine lineare Isometrie des Vektorraums Sym_n auf sich selbst (siehe Bemerkung (5) oben) und es gilt $U^T S_n^+ U = S_n^+$. Dabei liegt $U^T A U = D$ im relativen Inneren von F_L (denn die ersten r Einträge von D sind strikt positiv). Also ist $U F_L U^T = S_n^+ \cap (U H U^T)$ eine exponierte Seite, die A in ihrem relativen Inneren enthält. Nach Prop. 3.8 gibt es höchstens eine solche Seite. Außerdem gilt $L = \ker(D) = \ker(U^T A U) = U^T \ker(A)$ und damit $U F_L U^T = F_{\ker(A)}$.

(2) Sei F eine nicht-leere Seite von S_n^+ . Dann gilt $\text{relint}(F) \neq \emptyset$ (nach Prop. 2.3) und für $A \in \text{relint}(F)$ haben wir in (1) gezeigt, dass $F = F_{\ker(A)}$ exponiert ist.

(3) Zu jedem linearen Unterraum L von \mathbb{R}^n existiert $A \in S_n^+$ mit $\ker(A) = L$. Wir haben im Beweis von (1) gesehen, dass dann $U \in O_n$ mit $U^T A U = \text{Diag}(\lambda_1, \dots, \lambda_{\text{rk}(A)}, 0, \dots, 0)$ die gewünschte Eigenschaft hat.

(4) Wir haben in (2) gezeigt, dass die Abbildung $L \mapsto F_L$ surjektiv ist. Sind $L_1 \neq L_2$ Unterräume, etwa $L_1 \not\subset L_2$, dann folgt $\text{relint}(L_2) \cap F_{L_1} = \emptyset$, also $F_{L_1} \neq F_{L_2}$. Für $A \in \text{Mat}_n$ gilt außerdem $(L_1 \cup L_2) \subset \ker(A)$ genau dann, wenn $L_1 + L_2 \subset \ker(A)$. Daraus folgt (a). Die Eigenschaft (b) ist klar aus der Definition und (c) folgt aus (b) und der Tatsache, dass $L_1 \cap L_2$ der größte Unterraum ist, der in L_1 und L_2 enthalten ist. \square

Übung 6.5. Zeige, dass zu jedem linearen Unterraum L von \mathbb{R}^n ein $A \in S_n^+$ mit $\ker(A) = L$ existiert (was oben im Beweis von (3) verwendet wurde).

Übung 6.6. Zeige: Die Extremalstrahlen von S_n^+ sind genau die Kegel $\text{cone}(xx^T)$ für $x \in \mathbb{R}^n \setminus \{0\}$.

Definition 6.5. Sei $C \subset \mathbb{R}^d$ ein Kegel. Dann ist

$$C^* = \{x \in \mathbb{R}^d \mid \text{Für alle } y \in C: \langle x, y \rangle \geq 0\}$$

ein Kegel, der *duale Kegel* von C . Gilt $C^* = C$, so heißt C *selbstdual*.

Übung 6.7. Zeige, dass der positive Orthant in \mathbb{R}^d selbstdual ist.

Proposition 6.6. Der psd-Kegel S_n^+ ist selbstdual.

Beweis. Seien $A, B \in S_n^+$. Nach Prop. 6.2 gibt es $P, Q \in \text{Mat}_d$ mit $A = PP^T, B = QQ^T$. Dann folgt $\langle A, B \rangle = \langle PP^T, QQ^T \rangle = \text{tr}(PP^TQQ^T) = \text{tr}(Q^T P P^T Q) = \text{tr}(Q^T P (Q^T P)^T) \geq 0$. Ist umgekehrt $A \in \text{Sym}_n$ mit $\langle A, X \rangle \geq 0$ für alle $X \in S_n^+$, so folgt $x^T A x = \text{tr}(x^T A x) = \text{tr}(A x x^T) = \langle A, x x^T \rangle \geq 0$, für alle $x \in \mathbb{R}^d$. \square

7. POSITIV SEMIDEFINITE MATRIZEN II

7.1. SPEKTRAEDER UND LINEARE GLEICHUNGEN IN MATRIZEN

Definition 7.1. Ein Schnitt des psd-Kegels mit einem affin-linearen Unterraum von Sym_n heißt ein *Spektraeder*.¹

Der Name suggeriert schon, dass man sich Spektraeder in gewisser Hinsicht als Verallgemeinerungen von Polyedern vorstellen soll. Wir werden uns später noch mit der Geometrie von Spektraedern beschäftigen. Aber erst einmal setzen wir das Studium des psd-Kegels fort, und das hübsche Wort Spektraeder ist dabei nur ein Name.

Definition 7.2. Sei V ein endlich-dimensionaler \mathbb{R} -Vektorraum und $L \subset V$ ein affin-linearer Unterraum. Dann heißt $\text{codim}_V(L) = \dim(V) - \dim(L)$ die *Codimension* von L in V .

Übung 7.1. Es sei V ein \mathbb{R} -Vektorraum, $W \subset V$ ein linearer und $L \subset V$ ein affin-linearer Unterraum. Dann gilt $\text{codim}_W(L \cap W) \leq \text{codim}_V(L)$.

Proposition 7.3. *Es sei $S = L \cap S_n^+$ ein nicht-leeres Spektraeder. Sei $r \geq 0$ eine ganze Zahl mit $\text{codim}_{\text{Sym}_n}(L) < \frac{(r+2)(r+1)}{2}$. Dann enthält S eine Matrix vom Rang höchstens r . Genauer gilt $\text{rk}(X) \leq r$ für alle $X \in \text{ex}(S)$.*

Beweis. Da S_n^+ spitz ist, enthält S keine Gerade. Damit gilt $\text{ex}(S) \neq \emptyset$ nach Prop. 3.12. Sei also $X \in \text{ex}(S)$ und setze $m = \text{rk}(X)$. Die eindeutig bestimmte Seite F von S_n^+ , die X in ihrem relativen Innern enthält, hat nach Satz 6.4 die Dimension $\frac{(m+1)m}{2}$. Außerdem ist X auch ein relativ innerer Punkt von $L \cap F \subset S$. Da X andererseits ein Extrempunkt von S ist, folgt $L \cap \text{aff}(F) = \{X\}$ und damit $\dim(L \cap F) = 0$. Das geht nur, wenn $\dim(L) + \dim(F) \leq \dim(\text{Sym}_n)$ ist², also $\frac{(m+1)m}{2} = \dim(F) \leq \text{codim}(L) < \frac{(r+1)(r+2)}{2}$ und damit $m < r + 1$. \square

Übung 7.2. Seien $A_1, \dots, A_k \in \text{Sym}_n$ und $\alpha_1, \dots, \alpha_k \in \mathbb{R}$. Zeige: Wenn es eine Matrix $X \in S^+$ mit

$$\langle A_i, X \rangle = \alpha_i \text{ für } i = 1, \dots, k$$

gibt, dann gibt es eine solche Matrix mit

$$\text{rk}(X) \leq \lfloor \frac{\sqrt{8k+1}-1}{2} \rfloor.$$

Unser Hauptziel für diesen Abschnitt ist die folgende Präzisierung im kompakten Fall.

Satz 7.4. *Es sei $S = S_n^+ \cap L$ ein kompaktes nicht-leeres Spektraeder. Es gebe eine ganze Zahl $r > 0$ mit $\text{codim}_{\text{Sym}_n}(L) \leq \frac{(r+2)(r+1)}{2}$ und $n \geq r + 2$. Dann gibt es $X \in S$ mit $\text{rk}(X) \leq r$.*

¹Diese Definition wird in Abschnitt 13 korrigiert und leicht erweitert: Ein Spektraeder ist dort das Urbild eines psd-Kegels unter einer affin-linearen Abbildung.

²Die Dimensionsformel $\dim(L + W) = \dim(L) + \dim(W) - \dim(L \cap W)$ stimmt auch für affin-lineare Unterräume L und W , sobald $L \cap W \neq \emptyset$. Denn in diesem Fall kann man $v \in L \cap W$ nehmen und zu den linearen Räumen $L - v$, $W - v$ übergehen.

Für den Beweis brauchen wir ein wenig algebraische Topologie. Es sei $m \geq 0$ und $\mathbb{S}^m = \{x \in \mathbb{R}^{m+1} \mid \|x\| = 1\}$ die Einheitssphäre in \mathbb{R}^{m+1} . Wir schreiben \mathbb{P}^m für die Menge aller Ursprungsgeraden in \mathbb{R}^{m+1} (genannt der m -dimensionale reelle projektive Raum). Jede Ursprungsgerade in \mathbb{R}^{m+1} schneidet \mathbb{S}^m in genau zwei Punkten, deshalb hat jedes Element von \mathbb{P}^m genau zwei Repräsentanten in \mathbb{S}^m . Man versieht \mathbb{P}^m mit der *Quotiententopologie*, die durch die euklidische Topologie auf \mathbb{S}^m induziert wird.³ Der projektive Raum spielt eine wichtige Rolle in der Geometrie, auch wenn er zunächst nicht unbedingt anschaulich ist. (Zum Beispiel ist \mathbb{P}^2 eine nicht-orientierbare Fläche, die keine Einbettung in \mathbb{R}^3 besitzt.) Wir brauchen aber nur den folgenden Fakt aus der algebraischen Topologie.

Satz 7.5. *Es sei $m \geq 2$. Dann gibt es keine stetige, injektive Abbildung $\mathbb{S}^m \rightarrow \mathbb{P}^m$.*

Beweis. Die Aussage folgt aus klassischen Sätzen der algebraischen Topologie. Siehe zum Beispiel Hatcher [4], Kap. 2, insbesondere Kor. 2.B4. \square

Beweis von Satz 7.4. Falls $\text{codim}(L) < \frac{(r+2)(r+1)}{2}$, dann folgt die Aussage sofort aus Prop. 7.3. Es gelte also $\text{codim}(L) = \frac{(r+2)(r+1)}{2}$.

1. *Schritt: Reduktion auf den Fall $n = r + 2$.* Nach Prop. 7.3 gibt es $Y \in S$ derart, dass $\text{rk}(Y) \leq r+1$, also $\dim(\ker(Y)) \geq n-r-1$. Sei $R \subset \ker(Y)$ ein Unterraum mit $\dim(R) = n-r-2$ und \mathcal{F}_R die zugehörige Seite von S_n^+ . Nach Satz 6.4 wird \mathcal{F}_R durch Konjugation mit einer orthogonalen Matrix in den psd-Kegel S_{r+2}^+ überführt. Es genügt, die Aussage für $\mathcal{F}_R \cap L$ zu beweisen. (Beachte dabei, dass $\text{codim}_W(L \cap W) \leq \text{codim}_{\text{Sym}_n}(L)$ mit $W = \text{span}(\mathcal{F}_R)$ gilt, nach Aufgabe 7.1. Die Voraussetzung an die Codimension bleibt also erfüllt.)

2. *Schritt: Reduktion auf den Fall $S \cap S_{r+2}^{++} \neq \emptyset$.* Falls $S \cap S_{r+2}^{++} = \emptyset$, dann ist S in einer echten exponierten Seite $F = W \cap S_{r+2}^+$ von S_{r+2}^+ enthalten. Dann ist $\text{codim}_W(L) < \text{codim}_{\text{Sym}_{r+2}^+}(L)$, und da F in W zu $S_{\dim(W)}^+$ isomorph ist, folgt die Behauptung aus Prop. 7.3.

3. *Schritt: Falls es $X, Y \in S$ mit $\text{rk}(X) = \text{rk}(Y) = r + 1$ und $\ker(X) = \ker(Y)$ gibt, dann folgt die Behauptung.* Denn sei $W = \ker(X) = \ker(Y)$. Dann ist \mathcal{F}_W wie in Satz 6.4 eine Seite von S_{r+2}^+ , die X und Y in ihrem relativen Innern enthält. Dabei besteht $\text{relint}(\mathcal{F}_W)$ nach 6.4(3) aus psd-Matrizen vom Rang $r + 1$. Sei G die Gerade $\{\lambda X + (1 - \lambda)Y \mid \lambda \in \mathbb{R}\}$. Dann gilt $G \subset L$, und da S kompakt ist, gibt es eine Matrix X_0 in $G \cap \partial \mathcal{F}_W \subset S$. Für diese gilt $\text{rk}(X_0) < r + 1$.

4. *Schritt: Es gibt solche Matrizen X, Y .* Wir behaupten, dass es eine Matrix $X_0 \in \partial_L S$ gibt mit $\text{rk}(X_0) \leq r$. Nach Prop. 6.3 enthält $\partial_L S$ sicher keine Matrizen von vollem Rang, d.h. es gilt $\text{rk}(X) \leq r + 1$ für alle $X \in \partial_L S$. Angenommen, die Behauptung ist falsch. Dann folgt also $\text{rk}(X) = r + 1$ und damit $\dim(\ker(X)) = 1$ für alle $X \in \partial_L S$.

Nach Voraussetzung gilt $\text{codim}(L) = \frac{(r+2)(r+1)}{2}$, also $\dim(L) = \frac{(r+3)(r+2)}{2} - \frac{(r+2)(r+1)}{2} = r + 2$. Wir definieren eine stetige Abbildung $\varphi: \mathbb{S}^{r+1} \rightarrow \mathbb{P}^{r+1}$ wie folgt: Es sei $Z \in S \cap S_{r+2}^{++}$ positiv definit und sei A_1, \dots, A_{r+2} eine Basis des linearen Unterraums $L - Z$. Für jedes $y \in \mathbb{S}^{r+1}$, sei G_y die Halbgerade $\{Z + \lambda(y_1 A_1 + \dots + y_{r+2} A_{r+2}) \mid \lambda > 0\}$. Da S kompakt ist, besteht $G_y \cap \partial_L S$ aus einer eindeutigen Matrix $X(y)$ (die Eindeutigkeit gilt nach Aufgabe 2.1(1)). Nun zeigt man, dass $X(y)$ und $\ker(X(y))$ stetig von y abhängen. Wir setzen

$$\varphi: \begin{cases} \mathbb{S}^{r+1} & \rightarrow & \mathbb{P}^{r+1} \\ y & \mapsto & \ker(X(y)). \end{cases}$$

Nach Satz 7.5 kann φ nicht injektiv sein, d.h. es gibt $y_1, y_2 \in \mathbb{S}^{r+1}$ mit $\varphi(y_1) = \varphi(y_2)$. Dann sind $X = X(y_1)$ und $Y = X(y_2)$ die gesuchten Matrizen. \square

³Genauer kann man das so sagen: Betrachte die Abbildung $\pi: \mathbb{S}^m \rightarrow \mathbb{P}^m$ gegeben durch $\pi(x) = \text{span}(x)$. Dann versehen wir \mathbb{P}^m mit der feinsten Topologie derart, dass π stetig ist.

Übung 7.3. Zeige durch geeignet gewählte Gegenbeispiele, dass in Satz 7.4 keine der Voraussetzungen S kompakt, $r > 0$ und $n \geq r + 2$ überflüssig ist.

7.2. ANWENDUNG: QUADRATISCHE KONVEXITÄT

Lineare Bilder konvexer Mengen sind konvex. Die Frage, wann *quadratische* Bilder konvexer Mengen wieder konvex sind, ist dagegen nicht so leicht zu beantworten. Wir beginnen mit einer Folgerung aus Prop. 7.3.

Korollar 7.6. Es seien $A_1, A_2 \in \text{Sym}_n$ und sei $q_i: \mathbb{R}^n \rightarrow \mathbb{R}$ die quadratische Form $x \mapsto x^T A_i x$ ($i = 1, 2$). Für $\alpha_1, \alpha_2 \in \mathbb{R}$ haben die beiden quadratischen Gleichungen

$$q_1(x) = \alpha_1 \quad \text{und} \quad q_2(x) = \alpha_2$$

in x genau dann eine gemeinsame Lösung in \mathbb{R}^n , wenn die beiden linearen Gleichungen

$$\langle A_1, X \rangle = \alpha_1 \quad \text{und} \quad \langle A_2, X \rangle = \alpha_2$$

in X eine gemeinsame positiv semidefinite Lösung haben.

Beweis. Sei x eine gemeinsame Lösung der beiden quadratischen Gleichungen und setze $X = xx^T$. Dann ist $X \in S_n^+$ und es gilt $\langle A_i, X \rangle = \text{tr}(A_i xx^T) = \text{tr}(x^T A_i x) = \alpha_i$ für $i = 1, 2$.

Sei umgekehrt $X \in S_n^+$ eine gemeinsame Lösung der beiden linearen Gleichungen. Dann sagt Satz 7.3, dass es solches X mit $\text{rk}(X) \leq 1$ gibt. Nach Übung 6.1 gibt es $x \in \mathbb{R}^n$ mit $X = xx^T$. Es folgt $x^T A_i x = \text{tr}(x^T A_i x) = \text{tr}(A_i xx^T) = \text{tr}(A_i X) = \langle A_i, X \rangle = \alpha_i$. \square

Übung 7.4. Beweise die folgende Verschärfung: Genau dann haben die beiden quadratischen Gleichungen in Kor. 7.6 eine Lösung x mit $\|x\| \leq 1$, wenn die beiden linearen Gleichungen eine positiv semidefinite Lösung X mit $\text{tr}(X) \leq 1$ haben.

Übung 7.5. Es sei $C \subset \mathbb{R}^n$ ein Kegel und $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^d$ eine lineare Abbildung. Zeige, dass das Bild $\psi(C)$ ein Kegel in \mathbb{R}^d ist.

Korollar 7.7 (Dines 1941). Es seien $A_1, A_2 \in \text{Sym}_n$ und sei $q_i: \mathbb{R}^n \rightarrow \mathbb{R}$ die quadratische Form $x \mapsto x^T A_i x$ ($i = 1, 2$). Setze

$$Q: \begin{cases} \mathbb{R}^n & \rightarrow & \mathbb{R}^2 \\ x & \mapsto & (q_1(x), q_2(x)). \end{cases}$$

Dann ist das Bild $Q(\mathbb{R}^n)$ ein konvexer Kegel in \mathbb{R}^2 .

Beweis. Sei $\psi: \text{Sym}_n \rightarrow \mathbb{R}^2$ die lineare Abbildung gegeben durch $\psi(X) = (\langle A_1, X \rangle, \langle A_2, X \rangle)$. Dann ist $\psi(S_n^+)$ nach der vorangehenden Aufgabe ein Kegel in \mathbb{R}^2 , und nach Kor. 7.6 gilt $Q(\mathbb{R}^n) = \psi(S_n^+)$. \square

Übung 7.6. Finde drei quadratische Formen q_1, q_2, q_3 in zwei Variablen ($n = 2$) derart, dass das Bild von \mathbb{R}^2 unter der Abbildung $Q: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $Q(x) = (q_1(x), q_2(x), q_3(x))$ nicht konvex ist.

Das haben wir ziemlich direkt aus der „trivialen“ Rangabschätzung in Prop. 7.3 erhalten. Aber die ganze Mühe mit dem Beweis von Satz 7.4 soll sich natürlich auch lohnen:

Satz 7.8. Es sei $n > 2$ und $S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$. Seien $A_1, A_2 \in \text{Sym}_n$ und sei $q_i: \mathbb{R}^n \rightarrow \mathbb{R}$ die quadratische Form $x \mapsto x^T A_i x$ ($i = 1, 2$). Setze

$$Q: \begin{cases} \mathbb{R}^n & \rightarrow & \mathbb{R}^2 \\ x & \mapsto & (q_1(x), q_2(x)). \end{cases}$$

Dann ist das Bild $Q(S^{n-1})$ in \mathbb{R}^2 kompakt und konvex.

Beweis. Es sei B das Spektraeder $\{X \in S_n^+ \mid \text{tr}(X) = 1\}$, und wir setzen

$$\psi: \begin{cases} \text{Sym}_n & \rightarrow & \mathbb{R}^2 \\ X & \mapsto & (\langle A_1, X \rangle, \langle A_2, X \rangle). \end{cases}$$

Wir behaupten, dass $Q(S^{n-1}) = \psi(B)$ gilt und damit konvex ist (denn ψ ist linear). Für die eine Inklusion, sei $(\alpha_1, \alpha_2) \in Q(S^{n-1})$ und sei $x \in S^{n-1}$ mit $q_i(x) = \alpha_i$ ($i = 1, 2$). Setze $X = xx^T$, dann gilt $\text{tr}(X) = \text{tr}(xx^T) = \text{tr}(x^T x) = 1$, also $X \in B$, außerdem $\langle A_i, X \rangle = \text{tr}(A_i xx^T) = \alpha_i$ und somit $\psi(X) = (\alpha_1, \alpha_2)$.

Für die umgekehrte Inklusion, sei $(\alpha_1, \alpha_2) \in \psi(B)$. Es sei S das Spektraeder

$$\{X \in S_n^+ \mid \langle A_1, X \rangle = \alpha_1, \langle A_2, X \rangle = \alpha_2 \text{ und } \text{tr}(X) = 1\}$$

Dann ist S kompakt und nicht-leer (warum?). Wir wenden Satz 7.4 mit $r = 1$ an und erhalten die Existenz von $X \in S$ mit $\text{rk}(X) \leq 1$. Schreibe $X = xx^T$ mit $x \in \mathbb{R}^n$, dann folgt $q_i(x) = \alpha_i$, wie zuvor. \square

Übung 7.7. Zeige, dass die analoge Aussage zu Satz 7.8 für $n = 2$ nicht gilt.

Übung 7.8. Beweise die folgende Verbesserung von Kor. 7.7:

Es seien $A_1, A_2 \in \text{Sym}_n$ und sei $q_i: \mathbb{R}^n \rightarrow \mathbb{R}$ die quadratische Form $x \mapsto x^T A_i x$ ($i = 1, 2$). Setze

$$Q: \begin{cases} \mathbb{R}^n & \rightarrow & \mathbb{R}^2 \\ x & \mapsto & (q_1(x), q_2(x)) \end{cases}$$

und sei $B = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$. Dann ist $Q(B)$ eine kompakte und konvexe Teilmenge von \mathbb{R}^2 .

Definition 7.9. Es sei $A \in \text{Mat}_n(\mathbb{C})$ eine komplexe $n \times n$ -Matrix. Die Menge

$$R(A) = \{z^T A \bar{z} \mid z \in \mathbb{C}^n \text{ mit } z^T \bar{z} = 1\} \subset \mathbb{C}$$

heißt der *numerische Wertebereich* von A .

Korollar 7.10 (Satz von Toeplitz-Hausdorff). *Der numerische Wertebereich einer komplexen Matrix ist eine konvexe Teilmenge der komplexen Ebene.*

Beweis. Für $n = 1$ ist der numerische Wertebereich ein Punkt. Sei $n \geq 1$. Wir identifizieren \mathbb{C}^n mit \mathbb{R}^{2n} , indem wir $z = (z_1, \dots, z_n)^T \in \mathbb{C}^n$ auf $(\xi, \eta) \in \mathbb{R}^{2n}$ mit $z_j = \xi_j + i\eta_j$ abbilden. Unter dieser Identifikation ist die Menge aller $z \in \mathbb{C}^n$ mit $z^T \bar{z} = 1$ gerade die Einheitskugel

$$S^{2n-1} = \{(\xi, \eta) \in \mathbb{R}^{2n} \mid \xi^T \xi + \eta^T \eta = 1\}.$$

Nun ist $R(A)$ für jedes $A \in \text{Mat}_n(\mathbb{C})$ das Bild von S^{2n-1} unter einer quadratischen Abbildung $S^{2n-1} \rightarrow \mathbb{R}^2$. \square

Übung 7.9. Gib die quadratische Abbildung im vorangehenden Beweis explizit an.

8. DUALITÄT

8.1. DER BIDUALSATZ

Der Begriff *Dualität* kommt in vielen mathematischen Theorien vor. Es ist kaum möglich, sie alle auf einen gemeinsamen Nenner zu bringen und zu erklären, was Dualität schlechthin ist. Die Dualität in der Konvexität ist zunächst relativ nah an der der linearen Algebra. In 6.5 haben wir bereits den dualen Kegel definiert. Dies verallgemeinert sich auf beliebige konvexe Mengen wie folgt:

Definition 8.1. Es sei $A \subset \mathbb{R}^d$ eine Teilmenge. Die Menge

$$A^\circ = \{c \in \mathbb{R}^d \mid \langle c, x \rangle \leq 1 \text{ für alle } x \in A\}$$

heißt das *Dual* von A .¹

Proposition 8.2. Seien $A, B, A_i \subset \mathbb{R}^d$ Teilmengen ($i \in I$).

- (1) A° ist eine abgeschlossene konvexe Teilmenge, die den Ursprung enthält.
- (2) Es gilt $A \subset (A^\circ)^\circ$.
- (3) $(\mathbb{R}^d)^\circ = \{0\}$ und $\{0\}^\circ = \mathbb{R}^d$.
- (4) Wenn $A \subset B$, dann $B^\circ \subset A^\circ$.
- (5) Es gilt $(\bigcup_{i \in I} A_i)^\circ = \bigcap_{i \in I} A_i^\circ$.
- (6) Für alle $\alpha \in \mathbb{R}$ gilt $(\alpha A)^\circ = \alpha^{-1} A^\circ$.
- (7) Ist $C \subset \mathbb{R}^d$ ein Kegel, so ist das Dual bis auf Vorzeichen der duale Kegel (Def. 6.5):

$$-C^\circ = C^* = \{c \in \mathbb{R}^d \mid \langle c, x \rangle \geq 0\}.$$

- (8) Ist $L \subset \mathbb{R}^d$ ein linearer Unterraum, so gilt $L^\circ = L^\perp$.
- (9) Für $x_1, \dots, x_m \in \mathbb{R}^d$ gilt

$$(\text{conv}(x_1, \dots, x_m))^\circ = \{c \in \mathbb{R}^d \mid \langle c, x_i \rangle \leq 1 \text{ für } i = 1, \dots, m\}.$$

Beweis. (1) Für jedes feste $x \in \mathbb{R}^d$ ist der Halbraum $\{c \in \mathbb{R}^d \mid \langle c, x \rangle \leq 1\}$ abgeschlossen und konvex. Da A° nach Definition der Durchschnitt solcher Halbräume ist, ist auch A° konvex und abgeschlossen.

(7) Für $c \in C^*$ gilt $\langle -c, x \rangle \leq 0 < 1$ für alle $x \in C$. Ist umgekehrt $c \in C^\circ$ und $x \in C$ mit $0 < \langle c, x \rangle \leq 1$, so gibt es $\alpha \in \mathbb{R}$ mit $\langle c, \alpha x \rangle = \alpha \langle c, x \rangle > 1$, andererseits $\alpha x \in C$, ein Widerspruch. Also gilt $\langle c, x \rangle \leq 0$ und damit $-c \in C^*$.

(9) Die eine Inklusion ist klar. Für die andere, sei $c \in \mathbb{R}^d$ mit $\langle c, x_i \rangle \leq 1$ für $i = 1, \dots, m$ und sei $x = \sum_{i=1}^d \lambda_i x_i \in \text{conv}(x_1, \dots, x_m)$. Dann gilt $\langle c, x \rangle = \sum_{i=1}^d \lambda_i \langle c, x_i \rangle \leq 1$. \square

Übung 8.1. Beweise die restlichen Aussagen.

¹Im Buch von Barvinok wird A° die *Polare* genannt. Dafür spricht, dass das Wort *dual* in der Mathematik reichlich überbelegt ist. Andererseits erscheint die Trennung in Polarität und Dualität etwas künstlich.

Beispiele 8.3. Es sei

$$P = \{y \in \mathbb{R}^2 \mid y_1, y_2 \in [-1, 1]\} = \text{conv}((1, 1), (-1, 1), (1, -1), (-1, -1))$$

ein Quadrat. Dann gilt nach Prop. 8.2(9)

$$P^\circ = \{x \in \mathbb{R}^2 \mid |x_1| + |x_2| \leq 1\} = \text{conv}((1, 0), (-1, 0), (0, 1), (0, -1)),$$

ein gedrehtes, in P einbeschriebenes Quadrat.

Übung 8.2. Wie sehen $(P + (1/2, 0))^\circ$ und $(P + (1, 0))^\circ$ aus?

Übung 8.3. Zeige, dass $(P^\circ)^\circ = P$ gilt.

Satz 8.4 (Bidualsatz). *Es sei $A \subset \mathbb{R}^d$ eine abgeschlossene konvexe Teilmenge, die den Ursprung enthält. Dann gilt*

$$(A^\circ)^\circ = A.$$

Beweis. Die Inklusion $A \subset (A^\circ)^\circ$ war Prop. 8.2(2). Sei umgekehrt $u \in \mathbb{R}^d \setminus A$, dann müssen wir $u \notin (A^\circ)^\circ$ zeigen. Da A konvex und abgeschlossen ist, gibt es nach Satz 5.4 eine Hyperebene, die u und A strikt trennt, d.h. es gibt $c \in \mathbb{R}^d$ und $\alpha \in \mathbb{R}$ mit $\langle c, x \rangle < \alpha$ für alle $x \in A$ und $\langle c, u \rangle > \alpha$. Wegen $0 \in A$ folgt $\alpha > 0$. Setze $b = \alpha^{-1}c$. Dann gilt $\langle b, x \rangle < 1$ für alle $x \in A$, also $b \in A^\circ$. Andererseits ist $\langle b, u \rangle > 1$ und damit $u \notin (A^\circ)^\circ$. \square

Übung 8.4. Es sei $A \subset \mathbb{R}^d$ nicht leer. Dann gilt $(A^\circ)^\circ = \text{clos}(\text{conv}(A \cup \{0\}))$.

Übung 8.5.

- (1) Sei $B = \overline{B}_1(0)$ die abgeschlossene Einheitskugel. Zeige, dass $B = B^\circ$ gilt.
- (2) Ist $A \subset \mathbb{R}^d$ mit $A = A^\circ$, so folgt $A = \overline{B}_1(0)$.

Übung 8.6. Seien $C_1, C_2 \subset \mathbb{R}^d$ zwei Kegel. Zeige

- (1) Es gilt $(C_1 + C_2)^* = C_1^* \cap C_2^*$.
- (2) Sind C_1 und C_2 abgeschlossen, so gilt

$$(C_1 \cap C_2)^* = \text{clos}(C_1^* + C_2^*).$$

Bemerkung 8.5. In unserer Definition von Dualität lebt A° im selben Raum wie A . Das hat so seine Vor- und Nachteile. Eine abstraktere Definition geht so: Sei V ein \mathbb{R} -Vektorraum und $A \subset V$. Definiere $A^\vee = \{\ell \in V^\vee \mid \ell(x) \leq 1 \text{ für alle } x \in A\}$. Ist V ein endlich-dimensionaler euklidischer Raum, also ausgestattet mit einem (positiv definiten) Skalarprodukt $\langle -, - \rangle$, dann induziert jedes $v \in V$ ein lineares Funktional $\ell_v: V \rightarrow \mathbb{R}$ gegeben durch $\ell_v(x) = \langle v, x \rangle$. Damit erhält man einen Homomorphismus $\varphi: V \rightarrow V^\vee$ gegeben durch $\varphi(v) = \ell_v$. Da das Skalarprodukt nicht-ausgeartet ist, ist φ injektiv und damit auch surjektiv, weil V und V^\vee die gleiche endliche Dimension haben. Es gilt dann offensichtlich $\varphi(A^\circ) = A^\vee$.

Für unendlich-dimensionale Räume ist die Korrespondenz zwischen V und seinem Dualraum komplizierter. Andererseits ist jeder endlich-dimensionale euklidische Raum isomorph zu \mathbb{R}^d mit dem Standard-Skalarprodukt, so dass man durch die Abstraktion im endlich-dimensionalen Fall nicht viel gewinnt.

8.2. DUALITÄT VON POLYTOPEN

Das Folgende haben wir schon im Zusammenhang mit Aufgabe 4.1 diskutiert.

Proposition 8.6. *Es sei $S \subset \mathbb{R}^d$ eine endliche Teilmenge und P das Polytop $\text{conv}(S)$. Jede Seite von P hat die Form $\text{conv}(S')$ für $S' \subset S$. Insbesondere hat P nur endlich-viele Seiten.*

Beweis. Sei F eine Seite von P . Da F kompakt ist (Kor. 3.11), gilt $F = \text{conv}(\text{ex}(F))$ nach dem Satz von Krein-Milman 3.13. Andererseits gilt $\text{ex}(F) \subset \text{ex}(P) \subset S$. \square

Proposition 8.7. *Ist $P \subset \mathbb{R}^d$ ein Polytop mit $0 \in \text{int}(P)$, so ist auch P° ein Polytop mit $0 \in \text{int}(P^\circ)$.*

Beweis. Sei $P = \text{conv}(x_1, \dots, x_m)$ ein Polytop mit $0 \in \text{int}(P)$. Es gibt also $\varepsilon > 0$ mit $\overline{B}_\varepsilon(0) \subset P$. Nach Übung 8.5 und Prop. 8.2(6) gilt $(\overline{B}_\varepsilon(0))^\circ = (\varepsilon \overline{B}_1(0))^\circ = \varepsilon^{-1}(\overline{B}_1(0))^\circ = \overline{B}_{\varepsilon^{-1}}(0)$ und nach Prop. 8.2(4) damit $P^\circ \subset \overline{B}_{\varepsilon^{-1}}(0)$. Also ist P° kompakt. Nach Prop. 8.2(9) gilt außerdem

$$P^\circ = \{c \in \mathbb{R}^d \mid \langle c, x_i \rangle \leq 1 \text{ für } i = 1, \dots, m\}.$$

Damit ist P° ein kompaktes Polyeder und somit ein Polytop nach Kor. 4.4. \square

Korollar 8.8 (Satz von Weyl-Minkowski). *Eine Teilmenge von \mathbb{R}^d ist genau dann ein Polytop, wenn sie ein kompaktes Polyeder ist.*

Beweis. Wir haben schon gezeigt, dass jedes kompakte Polyeder ein Polytop ist (Kor. 4.4). Sei umgekehrt $P = \text{conv}(x_1, \dots, x_m)$ ein Polytop. Wir können ohne Einschränkung annehmen, dass $\text{int}(P) \neq \emptyset$ (sonst ersetzen wir, wie immer, \mathbb{R}^d durch $\text{aff}(P)$) und $0 \in \text{int}(P)$ gelten (sonst verschieben wir P). Nach Prop. 8.7 ist P° wieder ein Polytop, etwa $P^\circ = \text{conv}(u_1, \dots, u_k)$ für $u_1, \dots, u_k \in \mathbb{R}^d$. Nach dem Bidualsatz und Prop. 8.2(9) ist

$$P = (P^\circ)^\circ = \{x \in \mathbb{R}^d \mid \langle x, u_i \rangle \leq 1 \text{ für } i = 1, \dots, k\}$$

ein Polyeder. \square

Übung 8.7. Zeige, dass der Durchschnitt zweier Polytope ein Polytop ist.

Satz 8.9. *Es sei $P \subset \mathbb{R}^d$ ein Polytop, das den Ursprung in seinem Inneren enthält. Für jede Seite F von P , sei*

$$\widehat{F} = \{x \in P^\circ \mid \langle x, y \rangle = 1 \text{ für alle } y \in F\}.$$

(1) *Für jede Seite F von P ist \widehat{F} eine Seite von P° und es gilt*

$$\dim(F) + \dim(\widehat{F}) = d - 1.$$

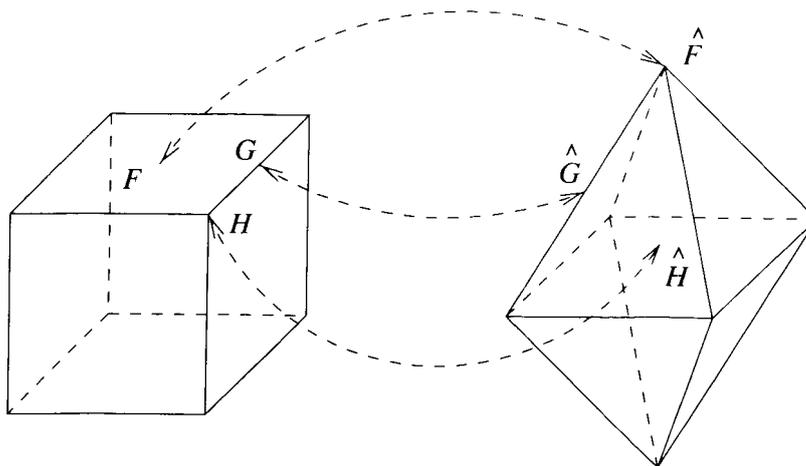
(2) *Sind $F \subset G$ Seiten von P , so folgt $\widehat{G} \subset \widehat{F}$.*

(3) *Es gilt $\widehat{\widehat{F}} = F$.*

Beweis. (1) Sei $P = \text{conv}(y_1, \dots, y_m)$ und sei F eine Seite von P . Für $F = \emptyset$ und $F = P$ ist die Behauptung trivial, wenn man die Konvention $\dim(\emptyset) = -1$ beachtet. Sei F also echt und seien $x_1, x_2 \in P^\circ$ derart, dass $1 = \langle \frac{1}{2}(x_1 + x_2), y \rangle = \frac{1}{2}(\langle x_1, y \rangle + \langle x_2, y \rangle)$ für alle $y \in F$. Wegen $\langle x_i, y \rangle \leq 1$ für $i = 1, 2$ gilt dann bereits $\langle x_1, y \rangle = \langle x_2, y \rangle = 1$ und damit $x_1, x_2 \in \widehat{F}$. Also ist \widehat{F} eine Seite.

Um die Dimensionsaussage zu zeigen, sei $c' \in \mathbb{R}^d$ und $\alpha \in \mathbb{R}$ mit $F = P \cap \{y \in \mathbb{R}^d \mid \langle c', y \rangle = \alpha\}$ und $\langle c', y \rangle < \alpha$ für alle $y \in P \setminus F$. Aus $0 \in \text{int}(P)$ folgt $\alpha > 0$, und wir setzen $c = \frac{1}{\alpha}c'$, so dass $c \in \widehat{F}$. Wir behaupten, dass

$$\text{span}(\widehat{F} - c) = (\text{span}(F))^\perp$$



Dualität zwischen den Seiten des Würfels und des Oktaeders in \mathbb{R}^3 .

Quelle: Barvinok, *A course in Convexity* [1], Seite 252.

gilt. Daraus wird die Behauptung folgen: Denn wegen $0 \notin F$ gilt $\dim(\text{span}(F)) = \dim(F) + 1$ und damit $\dim(\widehat{F}) = \dim(\text{span}(\widehat{F} - c)) = d - \dim(\text{span}(F)) = d - \dim(F) - 1$. Die Inklusion \subset ist klar. Für die Umkehrung, sei $y \in (\text{span}(F))^\perp$. Es gilt $F = \text{conv}(y_i \mid i \in I)$ mit $I = \{i \in \{1, \dots, m\} \mid y_i \in F\}$ nach Prop. 8.6. Ferner gilt

$$P^\circ = \{x \in \mathbb{R}^d \mid \langle x, y_i \rangle \leq 1 \text{ für alle } i = 1, \dots, m\}.$$

Deshalb gilt $c + \varepsilon y \in \widehat{F}$ für $\varepsilon > 0$ hinreichend klein. Denn für $i \notin I$ gilt $\langle c + \varepsilon y, y_i \rangle = \langle c, y_i \rangle + \varepsilon \langle y, y_i \rangle < 1 + \varepsilon \langle y, y_i \rangle$. Wir können also $\varepsilon > 0$ so wählen, dass $\langle c + \varepsilon y, y_i \rangle < 1$ für alle $i \notin I$ gilt. Für $i \in I$ gilt $\langle c + \varepsilon y, y_i \rangle = \langle c, y_i \rangle = 1$ wegen $c \in \widehat{F}$. Also gilt $x = c + \varepsilon y \in \widehat{F}$, wie behauptet, und damit $y = \frac{1}{\varepsilon}(x - c) \in \text{span}(\widehat{F} - c)$.

(2) ist klar.

(3) Nach (1) ist $\widehat{(\widehat{F})}$ eine Seite von $(P^\circ)^\circ = P$. Dabei ist $F \subset \widehat{(\widehat{F})}$ klar nach Definition. Es sei umgekehrt $z \in \widehat{(\widehat{F})}$. Wie in (2) können wir F mit einem Element aus \widehat{F} exponieren, d.h. wir wählen $c \in \widehat{F}$ mit $F = P \cap \{y \in \mathbb{R}^d \mid \langle c, y \rangle = 1\}$ und $\langle c, y \rangle < 1$ für alle $y \in P \setminus F$. Nach Definition gilt aber $\langle c, z \rangle = 1$ und damit $z \in F$. \square

Beispiel 8.10. Es sei $P = \{y \in \mathbb{R}^3 \mid y_i \in [-1, 1] \text{ für } i = 1, 2, 3\}$ ein Würfel. Dann gilt $P = \text{conv}((\pm 1, \pm 1, \pm 1))$. Damit ist

$$P^\circ = \{x \in \mathbb{R}^3 \mid |x_1| + |x_2| + |x_3| \leq 1\}$$

ein Oktaeder. Die Korrespondenz zwischen den Seiten von P und P° ist im Bild dargestellt.

Übung 8.8. Ein Polytop heißt *selbstdual*, wenn P durch eine lineare Transformation aus P° hervorgeht. Zeige:

- (1) Seien $x_1, \dots, x_d \in \mathbb{R}^d$ linear unabhängig und $x_{d+1} = -\sum_{i=1}^d x_i$. Dann ist das Polytop $\text{conv}(x_1, \dots, x_{d+1})$ selbstdual.
- (2) Zeige, dass der Hyperwürfel $[0, 1]^d$ für $d = 1, 2$ selbstdual ist, für $d \geq 3$ dagegen nicht.

9. KEGELPROGRAMMIERUNG

Unter dem Stichwort *Programmierung* laufen verschiedene konvexe Optimierungsprobleme. Im allgemeinen geht es darum, das Minimum (oder Maximum) eines linearen Funktionals auf einer konvexen Menge zu bestimmen. Wenn die konvexe Menge ein Polyeder ist, dann handelt es sich um die klassische *lineare Programmierung*. Ist sie ein Spektraeder, so spricht man von *semidefiniten Programmierung*. Diese beiden Programme sind wichtige Bausteine für die Optimierung, weil man sie einerseits theoretisch gut versteht und andererseits auch über effiziente numerische Lösungsverfahren verfügt. Wir werden in diesem Abschnitt zunächst beide gemeinsam von einem abstrakteren Standpunkt behandeln, der allgemeinen Kegelprogrammierung.

Notation 9.1. Sei V ein \mathbb{R} -Vektorraum und $K \subset V$ ein Kegel. Wir schreiben

$$x \succcurlyeq_K y \text{ genau dann wenn } x - y \in K.$$

Proposition 9.2. Für alle $x, y, z, w \in V$ und $\alpha \in \mathbb{R}$ gelten:

- (1) $x \succcurlyeq_K x$. (Reflexivität)
- (2) Wenn $x \succcurlyeq_K y$ und $y \succcurlyeq_K z$, dann $x \succcurlyeq_K z$. (Transitivität)
- (3) Wenn $x \succcurlyeq_K w$ und $y \succcurlyeq_K z$, dann $x + y \succcurlyeq_K w + z$.
- (4) Wenn $x \succcurlyeq_K y$ und $\alpha \geq 0$, dann $\alpha x \succcurlyeq_K \alpha y$.

Beweis. Die Eigenschaften folgen alle direkt aus der Definition eines Kegels. □

Beispiele 9.3.

- (1) Für den positiven Orthanten $K = \mathbb{R}_+^d$ ist $x \succcurlyeq_K y$ für $x, y \in \mathbb{R}^d$ genau dann, wenn $x_i \geq y_i$ für alle $i = 1, \dots, d$.
- (2) Es gilt $x \succcurlyeq_V y$ für alle $x, y \in V$, und $x \succcurlyeq_{\{0\}} y$ ist äquivalent zu $x = y$.

Proposition 9.4. Es seien V_1 und V_2 zwei endlich-dimensionale euklidische Vektorräume, mit Skalarprodukten $\langle -, - \rangle_1$ und $\langle -, - \rangle_2$. Zu jeder linearen Abbildung $\varphi: V_1 \rightarrow V_2$ gibt es eine eindeutig bestimmte lineare Abbildung $\varphi^*: V_2 \rightarrow V_1$ derart, dass

$$\langle \varphi(x), y \rangle_2 = \langle x, \varphi^*(y) \rangle_1$$

für alle $x \in V_1, y \in V_2$ gilt.

Beweis. Es sei $n = \dim(V_1)$ und x_1, \dots, x_n eine Orthonormalbasis von V_1 . Dann gilt $x = \sum_{i=1}^n \langle x_i, x \rangle_1 x_i$ für alle $x \in V_1$. Für $y \in V_2$ definieren wir daher $\varphi^*(y) = \sum_{i=1}^n \langle \varphi(x_i), y \rangle_2 x_i$. Nun rechnet man nach, dass φ^* linear ist und die behaupteten Eigenschaften hat. □

Übung 9.1. Wird φ in der vorangehenden Proposition bezüglich Orthonormalbasen auf V_1 und V_2 durch eine $m \times n$ -Matrix A dargestellt, so wird φ^* durch die transponierte Matrix A^T dargestellt.

Definition 9.5. Seien V_1, V_2 endlich-dimensionale euklidische Vektorräume und $K_1 \subset V_1, K_2 \subset V_2$ zwei Kegel, $\varphi: V_1 \rightarrow V_2$ eine lineare Abbildung. Fixiere $b \in V_1$ und $b^* \in V_2$. Es sei

$$\beta = \inf \{ \langle x, b \rangle_1 \mid x \in K_1 \text{ und } \varphi(x) - b^* \in K_2 \}.$$

Ein *Kegelprogramm* ist das Problem, β aus den Daten $(V_1, V_2, K_1, K_2, \varphi, b, b^*)$ zu bestimmen. Wie üblich kann β auch $-\infty$ oder ∞ sein. Das lineare Funktional $x \mapsto \langle x, b \rangle$ heißt die *Zielfunktion* (häufig auch *Kostenfunktion*). Die Elemente von

$$\{ x \in V_1 \mid x \in K_1 \text{ und } \varphi(x) - b^* \in K_2 \}$$

heißen *zulässige Punkte* des Kegelprogramms. Jeder zulässige Punkt x , in dem das Infimum angenommen wird, also $\beta = \langle x, b \rangle$, heißt ein *optimaler Punkt*.

Notation 9.6. Die übliche Notation für ein Kegelprogramm $(V_1, V_2, K_1, K_2, \varphi, b, b^*)$ ist

$$(P) \quad \left. \begin{array}{l} \text{Finde} \quad \beta = \inf \langle x, b \rangle_1 \\ \text{unter} \quad \varphi(x) \geq_{K_2} b^* \\ \quad \quad x \geq_{K_1} 0 \end{array} \right\} \text{ in der Variablen } x \in V_1.$$

Definition 9.7. Gegeben ein Kegelprogramm (P) wie oben. Das *duale Programm* ist

$$(D) \quad \left. \begin{array}{l} \text{Finde} \quad \beta^* = \sup \langle b^*, y \rangle_2 \\ \text{unter} \quad \varphi^*(y) \leq_{K_1^*} b \\ \quad \quad y \geq_{K_2^*} 0 \end{array} \right\} \text{ in der Variablen } y \in V_2.$$

Das ursprüngliche Programm (P) heißt das *primäre Programm*.

Übung 9.2. Gib der folgenden Aussage einen präzisen Sinn: Falls K_1 und K_2 abgeschlossen sind, so ist (P) bis auf Vorzeichen das duale Programm zu (D).

Satz 9.8. Gegeben sei ein Kegelprogramm (P) wie oben. Sei x ein zulässiger Punkt des primären Programms (P) und y ein zulässiger Punkt des dualen Programms (D).

(1) (Schwache Dualität) Es gilt

$$\langle x, b \rangle_1 \geq \langle b^*, y \rangle_2$$

und damit insbesondere

$$\beta \geq \beta^*.$$

Ist $\langle x, b \rangle_1 = \langle b^*, y \rangle_2$, dann ist x ein optimaler Punkt des primären und y ein optimaler Punkt des dualen Programms.

(2) (Optimalitätskriterium) Falls

$$\langle x, b - \varphi^*(y) \rangle_1 = 0 \quad \text{und} \quad \langle \varphi(x) - b^*, y \rangle_2 = 0$$

gelten, dann ist x ein optimaler Punkt des primären und y ein optimaler Punkt des dualen Programms, d.h. es gelten $\langle x, b \rangle = \beta$ und $\langle b^*, y \rangle = \beta^*$.

(3) (Komplementärer Schlupf¹) Ist umgekehrt x ein optimaler Punkt des primären und y ein optimaler Punkt des dualen Programms und ist außerdem $\beta = \beta^*$, dann gilt

$$\langle x, b - \varphi^*(y) \rangle_1 = 0 \quad \text{und} \quad \langle \varphi(x) - b^*, y \rangle_2 = 0.$$

Beweis. (1) Nach Voraussetzung gelten $x \in K_1$ und $b - \varphi^*(y) \in K_1^*$ und damit $\langle x, b \rangle_1 \geq \langle x, \varphi^*(y) \rangle_1$. Nach Definition von φ^* gilt weiter $\langle x, \varphi^*(y) \rangle_1 = \langle \varphi(x), y \rangle_2$. Wegen $y \in K_2$ und $\varphi(x) - b^* \in K_2^*$ folgt $\langle \varphi(x), y \rangle_2 \geq \langle b^*, y \rangle_2$. Also insgesamt $\langle x, b \rangle_1 \geq \langle b^*, y \rangle_2$, wie behauptet. Die zusätzlichen Aussagen sind daraus ebenfalls klar.

¹Eine eher holprige, aber anscheinend übliche, Übersetzung für *complementary slackness*.

(2) Aus $\langle x, b - \varphi^*(y) \rangle_1 = 0$ und $\langle \varphi(x) - b^*, y \rangle_2 = 0$ folgt

$$\langle x, b \rangle_1 = \langle x, \varphi^*(y) \rangle_1 = \langle \varphi(x), y \rangle_2 = \langle b^*, y \rangle_2$$

und damit die Optimalität von x und y nach (1).

(3) Im Beweis von (1) haben wir die Ungleichungskette

$$\langle x, b \rangle_1 \geq \langle x, \varphi^*(y) \rangle_1 = \langle \varphi(x), y \rangle_2 \geq \langle b^*, y \rangle_2$$

gesehen. Sind also die äußeren beiden Terme gleich, dann sind auch die beiden Ungleichungen Gleichheiten, was genau die Behauptung ist. \square

Obwohl der Beweis dieses Satzes sehr einfach war, ist er für die praktische Behandlung von Kegelprogrammen sehr nützlich und zeigt die Bedeutung der Dualität: *Jeder* zulässige Punkt des dualen Programms liefert eine untere Schranke für den optimalen Wert des primären Programms und umgekehrt (nach (1)) und man kann obendrein noch messen, wie weit man noch von der richtigen Antwort entfernt ist (nach (2)).

Falls $\beta > -\infty$ bzw. $\beta^* < \infty$, so heißt das Programm (P) bzw. (D) *beschränkt*, andernfalls *unbeschränkt*.

Korollar 9.9. *Wenn das primäre (bzw. das duale) Programm unbeschränkt ist, dann besitzt das duale (bzw. das primäre) Programm keinen zulässigen Punkt.* \square

Definition 9.10. Die Differenz $\beta - \beta^*$ heißt die *Dualitätslücke*. Ist sie gleich Null, so sagt man, das Programm habe *keine Dualitätslücke*.

Definition 9.11. Ein Kegelprogramm (P) wie oben mit $K_2 = \{0\}$ heißt ein *Kegelprogramm in kanonischer Form*.

Ein Kegelprogramm in kanonischer Form und sein Dual sehen dann also so aus:

Primäres Programm in kanonischer Form:

$$(P) \quad \left. \begin{array}{l} \text{Finde } \beta = \inf \langle x, b \rangle_1 \\ \text{unter } \varphi(x) = b^* \\ \quad \quad x \geq_K 0 \end{array} \right\} \text{ in der Variablen } x \in V_1.$$

Duales Programm in kanonischer Form:

$$(D) \quad \left. \begin{array}{l} \text{Finde } \beta^* = \sup \langle b^*, y \rangle_2 \\ \text{unter } \varphi^*(y) \leq_{K^*} b \end{array} \right\} \text{ in der Variablen } y \in V_2.$$

Übung 9.3. Jedes Kegelprogramm lässt sich in ein Kegelprogramm in kanonischer Form übersetzen (was die Bezeichnung rechtfertigt). Sei (P) ein allgemeines Kegelprogramm wie in Notation 9.6. Setze $V = V_1 \oplus V_2$ und $K = K_1 \oplus K_2 \subset V$ und definiere das Skalarprodukt auf V durch

$$\langle (x_1, y_1), (x_2, y_2) \rangle = \langle x_1, x_2 \rangle_1 + \langle y_1, y_2 \rangle_2.$$

Setze $\hat{b} = (b, 0) \in V$ und definiere $\hat{\varphi}: V \rightarrow V_2$ durch $\hat{\varphi}(x, y) = \varphi(x) - y$ für $x \in V_1, y \in V_2$.

Zeige, dass (P) äquivalent zu folgendem Problem in kanonischer Form ist:

$$\left. \begin{array}{l} \text{Finde } \beta = \inf \langle x, \hat{b} \rangle \\ \text{unter } \hat{\varphi}(x) = b^* \\ \quad \quad x \geq_K 0 \end{array} \right\} \text{ in der Variablen } x \in V.$$

Satz 9.12 (Kriterium für starke Dualität). *Es sei (P) ein Kegelprogramm in kanonischer Form wie oben. Angenommen der Kegel*

$$\widehat{\varphi}(K) = \{(\varphi(x), \langle x, b \rangle_1) \mid x \in K\}$$

ist abgeschlossen in $V_2 \oplus \mathbb{R}$ und (P) besitzt einen zulässigen Punkt. Dann gelten:

- (1) *Falls (P) beschränkt ist, so besitzt (P) einen optimalen Punkt.*
- (2) *Das Programm (P) hat keine Dualitätslücke.*

Beweis. (1) Es sei $L \subset V_2 \oplus \mathbb{R}$ die Gerade $L = \{(b^*, \lambda) \mid \lambda \in \mathbb{R}\}$. Der Schnitt $L \cap \widehat{\varphi}(K)$ besteht aus Punkten der Form $(b^*, \langle x, b \rangle_1)$, wobei x ein zulässiger Punkt ist, und ist damit nach Voraussetzung nicht leer. Da (P) beschränkt ist und $\widehat{\varphi}(K)$ abgeschlossen, gibt es einen zulässigen Punkt x mit $\langle x, b \rangle = \beta$, d.h. x ist ein optimaler Punkt.

(2) Falls (P) unbeschränkt ist, also $\beta = -\infty$, so besitzt (D) keinen zulässigen Punkt (Kor. 9.9), so dass auch $\beta^* = -\infty$ gilt. Wir können also annehmen, dass (P) beschränkt ist. Nach schwacher Dualität (Satz 9.8(1)) gilt $\beta \geq \beta^*$. Wir zeigen, dass für jedes $\varepsilon > 0$ ein zulässiger Punkt y von (D) existiert derart, dass $\langle b^*, y \rangle_2 \geq \beta - \varepsilon$. Daraus wird $\beta^* \geq \beta - \varepsilon$ folgen und damit $\beta = \beta^*$. Wegen $\beta - \varepsilon < \beta$ gilt

$$(b^*, \beta - \varepsilon) \notin \widehat{\varphi}(K).$$

Da $\widehat{\varphi}(K)$ abgeschlossen ist, gibt es nach Trennungssatz 5.4 eine Hyperebene, die $(b^*, \beta - \varepsilon)$ und $\widehat{\varphi}(K)$ strikt trennt. Definiert man das Skalarprodukt auf $V_2 \oplus \mathbb{R}$ durch $\langle (v_2, \alpha), (v'_2, \alpha') \rangle = \langle v_2, v'_2 \rangle_2 + \alpha\alpha'$, so bedeutet dies: Es gibt $(z, \rho) \in V_2 \oplus \mathbb{R}$ und $\alpha \in \mathbb{R}$ derart, dass

$$\begin{aligned} \langle b^*, z \rangle_2 + \rho(\beta - \varepsilon) &> \alpha \text{ und} \\ \langle \varphi(x), z \rangle_2 + \rho\langle x, b \rangle_1 &< \alpha \text{ für alle } x \in K. \end{aligned}$$

Wegen $0 \in K$ ist $\alpha > 0$. Da K ein Kegel ist, muss außerdem

$$\langle \varphi(x), z \rangle_2 + \rho\langle x, b \rangle_1 \leq 0 \text{ für alle } x \in K$$

gelten. Wir haben also

$$\begin{aligned} \langle b^*, z \rangle_2 + \rho(\beta - \varepsilon) &> 0 \text{ und} \\ \langle \varphi(x), z \rangle_2 + \rho\langle x, b \rangle_1 &\leq 0 \text{ für alle } x \in K. \end{aligned}$$

Nach (1) besitzt (P) einen optimalen Punkt x_0 , also $\langle x_0, b \rangle = \beta$ und $\varphi(x_0) = b^*$ und damit

$$\langle b^*, z \rangle_2 + \rho\beta \leq 0$$

aus der unteren Gleichung. Zusammen mit der oberen folgt $\rho < 0$. Setze $y = -\rho^{-1}z$, dann

$$\begin{aligned} \langle b^*, y \rangle_2 - (\beta - \varepsilon) &= -\rho^{-1}(\langle b^*, z \rangle_2 + \rho(\beta - \varepsilon)) > 0 \quad \text{und} \\ \langle x, \varphi^*(y) - b \rangle_1 &= \langle x, \varphi^*(y) \rangle_1 - \langle x, b \rangle_1 = \langle \varphi(x), y \rangle_2 - \langle x, b \rangle_1 \\ &= -\rho^{-1}(\langle \varphi(x), z \rangle_2 + \rho\langle x, b \rangle_1) \leq 0 \end{aligned}$$

für alle $x \in K$. Dies impliziert $b - \varphi^*(y) \in K^*$, also $\varphi^*(y) \leq_{K^*} b$. Also ist y ein zulässiger Punkt von (D) und $\langle b^*, y \rangle_2 > \beta - \varepsilon$, wie gewünscht. \square

Die Bedingung im Satz, dass $\widehat{\varphi}(K)$ abgeschlossen sei, ist nicht automatisch erfüllt, wenn K abgeschlossen ist:

Übung 9.4. Finde ein Beispiel eines abgeschlossenen spitzen Kegels in \mathbb{R}^3 derart, dass die Projektion $\pi: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ auf die ersten beiden Koordinaten nicht abgeschlossen ist.

10. LINEARE PROGRAMMIERUNG

In diesem Teil wenden wir die allgemeinen Aussagen über Kegelprogrammierung aus der vorigen Woche auf die klassische lineare Programmierung an. Als Vorbereitung brauchen wir erst noch ein wichtiges Resultat über Polyeder:

Satz 10.1 (Fourier-Motzkin-Elimination). Sei $d \geq 1$, P ein Polyeder in \mathbb{R}^d und $\pi: \mathbb{R}^d \rightarrow \mathbb{R}^{d-1}$ die Projektion auf die ersten $d - 1$ Komponenten. Dann ist $\pi(P)$ ein Polyeder in \mathbb{R}^{d-1} .

Beweis. Es sei

$$P = \left\{ x \in \mathbb{R}^d \mid \langle c_1, x \rangle \leq \beta_1, \dots, \langle c_k, x \rangle \leq \beta_k \right\}$$

ein Polyeder in \mathbb{R}^d . Sei $I = \{1, \dots, k\}$, dann definieren wir Indexmengen

$$I_0 = \{i \in I \mid c_{id} = 0\}$$

$$I_+ = \{i \in I \mid c_{id} > 0\}$$

$$I_- = \{i \in I \mid c_{id} < 0\}.$$

Ein Punkt $x' = (x_1, \dots, x_{d-1})^T \in \mathbb{R}^{d-1}$ liegt genau dann in $\pi(P)$, wenn es $x_d \in \mathbb{R}$ gibt mit

$$\sum_{j=1}^{d-1} c_{ij} x_j \leq \beta_i \quad \text{für alle } i \in I_0,$$

$$x_d \leq \frac{\beta_i}{c_{id}} - \sum_{j=1}^{d-1} \frac{c_{ij}}{c_{id}} x_j \quad \text{für alle } i \in I_+ \text{ und}$$

$$x_d \geq \frac{\beta_i}{c_{id}} - \sum_{j=1}^{d-1} \frac{c_{ij}}{c_{id}} x_j \quad \text{für alle } i \in I_-$$

Es gibt solches x_d genau dann, wenn

$$\sum_{j=1}^{d-1} c_{ij} x_j \leq \beta_i \quad \text{für alle } i \in I_0 \text{ und}$$

$$\frac{\beta_i}{c_{id}} - \sum_{j=1}^{d-1} \frac{c_{ij}}{c_{id}} x_j \leq \frac{\beta_k}{c_{kd}} - \sum_{l=1}^{d-1} \frac{c_{kl}}{c_{kd}} x_l \quad \text{für alle } i \in I_- \text{ und } k \in I_+$$

Damit ist $x' \in \pi(P)$ genau dann, wenn x' diese beiden Systeme linearer Ungleichungen erfüllt. Also ist $\pi(P)$ ein Polyeder. \square

Übung 10.1. Wenn $P \subset \mathbb{R}^d$ durch k Ungleichungen beschrieben ist, was lässt sich aus dem Beweis des Satzes über die minimale Anzahl der Ungleichungen sagen, die $\pi(P) \subset \mathbb{R}^{d-1}$ beschreiben?

Übung 10.2. Sei $P \subset \mathbb{R}^d$ ein Polyeder und sei $\varphi: \mathbb{R}^d \rightarrow \mathbb{R}^e$ eine injektive lineare Abbildung. Dann ist $\varphi(P)$ ein Polyeder in \mathbb{R}^e .

Korollar 10.2. *Das Bild eines Polyeders in \mathbb{R}^d unter einer linearen Abbildung ist ein Polyeder.*

Beweis. Es sei $P \subset \mathbb{R}^d$ ein Polyeder und $\varphi: \mathbb{R}^d \rightarrow \mathbb{R}^e$ linear. Setze

$$\widehat{\varphi}: \begin{cases} \mathbb{R}^d & \rightarrow & \mathbb{R}^{d+e} \\ x & \mapsto & (x, \varphi(x)) \end{cases}$$

Offenbar ist $\widehat{\varphi}$ injektiv und damit $\widehat{\varphi}(P)$ ein Polyeder in \mathbb{R}^{d+e} nach Aufgabe 10.2. Sei $\pi: \mathbb{R}^{d+e} \rightarrow \mathbb{R}^e$ die Projektion auf die hinteren e Komponenten, so dass $\varphi(P) = \pi(\widehat{\varphi}(P))$ gilt. Durch wiederholte Anwendung von Satz 10.1 sieht man, dass $\pi(\widehat{\varphi}(P))$ ein Polyeder ist. \square

Wir wenden uns nun der linearen Programmierung zu: Wir schreiben $x \geq y$ für $x, y \in \mathbb{R}^n$, falls $x - y \in \mathbb{R}_+^n$ gilt. Gegeben seien eine $m \times n$ -Matrix A und zwei Vektoren $b \in \mathbb{R}^n$ und $b^* \in \mathbb{R}^m$. Das zugehörige *lineare Programm* ist das Kegelprogramm in kanonischer Form:

$$(P) \quad \left. \begin{array}{l} \text{Finde} \quad \beta = \inf \langle x, b \rangle \\ \text{unter} \quad Ax = b^* \\ \quad \quad x \geq 0 \end{array} \right\} \text{ in der Variablen } x \in \mathbb{R}^n.$$

Das duale Problem ist

$$(D) \quad \left. \begin{array}{l} \text{Finde} \quad \beta^* = \sup \langle b^*, y \rangle \\ \text{unter} \quad A^T y \leq b \end{array} \right\} \text{ in der Variablen } y \in \mathbb{R}^m.$$

Satz 10.3 (Starke Dualität). *Es sei (P) ein lineares Programm wie oben. Wenn (P) einen zulässigen Punkt besitzt, dann gibt es keine Dualitätslücke, d.h. es gilt $\beta = \beta^*$. Ist (P) zusätzlich beschränkt, so besitzen (P) und (D) optimale Punkte.*

Beweis. Da \mathbb{R}_+^n ein Polyeder ist, ist der Kegel $\{(Ax, \langle x, b \rangle) \mid x \geq 0\}$ nach Kor. 10.2 ein Polyeder in \mathbb{R}^{n+1} und damit abgeschlossen. Nach Satz 9.12 hat (P) also keine Dualitätslücke, und falls (P) beschränkt ist, besitzt (P) einen optimalen Punkt. Es bleibt nur zu zeigen, dass dann auch (D) einen optimalen Punkt besitzt. Dazu schreiben wir (D) in kanonische Form um: Betrachte das Programm

$$(D') \quad \left. \begin{array}{l} \text{Finde} \quad y = \inf \langle -b^*, w_1 - w_2 \rangle \\ \text{unter} \quad A^T(w_1 - w_2) + z = b \\ \quad \quad w_1, w_2, z \geq 0 \end{array} \right\} \text{ in der Variablen } (w_1, w_2, z)^T \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n.$$

Das Programm (D') ist im wesentlichen ein lineares Programm auf $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n$ in kanonischer Form. (Denn in der ersten Zeile könnten wir $\langle -b^*, w_1 - w_2 \rangle$ auch umständlicher als $\langle ((-b^*, b^*, 0)^T, (w_1, w_2, z))^T \rangle$ schreiben, und in der zweiten Zeile statt $A^T(w_1 - w_2) + z$ auch $B(w_1, w_2, z)^T$, wobei $B = (A^T, -A^T, I_n) \in \text{Mat}(n \times (2m + n))$.) Ist $y \in \mathbb{R}^m$ ein zulässiger Punkt von (D), so wähle $w_1, w_2 \in \mathbb{R}_+^m$ mit $y = w_1 - w_2$ und setze $z = b - A^T y$. Dann ist $(w_1, w_2, z)^T$ ein zulässiger Punkt von (D'). Umgekehrt ist $y = w_1 - w_2$ für jeden zulässigen Punkt $(w_1, w_2, z)^T$ von (D') ein zulässiger Punkt von (D). Dies zeigt $\beta^* = -\gamma$, insbesondere ist (D') beschränkt. Da auch (P) beschränkt ist, besitzt (D) nach Kor. 9.9 einen zulässigen Punkt und damit auch (D'). Also besitzen (D') und (D) einen optimalen Punkt. \square

Korollar 10.4. *Angenommen, das primäre Programm (P) und das duale Programm (D) sind beide beschränkt. Dann gelten*

- (1) (P) und (D) besitzen optimale Punkte.
- (2) (Optimalitätskriterium und Komplementärer Schlumpf) Seien $x \in \mathbb{R}^n$ bzw. $y \in \mathbb{R}^m$ zulässige Punkte von (P) bzw. (D). Genau dann sind x und y optimal, wenn gilt:

$$x_i > 0 \quad \implies \quad \sum_{j=1}^m a_{ji} y_j = b_i.$$

Beweis. Dies ist die Kombination von Satz 9.8 und Satz 10.3. □

Übung 10.3. Beweise das *Farkassche Lemma*: Es seien $c_1, \dots, c_m \in \mathbb{R}^n$ und

$$P = \{x \in \mathbb{R}^n \mid \langle c_i, x \rangle \leq 0 \text{ für } i = 1, \dots, m\}.$$

Für $c \in \mathbb{R}^n$ gilt: Genau dann ist $\langle c, x \rangle \leq 0$ für alle $x \in P$, wenn es $\lambda_1, \dots, \lambda_m \geq 0$ gibt derart, dass $c = \lambda_1 c_1 + \dots + \lambda_m c_m$ gilt.

ANWENDUNG: DAS TRANSPORT-PROBLEM

Ein *gerichteter Graph* ist ein Paar (N, A) bestehend aus einer Menge N („Knoten“) und einer Teilmenge $A \subset N \times N$ („Pfeile“). Für $n_1, n_2 \in N$ schreiben wir $n_1 \rightarrow n_2$, falls $(n_1, n_2) \in A$. Eine *Schleife* ist ein Pfeil der Form $n \rightarrow n$ für $n \in N$. Das Transport-Problem ist das folgende allgemeine Optimierungsproblem. Gegeben sind:

Objekt	Interpretation
Ein endlicher gerichteter Graph $\Gamma = (N, A)$ mit $N = \{1, \dots, n\}$, $n \in \mathbb{N}$, der keine Schleifen enthält.	Die Knoten repräsentieren verschiedene Orte, zwischen denen ein Gut den Pfeilen entsprechend transportiert wird.
Ein Vektor $b \in \mathbb{R}^n$.	Der Wert b_i repräsentiert die Differenz zwischen <i>Nachfrage</i> und <i>Angebot</i> am Knoten i . Falls $b_i > 0$, so übersteigt die Nachfrage das Angebot, sonst umgekehrt.
Eine Funktion $c: A \rightarrow \mathbb{R}$, notiert in Indexschreibweise $c(i \rightarrow j) =: c_{ij}$, mit $c_{ij} \geq 0$ für alle $(i \rightarrow j) \in A$.	Der Wert c_{ij} repräsentiert die <i>Transportkosten</i> von i nach j pro transportierter Einheit.

Wir betrachten das lineare Programm

$$(P) \quad \left. \begin{array}{l} \text{Finde } \beta = \inf \sum_{(i \rightarrow j) \in A} c_{ij} x_{ij} \\ \text{unter } \sum_{j: (j \rightarrow i) \in A} x_{ji} - \sum_{j: (i \rightarrow j) \in A} x_{ij} = b_i \text{ für } i \in N \\ \quad \quad \quad x_{ij} \geq 0 \text{ für alle } (i \rightarrow j) \in A. \end{array} \right\} \text{ in der Variablen } x \in \mathbb{R}^A.$$

Ein Punkt $x \in \mathbb{R}^A$ wird als *Fluß* bezeichnet und repräsentiert die Gesamtbewegung des Guts zwischen den Knoten. Ein Fluß ist für das Programm zulässig, wenn Angebot und Nachfrage an jedem Knoten ausgeglichen werden (*Balancebedingung*). Der Wert β repräsentiert die gesamten Transportkosten, und das Optimierungsproblem besteht darin, diese zu minimieren. Das duale Programm sieht wie folgt aus:

$$(D) \quad \left. \begin{array}{l} \text{Finde } \beta^* = \sup \sum_{i=1}^n b_i y_i \\ \text{unter } y_j - y_i \leq c_{ij} \text{ für alle } (i \rightarrow j) \in A \end{array} \right\} \text{ in der Variablen } y \in \mathbb{R}^n.$$

Übung 10.4. Überprüfe, dass (D) tatsächlich das duale Programm zu (P) ist.

Dieses duale Problem, das rein formal gebildet wurde, kann man wie folgt interpretieren: Der Wert y_i repräsentiert den *Preis* des Guts am Knoten i und damit β^* den Gesamtgewinn, der durch Kauf und Verkauf des Guts an den verschiedenen Knoten erzielt wird. Die Zulässigkeitsbedingung für die Preisverteilung y sagt dann, dass unter Berücksichtigung der Transportkosten kein Gewinn durch Kauf an einem und Verkauf an einem anderen Knoten erzielt werden kann.

Korollar 10.5. Sei (P) das obige Transportproblem. Falls (P) einen zulässigen Fluß besitzt, dann gibt es einen optimalen Fluss und eine optimale Preisverteilung. Ist $x = (x_{ij}: (i \rightarrow j) \in A)$ ein zulässiger Fluß und y eine zulässige Preisverteilung, so sind x und y genau dann optimal, wenn

$$x_{ij} > 0 \implies y_j - y_i = c_{ij}$$

für alle $(i \rightarrow j) \in A$ gilt.

Beweis. Da $c_{ij} \geq 0$ nach Voraussetzung gilt, ist (P) sicher beschränkt. Außerdem ist $y = 0$ eine zulässige Preisverteilung. Damit folgt die Behauptung aus Kor. 10.4. \square

Es sei $\Gamma = (N, A)$ ein gerichteter Graph. Ein *Zykel* in Γ ist eine endliche Folge von Pfeilen $n_1 \leftrightarrow n_2 \leftrightarrow \dots \leftrightarrow n_m \leftrightarrow n_1$ in A mit $m \geq 2$, wobei die Notation $n_1 \leftrightarrow n_2$ bedeutet, dass der Pfeil entweder $n_1 \rightarrow n_2$ oder $n_2 \rightarrow n_1$ ist. Wenn in einem Fluß des Transportproblems ein echter Transport entlang eines Zyklus stattfindet, dann bekommt man intuitiv den Eindruck, dass der Fluß nicht optimal sein kann. Dieser Eindruck wird durch die folgende Aussage im Wesentlichen bestätigt:

Proposition 10.6. Es sei $\Gamma = (N, A)$ ein gerichteter Graph ohne Schleifen, $b \in \mathbb{R}^n$ und

$$T(\Gamma, b) = \left\{ x_{ij} \in \mathbb{R}^A \mid \sum_{j:(j \rightarrow i) \in A} x_{ji} - \sum_{j:(i \rightarrow j) \in A} x_{ij} = b_i \text{ für alle } i \in N \text{ und } x_{ij} \geq 0 \text{ für alle } (i \rightarrow j) \in A \right\}$$

das zugehörige Transportpolyeder. Sei $x \in T(\Gamma, b)$ ein Extrempunkt und $S \subset A$ die Menge aller Pfeile $(i \rightarrow j)$ mit $x_{ij} > 0$. Dann enthält S keine Zykel.

Beweis. Sei $x \in T$ und angenommen $C \subset A$ ist ein Zykel mit $x_{ij} > 0$ für $(i \rightarrow j) \in C$. Es seien $n_1 \leftrightarrow n_2 \leftrightarrow \dots \leftrightarrow n_m \leftrightarrow n_1$ die in C verbundenen Knoten und sei $\varepsilon > 0$ mit $\varepsilon < \min\{x_{ij} \mid (i \rightarrow j) \in C\}$. Wir definieren neue Flüsse

$$u_{ij} = \begin{cases} x_{ij} & \text{falls } (i \rightarrow j) \notin C \text{ und } (j \rightarrow i) \notin C \\ x_{ij} + \varepsilon & \text{falls } i = n_k, j = n_{k+1} \\ x_{ij} - \varepsilon & \text{falls } i = n_{k+1}, j = n_k \end{cases} \quad \begin{array}{l} \text{für ein } k \in \{1, \dots, m\} \\ \text{für ein } k \in \{1, \dots, m\} \end{array}$$

$$v_{ij} = \begin{cases} x_{ij} & \text{falls } (i \rightarrow j) \notin C \text{ und } (j \rightarrow i) \notin C \\ x_{ij} - \varepsilon & \text{falls } i = n_k, j = n_{k+1} \\ x_{ij} + \varepsilon & \text{falls } i = n_{k+1}, j = n_k \end{cases} \quad \begin{array}{l} \text{für ein } k \in \{1, \dots, m\} \\ \text{für ein } k \in \{1, \dots, m\}. \end{array}$$

Das heißt also, wir erhöhen und reduzieren den Fluß auf C um ε in Abhängigkeit von der Orientierung der Pfeile. Nach Konstruktion gilt dann $u, v \in T(\Gamma, b)$ und $x = \frac{1}{2}(u + v)$, aber $u, v \neq x$. Also ist x nicht extremal. \square

Übung 10.5. Zeige: Ist $T(\Gamma, b)$ wie oben nicht leer, so gilt $\sum_{i=1}^n b_i = 0$. Zeige durch ein Beispiel, dass die Umkehrung falsch ist.

Übung 10.6. Sei $\Gamma = (N, A)$ ein gerichteter Graph ohne Schleifen. Ein *gerichteter Zykel* ist eine Folge von Knoten und Pfeilen $n_1 \rightarrow n_2 \rightarrow \dots \rightarrow n_m \rightarrow n_1$ mit $m \geq 2$.

Zeige: Genau dann ist $T(\Gamma, b)$ für alle $b \in \mathbb{R}^n$ kompakt, wenn Γ keine gerichteten Zykel enthält.

11. SEMIDEFINITE PROGRAMMIERUNG

11.1. DUALITÄT FÜR SEMIDEFINITE PROGRAMME

Ein *semidefinites Programm* ist ein Kegelprogramm für den Kegel der positiv semidefiniten Matrizen (psd-Kegel). Wir erinnern an die folgenden Tatsachen aus Abschnitt 6:

- (1) Auf dem Vektorraum Sym_n der symmetrischen $n \times n$ -Matrizen hat das kanonische euklidische Skalarprodukt die Form

$$\langle A, B \rangle = \text{tr}(AB).$$

- (2) Der psd-Kegel ist unter diesem Skalarprodukt selbstdual, d.h. es gilt $(S_n^+)^* = S_n^+$.
 (3) Wir schreiben die Ordnungsrelation $\succcurlyeq_{S_n^+}$ als \succeq . Wie zuvor bedeutet also $A \succeq 0$, dass A positiv semidefinit ist.

Lemma 11.1. *Es sei $\varphi: \text{Sym}_n \rightarrow \mathbb{R}^m$ eine lineare Abbildung, $m, n \geq 1$.*

- (1) *Es gibt Matrizen $A_1, \dots, A_m \in \text{Sym}_n$ derart, dass*

$$\varphi(X) = (\langle A_1, X \rangle, \dots, \langle A_m, X \rangle)$$

für alle $X \in \text{Sym}_n$ gilt.

- (2) *Die duale Abbildung $\varphi^*: \mathbb{R}^m \rightarrow \text{Sym}_n$ ist dann gegeben durch*

$$\varphi^*(y) = y_1 A_1 + \dots + y_m A_m$$

für $y \in \mathbb{R}^m$.

Beweis. (1) Ist $\varphi = (\varphi_1, \dots, \varphi_m)$, dann sind $\varphi_1, \dots, \varphi_m$ lineare Funktionale auf Sym_n . Da Sym_n ein endlich-dimensionaler Vektorraum ist, ist jedes lineare Funktional durch das Skalarprodukt mit einem Element von Sym_n gegeben.

- (2) Für alle $X \in \text{Sym}_n$ und $y \in \mathbb{R}^m$ gilt

$$\begin{aligned} \langle X, \varphi^*(y) \rangle &= \langle \varphi(X), y \rangle = \langle (\langle A_1, X \rangle, \dots, \langle A_m, X \rangle), y \rangle = y_1 \langle A_1, X \rangle + \dots + y_m \langle A_m, X \rangle \\ &= \langle X, y_1 A_1 + \dots + y_m A_m \rangle. \end{aligned}$$

Da die duale Abbildung nach Prop. 9.4 eindeutig ist, folgt $\varphi^*(y) = y_1 A_1 + \dots + y_m A_m$. \square

Damit haben wir alles beisammen, um ein semidefinites Programm und sein Dual hinzuschreiben. Seien $A_1, \dots, A_m, B \in \text{Sym}_n$ und $b^* \in \mathbb{R}^m$.

Primäres Programm:

$$(P) \quad \left. \begin{array}{l} \text{Finde } \beta = \inf \langle B, X \rangle \\ \text{unter } \langle A_i, X \rangle = b_i^* \text{ für } i = 1, \dots, m \\ X \succeq 0 \end{array} \right\} \text{ in der Variablen } X \in \text{Sym}_n.$$

Duales Programm:

$$(D) \quad \left. \begin{array}{l} \text{Finde } \beta^* = \sup \langle b^*, y \rangle \\ \text{unter } \sum_{i=1}^m y_i A_i \leq B \end{array} \right\} \text{ in der Variablen } y \in \mathbb{R}^m.$$

Übung 11.1. Es sei $n = 2$, $m = 1$. Setze

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad b^* = 1.$$

Zeige: (1) $\beta = \beta^* = 0$; (2) $y_1 = 0$ ist optimaler Punkt von (D); (3) (P) besitzt keinen optimalen Punkt.

Übung 11.2. Es sei $n = 2$, $m = 1$. Setze

$$B = \begin{bmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad b^* = 0.$$

Zeige: (1) $\beta = 0$; (2) (P) besitzt einen optimalen Punkt. (3) (D) besitzt keinen zulässigen Punkt.

Um ein Kriterium für starke Dualität zu bekommen, brauchen wir das folgende

Lemma 11.2. *Es sei $K \subset \mathbb{R}^n$ ein spitzer abgeschlossener Kegel und $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung mit $\ker(\varphi) \cap K = \{0\}$. Dann ist $\varphi(K)$ wieder ein abgeschlossener Kegel.*

Beweis. Nach Prop. 5.12(2) besitzt K eine kompakte Kegelbasis $B \subset \mathbb{R}^n$. Aus der Voraussetzung folgt $0 \notin \varphi(B)$. Deshalb ist auch $\varphi(B)$ eine kompakte Kegelbasis von $\varphi(K)$, so dass $\varphi(K)$ (wiederum nach Prop. 5.12) abgeschlossen ist. \square

Satz 11.3 (Starke Dualität). *Es sei (P), (D) ein duales Paar von semidefiniten Programmen wie oben und sei $L = \{y_1 A_1 + \dots + y_m A_m + y_{m+1} B \mid y \in \mathbb{R}^{m+1}\}$. Falls L eine positiv definite Matrix enthält und (P) einen zulässigen Punkt besitzt, so gibt es keine Dualitätslücke. Ist (P) außerdem beschränkt, so besitzt (P) einen optimalen Punkt.*

Beweis. Wir wollen Satz 9.12 anwenden. Es sei $\varphi: \text{Sym}_n \rightarrow \mathbb{R}^m$ gegeben durch $\varphi(X) = (\langle A_1, X \rangle, \dots, \langle A_m, X \rangle)$ und $\widehat{\varphi}: \text{Sym}_n \rightarrow \mathbb{R}^{m+1}$ durch $\widehat{\varphi}(X) = (\varphi(X), \langle B, X \rangle)$. Nach Voraussetzung gibt es $C \in L \cap S_n^{++}$. Dann gilt für jedes $X \in \ker(\widehat{\varphi})$, so folgt $\langle C, X \rangle = 0$. Aus Aufgabe 11.3 unten folgt damit $\ker(\widehat{\varphi}) \cap S_n^+ = \{0\}$. Nach dem vorangehenden Lemma ist $\widehat{\varphi}(S_n^+)$ abgeschlossen und die Aussagen des Satzes folgen damit aus Satz 9.12. \square

Übung 11.3. Es seien $X, Y \in \text{Sym}_n$ positiv semidefinit mit $\langle X, Y \rangle = 0$.

Zeige, dass dann $\text{rk } X + \text{rk } Y \leq n$ gilt. Insbesondere: Ist Y positiv definit, so ist $X = 0$.

Der Vollständigkeit halber notieren wir noch die Optimalitätskriterien für semidefinite Programme, ein Spezialfall von Thm. 9.8:

Korollar 11.4. *Es sei (P), (D) ein duales Paar von semidefiniten Programmen wie oben. Sei X ein zulässiger Punkt von (P) und y ein zulässiger Punkt von (D). Dann gelten:*

(1) (Optimalitätskriterium) Falls

$$\langle X, B - \sum_{i=1}^m y_i A_i \rangle = 0$$

gilt, dann sind X und y optimal.

(2) (Komplementärer Schlupf) Sind X und y optimal und gilt $\beta = \beta^*$, dann folgt

$$\langle X, B - \sum_{i=1}^m y_i A_i \rangle = 0. \quad \square$$

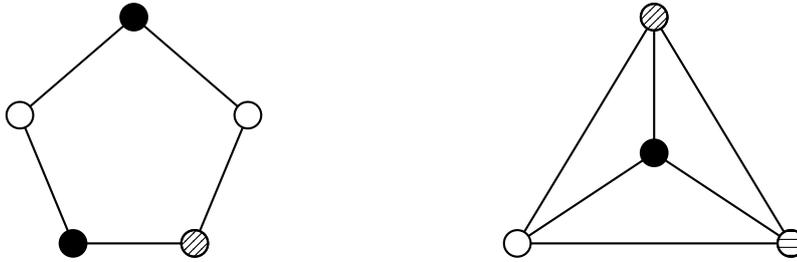
11.2. ANWENDUNG: CLIQUEN- UND FÄRBUNGSZAHL EINES GRAPHEN

Ein *Graph* (im Unterschied zu einem gerichteten Graph, was wir beim letzten Mal definiert hatten) besteht aus einer Menge N („Knoten“) und einer symmetrischen Relation $E \subset N \times N$ („Kanten“). Sind $m, n \in N$ Knoten mit $(m, n) \in E$, dann sagt man, dass m und n durch eine Kante *verbunden* sind. Wie zuvor nehmen wir stets an, dass es keine Schleifen gibt, also keine Kanten, die einen Knoten mit sich selbst verbinden.

Definition 11.5. Es sei $\Gamma = (N, E)$ ein Graph.

- (1) Eine *Clique* in Γ ist eine Teilmenge $C \subset N$ von Knoten derart, dass je zwei verschiedene Knoten in C durch eine Kante verbunden sind. Die Mächtigkeit einer maximalen Clique in Γ heißt die *Cliquenzahl* von Γ und wird mit $\omega(\Gamma)$ bezeichnet.
- (2) Eine *k-Färbung* von Γ ist eine Abbildung $\varphi: N \rightarrow \{1, \dots, k\}$ mit der Eigenschaft $\varphi(m) \neq \varphi(n)$ für alle $(m, n) \in E$. Die kleinste Zahl k derart, dass Γ eine *k-Färbung* besitzt heißt die *Färbungszahl* von Γ und wird mit $\chi(\Gamma)$ bezeichnet.

Es gilt $\omega(\Gamma) \leq \chi(\Gamma)$, da eine *k-Clique* keine Färbung mit weniger als k Farben haben kann.



Der berühmte *Vier-Farben-Satz* (bewiesen 1976 durch K. Appel und W. Haken) besagt, dass sich jede Landkarte (mit zusammenhängenden Ländern) durch vier Farben so färben lässt, dass je zwei Länder, die eine echte Grenze miteinander haben, verschiedene Farben haben. Formalisiert wird dies zu der Aussage, dass jeder ebene Graph (d.h. ein Graph der sich ohne Überkreuzungen von Kanten in \mathbb{R}^2 einbetten lässt) eine Vierfärbung besitzt. Im allgemeinen ist es rechnerisch sehr aufwendig („NP-vollständig“) die Cliquen- oder Färbungszahl für einen gegebenen Graph zu ermitteln oder auch nur gut zu approximieren. Zum Beispiel kann ein ebener Graph auch eine Dreifärbung besitzen, die aber im allgemeinen nicht leicht zu finden ist. (Diese Tatsache kann man sich auch in der Kryptographie zunutze machen.) Einer der ersten wichtigen Sätze (bewiesen von L. Lovász im Jahr 1978), der die Nützlichkeit der semidefiniten Programmierung für die kombinatorische Optimierung aufgezeigt hat, besagt jedoch, dass es ein duales Paar von semidefiniten Programmen gibt, deren optimaler Wert eine Zahl $\vartheta(\Gamma)$ mit der Eigenschaft $\omega(\Gamma) \leq \vartheta(\Gamma) \leq \chi(\Gamma)$ ist. Die Funktion $\vartheta: \Gamma \mapsto \vartheta(\Gamma)$ heißt die *Lovázsche Theta-Funktion*¹ Wir beschreiben nun, wie das entsprechende semidefinite Programm gebaut wird:

Konstruktion 11.6. Es sei $\Gamma = (N, E)$ ein Graph mit n Knoten $N = \{1, \dots, n\}$. Für $i, j \in \{1, \dots, n\}$ sei A_{ij} die Matrix mit $a_{ij} = a_{ji} = 1$ und alle anderen Einträge gleich 0. Sei J die $n \times n$ -Matrix, deren Einträge alle gleich 1 sind und I die $n \times n$ -Einheitsmatrix.

¹Es besteht kein Zusammenhang mit den diversen Typen von Theta-Funktionen in der komplexen Analysis, Zahlentheorie und algebraischen Geometrie.

Wir definieren das folgende *Primäre Programm*:

$$(P) \quad \left. \begin{array}{l} \text{Finde} \quad \vartheta(\Gamma) = \sup \langle J, X \rangle = - \inf \langle -J, X \rangle \\ \text{unter} \quad \langle A_{ij}, X \rangle = 0 \text{ für } i < j \text{ mit } (i, j) \notin E \\ \quad \quad \text{tr}(X) = \langle I, X \rangle = 1 \\ \quad \quad X \geq 0 \end{array} \right\} \text{ in der Variablen } X \in \text{Sym}_n.$$

Übung 11.4. Wie sieht das duale Programm dazu aus?

Bemerkung 11.7. Das Programm (P) und sein Dual haben keine Dualitätslücke. Denn die Matrix $\frac{1}{n}I$ ist ein zulässiger Punkt von (P). Außerdem ist I positiv definit und beschreibt eine der Nebenbedingungen in (P). Nach Satz 11.3 gibt es deshalb keine Dualitätslücke.

Satz 11.8 (Lovász). *Ist $\Gamma = (N, E)$ ein Graph und (P) wie oben, dann gilt*

$$\omega(\Gamma) \leq \vartheta(\Gamma) \leq \chi(\Gamma).$$

Beweis. Als erstes kümmern wir uns um die Cliquenzahl: Es sei $W \subset N$ eine Clique und sei $k = |W|$. Sei $x = (x_1, \dots, x_n)$ der Vektor

$$x_i = \begin{cases} 1 & \text{falls } i \in W \\ 0 & \text{falls } i \notin W \end{cases}$$

und setze $X = (1/k) \cdot xx^T$. Dann ist X ein zulässiger Punkt von (P) und es gilt $\langle J, X \rangle = k$ und damit $\vartheta(\Gamma) \geq \omega(\Gamma)$.

Jetzt kommt die Färbungszahl dran. Es sei X ein zulässiger Punkt von (P) (z.B. $X = \frac{1}{n}I$). Da X positiv semidefinit ist, gibt es nach Prop. 6.2 eine Matrix $U \in \text{Mat}_n$ mit $X = U^T U$. Es seien u_1, \dots, u_n die Spalten von U , so dass $\langle u_i, u_j \rangle = x_{ij}$ gilt.

Sei nun $\varphi: N \rightarrow \{1, \dots, k\}$ eine k -Färbung von Γ . Da X ein zulässiger Punkt ist, gilt $x_{ij} = 0$, wann immer $\varphi(i) \neq \varphi(j)$. Wegen $\text{tr}(X) = \langle I, X \rangle = 1$ gilt außerdem

$$\sum_{i=1}^n \|u_i\|^2 = 1.$$

Für jede Farbe $1 \leq j \leq k$ definieren wir einen Vektor

$$w_j = \sum_{i:\varphi(i)=j} u_i.$$

Es gilt dann

$$\|w_j\|^2 = \sum_{i_1, i_2: \varphi(i_1)=\varphi(i_2)=j} \langle u_{i_1}, u_{i_2} \rangle = \sum_{i:\varphi(i)=j} \|u_i\|^2,$$

die zweite Gleichheit weil die Summanden in der Mitte für $i_1 \neq i_2$ gleich 0 sind, und damit

$$\sum_{j=1}^k \|w_j\|^2 = 1.$$

Wegen $\sum_{j=1}^k w_j = \sum_{i=1}^n u_i$ gilt außerdem

$$\langle J, X \rangle = \left\| \sum_{j=1}^k w_j \right\|^2 \leq \left(\sum_{j=1}^k \|w_j\| \right)^2.$$

Nach der nachfolgenden Aufgabe folgt daraus nun $\langle J, X \rangle \leq k$ und damit $\vartheta(\Gamma) \leq \chi(\Gamma)$. \square

Übung 11.5. Es seien $\alpha_1, \dots, \alpha_k$ reelle Zahlen mit $\sum_{i=1}^k \alpha_i^2 = 1$. Zeige, dass $\sum_{i=1}^k \alpha_i \leq \sqrt{k}$ gilt.

Übung 11.6. Es sei Γ das Fünfeck mit Knoten 1, 2, 3, 4, 5 und Kanten (1, 2), (2, 3), (3, 4), (4, 5), (5, 1). Zeige, dass $\omega(\Gamma) = 2$, $\chi(\Gamma) = 3$, und $\vartheta(\Gamma) = \sqrt{5}$ gelten.

12. INNERE-PUNKTE-VERFAHREN

In diesem Abschnitt geben wir eine kurze Einleitung in einige der Methoden, die den Hintergrund für praxistaugliche Algorithmen zur Lösung von semidefiniten Programmen bilden. Dieser Exkurs beschränkt sich auf einige wesentliche Punkte und wird mathematisch zum Teil weniger präzise und vor allem weniger vollständig sein als der Rest der Vorlesung. Die Darstellung hier orientiert sich in erster Linie am Buch von Forst und Hoffmann *Optimization – Theory and Practice* [3], in dem sich natürlich auch die fehlenden Puzzleteile finden.

12.1. DAS PRIMÄR-DUALE SYSTEM

Im folgenden sei stets

$$\begin{array}{ll}
 \text{(P)} & \left. \begin{array}{l} \text{Finde } \beta = \inf \langle B, X \rangle \\ \text{unter } \langle A_i, X \rangle = b_i^* \text{ für } i = 1, \dots, m \\ X \geq 0 \end{array} \right\} \text{ in der Variablen } X \in \text{Sym}_n. \\
 \text{(D)} & \left. \begin{array}{l} \text{Finde } \beta^* = \sup \langle b^*, y \rangle \\ \text{unter } \sum_{i=1}^m y_i A_i \leq B \end{array} \right\} \text{ in der Variablen } y \in \mathbb{R}^m.
 \end{array}$$

ein duales Paar von semidefiniten Programmen. Wir machen zunächst geeignete Annahmen an das System, um starke Dualität und die Anwendbarkeit der betrachteten Methoden sicherzustellen.

Definition 12.1. Das Paar (P),(D) genügt der *Innere-Punkte-Bedingung*, wenn gelten:

- Die Matrizen A_1, \dots, A_m sind linear unabhängig.
- (P) hat einen *strikt zulässigen Punkt*, d.h. eine positiv definite zulässige Matrix X .
- (D) hat einen strikt zulässigen Punkt, d.h. einen Punkt y derart, dass $B - \sum_{i=1}^m y_i A_i$ positiv definit ist.

Die Innere-Punkte-Bedingung stellt unter anderem sicher, dass es keine Dualitätslücke gibt. Dies folgt aus einer allgemeineren Aussage über Kegelprogramme, die sich als Anwendung von Satz 9.8 ergibt (siehe Barvinok [1], S. 169ff).

Proposition 12.2. *Es sei*

$$\text{(Q)} \quad \left. \begin{array}{l} \text{Finde } \beta = \inf \langle x, b \rangle \\ \text{unter } \varphi(x) = b^* \\ x \succeq_K 0 \end{array} \right\} \text{ in der Variablen } x \in V_1.$$

ein Kegelprogramm in kanonischer Form (Notationen wie in Def. 9.11). Falls (Q) einen zulässigen Punkt x im Inneren von K besitzt, mit einer offenen Umgebung W von x in K derart, dass $0 \in V_2$ ein innerer Punkt von $\varphi(W)$ ist, dann gibt es keine Dualitätslücke. \square

Korollar 12.3. Falls das duale Paar $(P), (D)$ von semidefiniten Programmen der Innere-Punkte-Bedingung genügt, dann gibt es keine Dualitätslücke.

Beweis. Da A_1, \dots, A_m linear unabhängig sind, ist in (P) die lineare Abbildung $\varphi: X \mapsto (\langle A_1, X \rangle, \dots, \langle A_m, X \rangle)$ surjektiv. Surjektive lineare Abbildungen sind offen (d.h. Bilder offener Mengen sind offen), so dass die Behauptung aus dem vorangehenden Satz folgt. \square

Als nächstes möchten wir das duale Programm so umschreiben, dass es mehr der Form des primären entspricht. Das ist in der semidefiniten Programmierung nicht in derselben Weise möglich, wie in der linearen (vgl. Satz 10.3). Statt dessen benutzen wir das Folgende:

Proposition 12.4. Es sei $(P), (D)$ ein duales Paar von semidefiniten Programmen wie oben, das der Innere-Punkte-Bedingung genügt. Seien X bzw. y zulässige Punkte von (P) bzw. (D) . Genau dann sind X und y optimal, wenn es eine Matrix $Y \in \text{Sym}_n$ gibt derart, dass die folgenden Bedingungen erfüllt sind:

$$(PD) \quad \begin{cases} \langle A_i, X \rangle = b^* \text{ für } i = 1, \dots, m \\ Y + \sum_{i=1}^m y_i A_i = B \\ X \geq 0, Y \geq 0 \\ XY = 0. \end{cases}$$

Beweis. Nach Kor. 12.3 gibt es keine Dualitätslücke. Damit folgt die Behauptung aus Kor. 11.4. (Beachte auch die nachfolgende Aufgabe.) \square

Übung 12.1. Für $X, Y \in S_n^+$ gilt $\langle X, Y \rangle = 0$ genau dann, wenn $XY = 0$.

Definition 12.5. Das System (PD) in den Unbestimmten $(X, Y, y) \in \text{Sym}_n \times \text{Sym}_n \times \mathbb{R}^m$ heißt das *primär-duale System* zum Paar $(P), (D)$.

12.2. BARRIEREFUNKTIONEN UND DER ZENTRALE PFAD

Definition 12.6. Es sei $C \subset \mathbb{R}^n$ eine konvexe Teilmenge. Eine Funktion $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ heißt *konvex*, wenn

$$\lambda \varphi(x) + (1 - \lambda) \varphi(y) \geq \varphi(\lambda x + (1 - \lambda)y)$$

für alle $x, y \in C$ und $\lambda \in [0, 1]$ gilt. Die Funktion φ heißt *strikt konvex*, wenn die angegebene Ungleichung für $x \neq y$ und $\lambda \in (0, 1)$ strikt ist.

Übung 12.2. Ist φ stetig differenzierbar, so ist $\varphi: C \rightarrow \mathbb{R}$ genau dann konvex, wenn die Ungleichung

$$\varphi(y) \geq \varphi(x) + \langle (\nabla \varphi)(x), y - x \rangle$$

für alle $x, y \in C$ erfüllt ist.

Lemma 12.7. Die Funktion $-\log \circ \det$ ist strikt konvex und analytisch auf S_n^{++} und geht am Rand des psd-Kegels S_n^+ gegen ∞ . Ihre Ableitung in einem Punkt $X_0 \in S_n^{++}$ ist

$$\nabla(-\log \circ \det)(X_0) = -X_0^{-1}.$$

Beweisskizze. Setze $\varphi = -\log \circ \det$. Dass φ analytisch ist und am Rand des psd-Kegels gegen ∞ geht, ist nicht schwer zu sehen. Für die Ableitung der Determinante in einem Punkt $X_0 \in \text{Mat}_n$ mit $\det(X_0) \neq 0$ gilt nach der *Jacobi-Formel* $(\nabla \det)(X_0) = \det(X_0) \cdot X_0^{-1}$. Der Beweis findet sich z.B. in der englischen Wikipedia. Damit folgt die Aussage über die Ableitung von φ aus der Kettenregel.

Um zu zeigen, dass φ konvex ist, seien $X, Y \in S_n^{++}$ und seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte von $X^{-1}Y$. Diese sind reell und positiv (siehe nachfolgende Aufgabe). Es gilt

$$\begin{aligned}\varphi(Y) - \varphi(X) &= -\log \det(Y) + \log \det(X) = -\log \det(X^{-1}Y) \\ &= -\log \prod_{i=1}^n \lambda_i = -\sum_{i=1}^n \log(\lambda_i).\end{aligned}$$

Da $-\log$ konvex ist, und außerdem $\log(x) \leq x - 1$ für alle $x > 0$ ist, gilt

$$\frac{1}{n} \sum_{i=1}^n \log(\lambda_i) \leq \log\left(\frac{1}{n} \sum_{i=1}^n \lambda_i\right) \leq \frac{1}{n} \sum_{i=1}^n \lambda_i - 1.$$

Daraus ergibt sich insgesamt

$$\begin{aligned}\varphi(Y) - \varphi(X) &= -n\left(\frac{1}{n} \sum_{i=1}^n \log(\lambda_i)\right) \geq -\sum_{i=1}^n \lambda_i + n \\ &= \text{tr}(-X^{-1}Y + I_n) = \text{tr}(-X^{-1}(Y - X)) \\ &= \langle \nabla \varphi(X), Y - X \rangle.\end{aligned}$$

Nach Aufgabe 12.2 ist damit φ konvex. Es ist noch zu zeigen, dass φ auch strikt konvex ist. Angenommen, es gilt $\varphi(Y) - \varphi(X) = \langle (\nabla \varphi)(X), Y - X \rangle$ für $X, Y \in S_n^{++}$. Dann sind die obigen Ungleichungen allesamt Gleichungen. Da die Logarithmus-Funktion strikt konkav ist, folgt dann $\lambda_1 = \dots = \lambda_n$. Ferner gilt $\log(x) = x - 1$ genau für $x = 1$. Es folgt damit $\sum_{i=1}^n \lambda_i = n$ und schließlich $\lambda_1 = \dots = \lambda_n = 1$. Da $X^{-1}Y$ außerdem diagonalisierbar ist (s.u.), muss $X^{-1}Y = I_n$ gelten, also $X = Y$. \square

Übung 12.3. Seien $X, Y \in S_n^{++}$. Zeige, dass XY über \mathbb{R} diagonalisierbar ist und alle Eigenwerte positiv sind.

Definition 12.8. Sei $(P),(D)$ ein duales Paar von semidefiniten Programmen wie oben. Die Funktionen

$$\begin{aligned}F_\lambda(X) &= \langle B, X \rangle - \lambda \log(\det(X)) \\ G_\lambda(y) &= \langle b^*, y \rangle + \lambda \log(\det(B - \sum_{i=1}^m y_i A_i))\end{aligned}$$

heißen die *logarithmischen Barrierefunktionen des Paares $(P),(D)$ zum Parameter λ* .

Wir schreiben außerdem

$$\begin{aligned}S &= \left\{ X \in S_n^+ \mid \langle A_i, X \rangle = b_i^* \text{ für } i = 1, \dots, m \right\} \\ S^* &= \left\{ y \in \mathbb{R}^m \mid B - \sum_{i=1}^m y_i A_i \geq 0 \right\}.\end{aligned}$$

Satz 12.9. *Es sei $(P),(D)$ ein duales Paar von semidefiniten Programmen, das der Innere-Punkte-Bedingung genügt. Für $\lambda > 0$ und $(X, Y, y) \in \text{Sym}_n \times \text{Sym}_n \times \mathbb{R}^m$ sind die folgenden beiden Aussagen äquivalent:*

(a) *Es gelten*

$$(PD)_\lambda \quad \begin{cases} \langle A_i, X \rangle = b_i^* \text{ für } i = 1, \dots, m \\ Y + \sum_{i=1}^m y_i A_i = B \\ X \geq 0, Y \geq 0 \\ XY = \lambda I_n. \end{cases}$$

(b) *Es gelten:*

- (1) *Der Punkt X ist die eindeutig bestimmte Minimalstelle von F_λ auf S .*
- (2) *Der Punkt y ist die eindeutig bestimmte Maximalstelle von G_λ auf S^* .*
- (3) *$Y = B - \sum_{i=1}^m y_i A_i$.*

Insbesondere hat das System in (a) in den Unbestimmten (X, Y, y) eine eindeutige Lösung.

Beweisskizze. Wir schreiben $\varphi(X) = (\langle A_1, X \rangle, \dots, \langle A_m, X \rangle)$. Da die Funktion F_λ nach Lemma 12.7 strikt konvex ist und auf dem Rand von S gegen ∞ geht, nimmt sie ihr Minimum auf S in einem eindeutig bestimmten Punkt $X \in S_n^{++}$ an. Ferner ist das Minimum unter der Nebenbedingung $X \in S$, also $\varphi(X) = b^*$, durch die Lagrange-Bedingungen charakterisiert (die aus der Grundvorlesung Analysis in der einen oder anderen Form möglicherweise bekannt sind): Sei dazu $L(X, y) = F_\lambda(X) + \langle y, b^* - \varphi(X) \rangle$ die *Lagrange-Funktion* von F_λ . Dann nimmt F_λ ihr Minimum auf S genau dann in X an, wenn

$$\nabla_X L(X, y) = B - \lambda X^{-1} - \sum_{i=1}^m y_i A_i = 0$$

für ein $y \in \mathbb{R}^m$ gilt. (Die Einträge von y heißen die *Lagrange-Multiplikatoren*). Mit $Y = B - \sum_{i=1}^m y_i A_i$ bedeutet dies genau $XY = \lambda I_n$. (Aus $Y = \lambda X^{-1}$ folgt dabei insbesondere $Y \geq 0$). Wir sehen also: Genau dann ist (X, Y, y) eine Lösung von (a), wenn die Funktion F_λ in X ihr Minimum annimmt, mit zugehörigem Lagrange-Multiplikator y und daraus eindeutig bestimmtem Y (beachte, dass hier die lineare Unabhängigkeit von A_1, \dots, A_m eingehet). Das analoge Argument für G_λ zeigt die behauptete Äquivalenz. \square

Definition 12.10. Für jedes $\lambda > 0$ bezeichnen wir das eindeutige Minimum von F_λ auf S mit $X(\lambda)$, das Maximum von G_λ mit $y(\lambda)$, und die zugehörige komplementäre Matrix in (PD_λ) mit $Y(\lambda)$. Die Menge

$$C = \{(X(\lambda), Y(\lambda), y(\lambda)) \mid \lambda > 0\} \subset \text{Sym}_n \times \text{Sym}_n \times \mathbb{R}$$

heißt der *zentrale Pfad* des Paares (P),(D).

Bemerkung 12.11. Es ist eine bemerkenswerte Tatsache, dass der zentrale Pfad eine semi-algebraische Kurve ist. (Dies bedeutet im wesentlichen, dass der Zariski-Abschluss von C eine algebraische Kurve ist). Dies spielt für uns hier keine Rolle, die Geometrie dieser Kurve ist aber Gegenstand aktueller Forschung. (Siehe etwa deLoera, Sturmfels, Vinzant [2]: *The central curve in linear programming* (arXiv:1012.3978)).

Satz 12.12. *Es sei (P),(D) ein duales Paar von semidefiniten Programmen, das der Innere-Punkte-Bedingung genügt. Für jede Folge $(\lambda_k)_{k \in \mathbb{N}}$ von positiven Konstanten gibt es eine Teilfolge der Folge $(X(\lambda_k), Y(\lambda_k), y(\lambda_k))_{k \in \mathbb{N}}$ auf dem zentralen Pfad, die gegen ein optimales Tripel konvergiert, d.h. gegen eine Lösung des Systems (PD) in Prop. 12.4.*

Beweisskizze. Jedes Tripel $(X(\lambda_k), Y(\lambda_k), y(\lambda_k))$ ist eine Lösung des Systems (PD_λ) in Satz 12.9. Daher genügt es zu beweisen, dass die Folge $(X(\lambda_k), Y(\lambda_k), y(\lambda_k))$ beschränkt ist. Wir verweisen hierfür auf Forst/Hoffmann [3, Thm. 7.5.6]. Die Folge besitzt dann eine konvergente Teilfolge, so dass (X_k, Y_k) gegen Null und damit $(X(\lambda_k), Y(\lambda_k), y(\lambda_k))$ gegen eine Lösung von (PD) konvergiert. \square

Das folgende Korollar hätten wir mit etwas Mühe auch aus abstrakten Argumenten über Kegelprogrammierung erhalten können.

Korollar 12.13. *Falls das Paar (P),(D) der Innere-Punkte-Bedingung genügt, dann besitzen (P) und (D) optimale Punkte.* \square

Wir haben also gesehen, dass der zentrale Pfad auf einen optimalen Punkt zuläuft. Tatsächlich kann man auch noch zeigen, dass die Zielfunktion entlang des zentralen Pfades streng monoton wachsend ist. Aber im Hinblick auf tatsächliche Algorithmen fängt die eigentliche Arbeit natürlich jetzt erst an: Es gibt verschiedene numerische Verfahren, um den zentralen Pfad (stückweise linear) zu approximieren („path tracking“), etwa basierend auf dem Newton-Verfahren. Dies ist für sich genommen ein umfangreiches Thema, auf das wir nicht näher eingehen können.

12.3. EIN BEISPIEL MIT MATLAB/SEDUMI/YALMIP

Wir betrachten unser Beispiel aus Aufgabe 11.6 und berechnen die Lovasz-Zahl für den Graphen, der wie ein Fünfeck aussieht. Dazu geben wir das entsprechende Programm (direkt vor Bemerkung 11.7) in Matlab ein. Wir verwenden den SDP-solver SeDuMi und die Toolbox Yalmip (beide frei verfügbar).

```
>>J = [1,1,1,1,1;
       1,1,1,1,1;
       1,1,1,1,1;
       1,1,1,1,1;
       1,1,1,1,1];
>>A1 = [0,0,1,0,0;
        0,0,0,0,0;
        1,0,0,0,0;
        0,0,0,0,0;
        0,0,0,0,0];
>>A2 = [0,0,0,1,0;
        0,0,0,0,0;
        0,0,0,0,0;
        1,0,0,0,0;
        0,0,0,0,0];
>>A3 = [0,0,0,0,0;
        0,0,0,1,0;
        0,0,0,0,0;
        0,1,0,0,0;
        0,0,0,0,0];
>>A4 = [0,0,0,0,0;
        0,0,0,0,1;
        0,0,0,0,0;
        0,0,0,0,0;
        0,1,0,0,0];
>>A5 = [0,0,0,0,0;
        0,0,0,0,0;
        0,0,0,0,1;
        0,0,0,0,0;
        0,0,1,0,0];

%Definiere und loese das SDP in der Variablen X
>>X = sdpvar(5,5);
>>P = [X >= 0, trace(A1*X) == 0, trace(A2*X) == 0, trace(A3*X) == 0,
      trace(A4*X) == 0, trace(A5*X) == 0, trace(X) == 1 ];
>>h = -trace(J*X);
>>solvesdp(P,h);

SeDuMi 1.3 by AdvOL, 2005–2008 and Jos F. Sturm, 1998–2003.
Alg = 2: xz-corrector, theta = 0.250, beta = 0.500
Put 6 free variables in a quadratic cone
eqs m = 15, order n = 8, dim = 33, blocks = 3
nnz(A) = 25 + 0, nnz(ADA) = 225, nnz(L) = 120
it :      b*y      gap    delta rate  t/tP*  t/tD*   feas cg cg  prec
0 :              2.75E+00 0.000
1 :  4.88E+00 5.48E-01 0.000 0.1989 0.9000 0.9000  -0.13  1  1  1.7E+00
2 :  2.78E+00 1.47E-01 0.000 0.2680 0.9000 0.9000   1.57  1  1  3.5E-01
3 :  2.25E+00 6.31E-03 0.000 0.0430 0.9900 0.9900   1.48  1  1  1.1E-02
4 :  2.24E+00 2.45E-07 0.000 0.0000 1.0000 1.0000   1.01  1  1  4.1E-07
5 :  2.24E+00 2.85E-14 0.000 0.0000 1.0000 1.0000   1.00  1  1  4.8E-14

iter seconds digits      c*x      b*y
5         0.2  14.3  2.2360679775e+00  2.2360679775e+00
|Ax-b| = 9.5e-14, [Ay-c]_+ = 1.1E-14, |x|= 8.7e+00, |y|= 5.3e-01
```

Detailed timing (sec)

```
Pre      IPM      Post
1.808E-01  3.441E-01  4.832E-02
Max-norms: ||b||=2, ||c|| = 1,
Cholesky |add|=0, |skip| = 0, ||L.L|| = 5.23598.
```

In all dem Gewusel sehen wir unter der Zielfunktion (c^*x) tatsächlich eine gute Näherung der Zahl $\sqrt{5}$, wie in der Aufgabe behauptet.

Übung 12.4. Berechne das Lovasz-Theta numerisch für weitere Graphen Deiner Wahl.

13. SPEKTRAEDER UND STARRE KONVEXITÄT

13.1. SPEKTRAEDER ALS PUNKTMENGEN

Wie in Abschnitt 7 definiert, ist ein Spektraeder der Durchschnitt des psd-Kegels mit einem affin-linearen Unterraum im Raum der symmetrischen Matrizen. In diesem Abschnitt fassen wir Spektraeder nicht primär als Mengen von Matrizen, sondern als Punkt-mengen im euklidischen Raum auf. Dies macht technisch keinen großen Unterschied, in der Anschauung aber schon.

Es sei $L \subset \text{Sym}_k$ ein affin-linearer Unterraum. Wähle $A_0 \in L$ und schreibe $L - A_0 = \text{span}(A_1, \dots, A_n)$ mit $A_i \in \text{Sym}_k$ und $n \geq 0$. Betrachte die Abbildung

$$\varphi: \begin{cases} \mathbb{R}^n & \rightarrow L \\ (u_1, \dots, u_n)^T & \mapsto A_0 + u_1 A_1 + \dots + u_n A_n \end{cases}$$

Dann gilt

$$\varphi^{-1}(L \cap S_k^+) = \{u \in \mathbb{R}^n \mid A_0 + u_1 A_1 + \dots + u_n A_n \geq 0\}.$$

Wenn A_1, \dots, A_n linear unabhängig sind, dann ist φ bijektiv und wir können $\varphi^{-1}(L \cap S_k^+)$ mit dem Spektraeder $L \cap S_k^+$ identifizieren. Das führt zu einer leicht erweiterten Definition:

Definition 13.1. Eine Teilmenge $S \subset \mathbb{R}^n$ heißt ein *Spektraeder*, wenn es eine affin-lineare Abbildung $\varphi: \mathbb{R}^n \rightarrow \text{Sym}_k$ gibt derart, dass $S = \varphi^{-1}(\text{im}(\varphi) \cap S_k^+)$.

Übung 13.1. Es seien $A_0, \dots, A_n \in \text{Sym}_k$ und $S = \{u \in \mathbb{R}^n \mid A_0 + u_1 A_1 + \dots + u_n A_n \geq 0\}$. Seien $B_1, \dots, B_m \in \text{Sym}_k$ eine Basis von $\text{span}(A_1, \dots, A_n)$, $m \leq n$. Zeige: Es gibt eine bijektive lineare Abbildung $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^m \times \mathbb{R}^{n-m}$ mit $\psi(S) = \{(u, v) \in \mathbb{R}^m \times \mathbb{R}^{n-m} \mid A_0 + u_1 B_1 + \dots + u_m B_m \geq 0\}$.

Beispiel 13.2. Es sei $n = k = 2$ und

$$A_0 = I_2, \quad A_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Für $u = (u_1, u_2)^T \in \mathbb{R}^2$ ist dann

$$A_0 + u_1 A_1 + u_2 A_2 = \begin{bmatrix} 1 - u_1 & u_2 \\ u_2 & 1 + u_1 \end{bmatrix}.$$

Diese Matrix ist genau dann positiv semidefinit, wenn $-1 \leq u_1 \leq 1$ und $1 - u_1^2 - u_2^2 \geq 0$ gelten. Die zweite Bedingung (die von der Determinante kommt) impliziert dabei die erste. Wir sehen also, dass

$$S = \{u \in \mathbb{R}^2 \mid u_1^2 + u_2^2 \leq 1\}$$

die abgeschlossene Einheitskreisscheibe ist.

Übung 13.2. Zeige, dass jedes Polyeder auch ein Spektraeder ist.

Definition 13.3. Ein *lineares (symmetrisches) Matrixpolynom* der Größe k ist ein Polynom vom Grad 1 (in mehreren Veränderlichen) über dem Ring der symmetrischen $k \times k$ -Matrizen. Ist $\mathcal{M} = A_0 + x_1 A_1 + \dots + x_n A_n$ ein lineares Matrixpolynom in n Veränderlichen $x = (x_1, \dots, x_n)$, dann definiert \mathcal{M} das Spektraeder

$$\mathcal{S}(\mathcal{M}) = \{u \in \mathbb{R}^n \mid \mathcal{M}(u) \geq 0\}.$$

Das Matrixpolynom \mathcal{M} heißt *normiert*, wenn $\mathcal{M}(0) = I_k$ gilt.

Ein lineares Matrixpolynom kann man sich entweder als Polynom vom Grad 1 mit Matrixkoeffizienten vorstellen, oder auch als Matrix, deren Einträge Polynome vom Grad 1 sind, je nachdem, wie es gerade passt. In unserem Beispiel oben ist das Matrixpolynom

$$\mathcal{M} = A_0 + x_1 A_1 + x_2 A_2 = \begin{bmatrix} 1 - x_1 & x_2 \\ x_2 & 1 + x_1 \end{bmatrix}.$$

Bemerkung 13.4. Es sei \mathcal{M} ein lineares Matrixpolynom in n Veränderlichen und S das Spektraeder $\mathcal{S}(\mathcal{M})$. Für $u, v \in \mathbb{R}^n$ ist $v^T \mathcal{M}(u) v \geq 0$ eine lineare Ungleichung in u und

$$S = \{u \in \mathbb{R}^n \mid \text{Für alle } v \in \mathbb{R}^n: v^T \mathcal{M}(u) v \geq 0\}.$$

Also ist S durch unendlich viele lineare Ungleichungen beschrieben. Das ist noch nichts Besonderes: Nach dem Bidualsatz wird *jede* abgeschlossene konvexe Menge durch unendlich viele lineare Ungleichungen beschrieben (nämlich die Ungleichungen, die durch die Gesamtheit aller Stützhyperebenen gegeben sind). Bei Spektraedern sind diese Ungleichungen aber in spezieller Weise parametrisiert und können explizit angegeben werden.

Lemma 13.5. *Es sei \mathcal{M} ein lineares Matrixpolynom der Größe k in n Veränderlichen.*

- (1) *Falls \mathcal{M} normiert ist, so gilt $0 \in \text{int}(\mathcal{S}(\mathcal{M}))$.*
- (2) *Ist $0 \in \text{int}(\mathcal{S}(\mathcal{M}))$, so gibt es ein normiertes lineares Matrixpolynom \mathcal{N} der Größe $l \leq k$ in n Veränderlichen mit $\mathcal{S}(\mathcal{M}) = \mathcal{S}(\mathcal{N})$.*

Beweis. (1) folgt leicht daraus, dass I_k ein innerer Punkt von S_k^+ ist.

(2) Es sei $\mathcal{M} = A_0 + x_1 A_1 + \dots + x_n A_n$. Wegen $0 \in \mathcal{S}(\mathcal{M})$ gilt $A_0 \geq 0$. Deshalb gibt es eine invertierbare $k \times k$ -Matrix U derart, dass

$$U^T A_0 U = \begin{bmatrix} I_l & 0 \\ 0 & 0 \end{bmatrix}$$

mit $l \leq k$ gilt. Es sei nun

$$U^T \mathcal{M} U = U^T A_0 U + x_1 U^T A_1 U + \dots + x_n U^T A_n U = \begin{bmatrix} \mathcal{N} & \mathcal{A} \\ \mathcal{A}^T & \mathcal{N}' \end{bmatrix},$$

wobei \mathcal{N} bzw. \mathcal{N}' lineare Matrixpolynome der Größe l bzw. $k - l$ sind und \mathcal{A} ein nicht-symmetrisches lineares Matrixpolynom der Größe $l \times (k - l)$. Wir behaupten, dass $\mathcal{A} = 0$ und $\mathcal{N}' = 0$ gelten. Nach Wahl von U ist der konstante Term von \mathcal{N}' gleich Null, es ist also etwa $\mathcal{N}' = x_1 B_1 + \dots + x_n B_n$ mit $B_i \in \text{Sym}_{k-l}$. Da 0 ein innerer Punkt von $\mathcal{S}(\mathcal{M}) \subset \mathcal{S}(\mathcal{N}')$ ist, gibt es $\varepsilon > 0$ derart, dass $\pm \varepsilon B_i \geq 0$ für $i = 1, \dots, n$. Das geht nur, wenn $B_i = 0$ für $i = 1, \dots, n$ gilt. Ist nun W eine offene Umgebung von 0 mit $\mathcal{M}(x) \geq 0$ für alle $x \in W$, so folgt für $x \in W$ aus $\mathcal{N}'(x) = 0$ auch $\mathcal{A}(x) = 0$. Dies impliziert $\mathcal{A} = 0$.

Insgesamt haben wir $\mathcal{S}(\mathcal{M}) = \mathcal{S}(\mathcal{N})$ und damit die Behauptung gezeigt. \square

Lemma 13.6. *Für jedes normierte lineare Matrixpolynom \mathcal{M} in n Veränderlichen gilt*

$$\text{int}(\mathcal{S}(\mathcal{M})) = \{u \in \mathbb{R}^n \mid \mathcal{M}(u) > 0\}.$$

Beweis. Es sei $\mathcal{M} = I_k + x_1 A_1 + \dots + x_n A_n$, $L = I_k + \text{span}(A_1, \dots, A_n)$ und $\varphi: \mathbb{R}^n \rightarrow L$, $\varphi(u) = I_k + u_1 A_1 + \dots + u_n A_n$. Dann gilt $\text{int}(S) = \varphi^{-1}(\text{relint}(L \cap S_k^+)) = \varphi^{-1}(L \cap S_k^+)$ und damit die Behauptung. \square

Eine weitere wichtige Eigenschaft von Spektraedern steckt in der folgenden Aufgabe (deren Lösung allerdings nicht unbedingt etwas mit diesem Abschnitt zu tun hat).

Übung 13.3. Zeige, dass alle Seiten eines Spektraeders exponiert sind.

13.2. STARRE KONVEXITÄT

Wir haben gesehen, dass eine Vielzahl von konvexen Mengen (z.B. alle Polyeder) als Spektraeder realisiert werden können. Spektraeder besitzen aber auch einige sehr spezielle Eigenschaften, die sie von anderen konvexen Mengen unterscheiden (s. Übung 13.3). Die wichtigste ist die sogenannte starre Konvexität, die wir in diesem Abschnitt untersuchen.

Definition 13.7. Es sei $a \in \mathbb{R}^n$. Ein Polynom $p \in \mathbb{R}[x]$ in n Veränderlichen $x = (x_1, \dots, x_n)$ heißt ein *reelles Nullstellenpolynom* oder ein *RZ-Polynom* bezüglich a , wenn folgendes gilt: Für alle $v \in \mathbb{R}^n$ sind alle komplexen Nullstellen des Polynoms $p(a + tv) \in \mathbb{R}[t]$ reell.

Beispiele 13.8.

- (1) Das Polynom $p = 1 - x_1^2 - x_2^2$ ist ein RZ-Polynom bzgl. $(0, 0)$: Denn für jedes $v \in \mathbb{R}^d$ hat das quadratische Polynom $p(tv) = 1 - t^2 v_1^2 - t^2 v_2^2 = -(v_1^2 + v_2^2)t^2 + 1$ in der Veränderlichen t die beiden reellen Nullstellen $\pm \sqrt{(v_1^2 + v_2^2)^{-1}}$.
- (2) Das Polynom $p = 1 - x_1^4 - x_2^4$ ist kein RZ-Polynom bzgl. $(0, 0)$. Denn die Nullstellen von $p(tv)$ in t sind die vier vierten Wurzeln von $(v_1^4 + v_2^4)^{-1}$, von denen zwei reell und zwei nicht reell sind.

Proposition 13.9. Es sei \mathcal{M} ein normiertes lineares Matrixpolynom in n Veränderlichen $x = (x_1, \dots, x_n)$. Dann ist das Determinantenpolynom $\det(\mathcal{M}) \in \mathbb{R}[x]$ ein RZ-Polynom bzgl. 0 .

Beweis. Der Grund ist, dass die Eigenwerte einer symmetrischen Matrix alle reell sind: Schreibe $\mathcal{M} = I_k + \mathcal{N}$ mit $\mathcal{N} = x_1 A_1 + \dots + x_n A_n$, $p = \det(\mathcal{M}) \in \mathbb{R}[x]$ und $v \in \mathbb{R}^n$, dann gilt

$$p(tv) = \det(tv) = \det(I_k + \mathcal{N}(tv)) = \det(t(t^{-1}I_k + \mathcal{N}(v))) = t^k \det(sI_n + \mathcal{N}(v)),$$

wobei $s = t^{-1}$. Nun ist $\det(sI_k + \mathcal{N}(v))$ gerade das charakteristische Polynom der symmetrischen Matrix $-\mathcal{N}(v)$ und hat deshalb in der Variablen s nur reelle Nullstellen. Damit sind auch die Nullstellen von $p(tv)$ in der Variablen $t = s^{-1}$ alle reell. \square

Übung 13.4. Es sei $p = x_1^3 - x_1^2 - x_1 - x_2^2 + 1$. Finde ein normiertes lineares Matrixpolynom \mathcal{M} in x_1, x_2 mit $p = \det(\mathcal{M})$.

Definition 13.10. Es sei $a \in \mathbb{R}^n$ und p ein RZ-Polynom bzgl. a mit $p(a) > 0$. Sei A die Zusammenhangskomponente der Menge $\{u \in \mathbb{R}^n \mid p(u) \neq 0\}$, die den Punkt a enthält. Der Abschluss $\text{clos}(A)$ von A heißt das durch p definierte *starr konvexe Gebiet* und wird mit $\mathcal{C}_p(a)$ bezeichnet.

Diese Definition legt nahe, dass starr konvexe Gebiete konvex sind. Das werden wir später beweisen.

Übung 13.5. Es sei $p \in \mathbb{R}[x, y]$ ein RZ-Polynom bzgl. 0 in zwei Veränderlichen mit $p(0, 0) = 1$. Zeige durch ein Beispiel, dass $\mathcal{C}_p(a)$ im allgemeinen keine Zusammenhangskomponente der Menge $\{u \in \mathbb{R}^2 \mid p(u) \geq 0\}$ sein muss.

Proposition 13.11. *Jedes Spektraeder mit nicht-leerem Inneren ist starr konvex.*

Beweis. Es sei S ein Spektraeder mit nicht-leerem Inneren in \mathbb{R}^n . Nach Verschieben können wir annehmen, dass $0 \in \text{int}(S)$. Dann gibt es nach Lemma 13.5 ein normiertes lineares Matrixpolynom \mathcal{M} mit $S = \{u \in \mathbb{R}^n \mid \mathcal{M}(u) \geq 0\}$. Nach Prop. 13.9 ist dann $p = \det(\mathcal{M})$ ein RZ-Polynom bzgl. 0. Dabei ist S das durch p definierte starr konvexe Gebiet: Da p nach Lemma 13.6 in $\text{int}(S)$ nirgends verschwindet, gilt $S \subset \mathcal{C}_p(a)$. Andererseits kann ein Punkt in $v \in \mathbb{R}^n \setminus S$ nicht im Abschluss der Zusammenhangskomponente von 0 in $\{u \in \mathbb{R}^n \mid p(u) \neq 0\}$ liegen. Denn $\mathcal{M}(v)$ hat einen negativen Eigenwert und die Eigenwerte hängen stetig von v ab. Deshalb enthält jeder Weg von 0 nach v einen Punkt, in dem p verschwindet. Daraus folgt $v \notin \mathcal{C}_p(a)$ und insgesamt $\mathcal{C}_p(a) \subset S$. \square

Bemerkung 13.12. Aus der Tatsache, dass $p = 1 - x_1^4 - x_2^4$ kein RZ-Polynom bezüglich 0 ist, folgt, dass die Menge $C = \{u \in \mathbb{R}^d \mid u_1^4 + u_2^4 \leq 1\}$ kein Spektraeder ist. (Um dies sauber zu beweisen, muss man sich folgendes überlegen: Wäre \mathcal{M} ein lineares Matrixpolynom mit $C = \mathcal{S}(\mathcal{M})$, so wäre p ein Faktor von $\det(\mathcal{M})$, also $\det(\mathcal{M})$ kein RZ-Polynom bzgl. 0).

Es ist eine offene Frage der aktuellen Forschung, ob die Umkehrung von Prop. 13.11 gilt:

Vermutung 13.13 (Helton-Vinnikov). *Jede starr konvexe Menge ist ein Spektraeder.*

Die Vermutung geht auf den Beweis für den Fall $n = 2$ zurück:

Satz 13.14 (Helton-Vinnikov 2004 [7]). *Jede ebene starr konvexe Menge ist ein Spektraeder.*

Der Beweis dieses Satzes erfordert ziemlich tiefe Resultate der algebraischen Geometrie und sprengt den Rahmen dieser Vorlesung bei weitem.

14. SPEKTRAEDER II

14.1. STARRE KONVEXITÄT (FORTSETZUNG)

Dass starr konvexe Mengen konvex sind, ist aus der Definition nicht direkt ersichtlich. Wir geben hier den Beweis nach Renegar [13]. Zu einem Polynom $p \in \mathbb{R}[x]$ in n Veränderlichen $x = (x_1, \dots, x_n)$ betrachten wir die *Homogenisierung* von p bezüglich einer zusätzlichen Variable x_0 : Ist $d = \deg(p)$ und $p = \sum_{i_1+\dots+i_n \leq d} a_{i_1 \dots i_n} x_1^{i_1} \dots x_n^{i_n}$, so ist die Homogenisierung P von p bzgl. x_0 gegeben durch $P = \sum_{i_0+\dots+i_n=d} a_{i_1 \dots i_n} x_0^{i_0} x_1^{i_1} \dots x_n^{i_n} = x_0^d p(x_0^{-1}x)$. Alle Terme von P haben also dann den gleichen Grad d , und es gilt $P(\lambda v) = \lambda^d P(v)$ für alle $v \in \mathbb{R}^{n+1}$ und $\lambda \in \mathbb{R}$. Außerdem erhält man p aus P durch $p = P(1, x_1, \dots, x_n)$ zurück.

Beispiel 14.1. Es sei $p = x_2^2 - x_1^3 - x_1^2 + 1$. Dann ist $P = x_0 x_2^2 - x_1^3 - x_0 x_1^2 + x_0^3$.

Es sei $p \in \mathbb{R}[x]$ ein RZ-Polynom bezüglich $a \in \mathbb{R}^n$ mit $p(a) > 0$. Ergänzend zur Notation im letzten Abschnitt schreiben wir $C_p^\circ(a)$ für die Zusammenhangskomponente von $\{u \in \mathbb{R}^n \mid p(u) \neq 0\}$, die a enthält. Nach Definition gilt dann $C_p(a) = \text{clos}(C_p^\circ(a))$, wobei $C_p(a)$ die durch p definierte starr konvexe Menge ist.

Notation 14.2. Ist $u \in \mathbb{R}^n$, so schreiben wir im folgenden \bar{u} für den Vektor $(1, u) \in \mathbb{R}^{n+1}$.

Lemma 14.3. *Es sei $p \in \mathbb{R}[x]$ ein Polynom in n Veränderlichen und $a \in \mathbb{R}^n$ mit $p(a) > 0$. Sei P die Homogenisierung von p .*

(1) *Die folgenden Aussagen sind äquivalent:*

- (a) *Für alle $u \in \mathbb{R}^n$ hat $p(a + tu) \in \mathbb{R}[t]$ nur reelle Nullstellen (RZ-Eigenschaft).*
- (b) *Für alle $v \in \mathbb{R}^{n+1}$ hat $P(t\bar{a} + v) \in \mathbb{R}[t]$ nur reelle Nullstellen.¹*

(2) *Ist p ein RZ-Polynom bzgl. a , so gilt*

$$C_p^\circ(a) = \{u \in \mathbb{R}^n \mid \text{alle Nullstellen von } P(t\bar{a} + \bar{u}) \text{ sind negativ}\}.$$

Beweis. (1) Sei p ein RZ-Polynom bzgl. a und sei $v = (v_0, u) \in \mathbb{R}^{n+1}$. Sei τ eine Nullstelle von $P(t\bar{a} + v)$ in t . Falls $\tau = -v_0$, so ist τ reell, andernfalls schreiben wir

$$0 = P(\tau\bar{a} + v) = (\tau + v_0)^d p((\tau + v_0)^{-1}(\tau a + u)) = p(a + \sigma b)$$

wobei $b = v_0 a + u$ und $\sigma = (\tau + v_0)^{-1}$. Es folgt, dass σ reell ist und damit auch τ .

Umgekehrt habe $P(t\bar{a} + v) \in \mathbb{R}[t]$ für alle $v \in \mathbb{R}^{n+1}$ nur reelle Nullstellen. Sei $u \in \mathbb{R}^n$, dann

$$p(a + tu) = P(1, a + tu) = P(\bar{a} + t(0, u)) = t^d P(t^{-1}\bar{a} + (0, u)) = s^{-d} P(s\bar{a} + (0, u))$$

mit $s = t^{-1}$ und es folgt, dass die linke Seite nur reelle Nullstellen in t hat.

(2) Sei U die Menge auf der rechten Seite. Für $v = \bar{a}$ hat $P(t\bar{a} + \bar{a}) = P((t+1)\bar{a}) = (t+1)^d P(\bar{a})$ nur die negative Nullstelle $t = -1$. Also gilt $a \in U$. Da die Nullstellen von

¹Homogene Polynome, die dieser Bedingung genügen, nennt man *hyperbolisch* oder *reell stabil*.

$P(t\bar{a} + v)$ stetig von $v \in \mathbb{R}^{n+1}$ abhängen und $P(v) > 0$ für $v \in C_p^\circ(a)$ gilt, folgt $C_p^\circ(a) \subset U$. Für die umgekehrte Inklusion, sei $b \in U$. Wir konstruieren einen Weg von a nach b , auf dem p strikt positiv ist. Sei dazu $[b, a] = \{\lambda b + (1 - \lambda)a \mid \lambda \in [0, 1]\}$. Wegen $P(\bar{a}) > 0$ gibt es $\tau > 0$ derart, dass $P(\tau\bar{a} + \bar{u}) > 0$ für alle $u \in [b, a]$ gilt. Ferner gilt $P(t\bar{a} + \bar{b}) > 0$ für alle $t \geq 0$. Denn sonst gäbe es $t \geq 0$ mit $P(t\bar{a} + \bar{b}) = 0$ im Widerspruch zu $\bar{b} \in U$. Außerdem gilt $P(\bar{a} + t\bar{a}) > 0$ für alle $t > 0$ (s.o.). Also ist p auf den drei Geradensegmenten

$$\Gamma_1 = \{b + ta \mid 0 \leq t \leq \tau\} \quad \Gamma_2 = \{v + \tau a \mid v \in [b, a]\} \quad \Gamma_3 = \{a + ta \mid \tau \geq t \geq 0\}$$

positiv. Man sieht nun, dass $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ ein stetiger Weg ist, der b und a verbindet und in $C_p^\circ(a)$ verläuft. Also folgt $b \in C_p^\circ(a)$. \square

Satz 14.4 (Gårding). *Es sei $p \in \mathbb{R}[x]$ ein RZ-Polynom bzgl. $a \in \mathbb{R}^n$ mit $p(a) > 0$. Für jedes $b \in C_p^\circ(a)$ ist p auch ein RZ-Polynom bezüglich b .*

Beweis. Wir zeigen zunächst die folgende Aussage:

(*) Für alle $\alpha > 0$, $s \geq 0$ und $v \in \mathbb{R}^{n+1}$ hat jede komplexe Nullstelle in t des Polynoms $P(\alpha i\bar{a} + t\bar{b} + sv)$ negativen Imaginärteil.

Dabei ist $i = \sqrt{-1}$. Für $s = 0$ steht dort $P(\alpha i\bar{a} + t\bar{b}) = t^d P(\alpha i t^{-1}a + \bar{b})$. Nach Lemma 14.3(2) hat dieses Polynom in der Variablen $\alpha i t^{-1}$ nur negative reelle Nullstellen und damit in t nur negativ-imaginäre Nullstellen.

Gäbe es nun $s > 0$ derart, dass $P(\alpha i\bar{a} + t\bar{b} + sv)$ eine Nullstelle mit nicht-negativem Imaginärteil hätte, dann gäbe es aus Stetigkeitsgründen ein s' mit $0 < s' \leq s$ derart, dass $P(\alpha i\bar{a} + t\bar{b} + s'\nu)$ eine reelle Nullstelle τ hat. Setze $w = \tau\bar{b} + s'\nu$, dann ist also $w \in \mathbb{R}^{n+1}$ mit $P(\alpha i\bar{a} + w) = 0$. Es folgt, dass αi eine nicht-reelle Nullstelle des Polynoms $P(t\bar{a} + w)$ ist, ein Widerspruch zu Lemma 14.3(1).

Nach der obigen Hilfsaussage mit $s = 1$ haben alle Nullstellen von $P(\alpha i\bar{a} + t\bar{b} + v)$ negativen Imaginärteil. Da die Nullstellen stetig von $\alpha > 0$ abhängen, sehen wir für $\alpha \rightarrow 0$, dass alle Nullstellen von $P(t\bar{b} + v)$ nicht-positiven Imaginärteil (≤ 0) haben. Aber $P(t\bar{b} + v)$ ist ein reelles Polynom, so dass für jede Nullstelle auch ihr komplex-Konjugiertes eine Nullstelle ist. Gäbe es also eine nicht-reelle Nullstelle, dann auch eine mit positivem Imaginärteil. Es folgt, dass alle Nullstellen reell sind. Nach 14.3(1) ist p ein RZ-Polynom bezüglich b . \square

Korollar 14.5. *Jede starr konvexe Menge ist konvex.*

Beweis. Es sei p ein RZ-Polynom bezüglich $a \in \mathbb{R}^n$ mit $p(a) > 0$. Es genügt zu zeigen, dass $C_p^\circ(a)$ konvex ist. Seien dazu $b, c \in C_p^\circ(a)$ und $\lambda \in [0, 1]$. Wir müssen zeigen, dass $\lambda b + (1 - \lambda)c \in C_p^\circ(a)$ gilt. Nach dem vorangehenden Satz dürfen wir $b = a$ annehmen. Wir verwenden Lemma 14.3(2): Sei P die Homogenisierung von p , dann gilt

$$P(t\bar{a} + \lambda\bar{a} + (1 - \lambda)\bar{c}) = P((t + \lambda)\bar{a} + (1 - \lambda)\bar{c}) = (1 - \lambda)^d P\left(\frac{t + \lambda}{1 - \lambda}\bar{a} + \bar{c}\right).$$

Wegen $\bar{c} \in C_p^\circ(a)$ gilt für jede Nullstelle τ der linken Seite: $\tau + \lambda < 0$, also $\tau < -\lambda < 0$. Daraus folgt $\lambda a + (1 - \lambda)c \in C_p^\circ(a)$. \square

Bemerkung 14.6. Ein RZ-Polynom p bzgl. a hat keine Nullstellen in $C_p^\circ(a)$.

Übung 14.1. Es sei $p \in \mathbb{R}[x, y]$ ein Polynom in zwei Veränderlichen mit $p(0) \neq 0$ und sei A die Zusammenhangskomponente von $\{u \in \mathbb{R}^2 \mid p(u) \neq 0\}$, die 0 enthält. Gib ein Beispiel für ein solches p derart, dass $\text{clos}(A)$ konvex ist, p jedoch in $\text{int}(\text{clos}(A))$ eine Nullstelle besitzt.

14.2. SPEKTRALE SCHATTEN

In diesem Abschnitt untersuchen wir die linearen Bilder von Spektraedern.

Definition 14.7. Eine Teilmenge $R \subset \mathbb{R}^n$ heißt ein *spektraler Schatten*, wenn es ein Spektraeder $S \subset \mathbb{R}^m$ und eine lineare Abbildung $\varphi: \mathbb{R}^m \rightarrow \mathbb{R}^n$ mit $R = \varphi(S)$ gibt².

Übung 14.2. Seien R, S, φ wie in der Definition. Zeige: Ist φ injektiv, so ist R ein Spektraeder.

Korollar 14.8. Eine Teilmenge $R \subset \mathbb{R}^n$ ist genau dann ein spektraler Schatten, wenn es $m \geq n$ und ein Spektraeder $S \subset \mathbb{R}^n \times \mathbb{R}^m$ gibt derart, dass R die Projektion von S auf die ersten n Komponenten ist.

Beweis. Es sei $S' \subset \mathbb{R}^m$ ein Spektraeder und $R = \varphi(S')$. Sei $\psi: \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m$, $\psi(x) = (\varphi(x), x)$ die Graphenabbildung von φ . Dann ist ψ injektiv, so dass $S = \psi(S')$ ein Spektraeder ist (Übung 14.2). Nun ist R die Projektion von S auf die ersten n Komponenten. \square

Etwas expliziter ist ein spektraler Schatten also eine Menge der Form

$$R = \{u \in \mathbb{R}^n \mid \text{Es gibt } v \in \mathbb{R}^m \text{ mit } \mathcal{M}(u, v) \geq 0\},$$

gegeben durch ein lineares Matrixpolynom \mathcal{M} in $n + m$ Veränderlichen. Eine solche Darstellung, die durch das Einführen zusätzlicher Variablen möglich gemacht wird, wird häufig als *geliftete Darstellung* bezeichnet. Ist S das Spektraeder in \mathbb{R}^{n+m} und $\text{pr}: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ die Projektion mit $R = \text{pr}(S)$, so kann man eine lineare Zielfunktion über R optimieren, indem man sie durch ein semidefinites Programm über S optimiert. Außerdem sind spektrale Schatten abgeschlossen unter den üblichen Operationen, die Konvexität erhalten:

- Endliche Durchschnitte (Das stimmt auch für Spektraeder; s.u.)
- Konvexe Dualität (s.u.)
- Topologischer Abschluss (s.u.) und Inneres (Netzer [9])
- Minkowski-Summen und endliche kartesische Produkte
- Konvexe Hüllen von endlichen Vereinigungen (Helton-Nie [5], Netzer-Sinn [11])

Einige dieser Eigenschaften werden wir im folgenden beweisen.

Übung 14.3. Zeige:

- (1) Der Durchschnitt zweier Spektraeder ist ein Spektraeder.
- (2) Der Durchschnitt zweier spektraler Schatten ist ein spektraler Schatten.

Übung 14.4. Zeige, dass der TV-Bildschirm $\{u \in \mathbb{R}^2 \mid 1 - u_1^4 - u_2^4 \geq 0\}$ ein spektraler Schatten ist. (*Hinweis:* Verwende, dass die Einheitskreisscheibe ein Spektraeder ist.)

Dieses Beispiel zeigt insbesondere, dass spektrale Schatten nicht starr konvex sein müssen.

Proposition 14.9 (Nemirovski [8, §4.1.1]). *Das Dual eines spektralen Schattens ist wieder ein spektraler Schatten.*

Beweis. Es sei $R = \{u \in \mathbb{R}^n \mid \exists v \in \mathbb{R}^m: \mathcal{M}(u, v) \geq 0\}$ ein spektraler Schatten gegeben durch ein lineares Matrixpolynom \mathcal{M} . Wir können annehmen, dass \mathcal{M} normiert ist. (Indem wir \mathbb{R}^{n+m} ggf. durch die affine Hülle des Spektraeders $\{(u, v) \in \mathbb{R}^{n+m} \mid \mathcal{M}(u, v) \geq 0\}$ ersetzen, können wir annehmen, dass dieses nicht-leeres Inneres besitzt und damit nach Lemma 13.5 durch ein normiertes lineares Matrixpolynom beschrieben wird.) Wir schreiben

$$\mathcal{M}(x) = I_k + \sum_{i=1}^{n+m} x_i A_i$$

²In der Literatur hat sich noch kein einheitlicher Name etabliert. Es kursieren: Projections of spectrahedra, spectrahedral shadows, semidefinitely representable sets, lifted-LMI representable sets, etc.

in $n + m$ Variablen $x = (x_1, \dots, x_{n+m})$. Für $w \in \mathbb{R}^n$, betrachte das folgende duale Paar von semidefiniten Programmen:

$$\begin{array}{l}
 \text{(P)} \quad \left. \begin{array}{l}
 \text{Finde} \quad \inf \langle I_k, X \rangle = \text{tr}(X) \\
 \text{unter} \quad \langle A_i, X \rangle = w_i \text{ für } i = 1, \dots, n \\
 \quad \quad \langle A_i, X \rangle = 0 \text{ für } i = n + 1, \dots, n + m \\
 \quad \quad X \geq 0
 \end{array} \right\} \text{ in der Variablen } X \in \text{Sym}_k. \\
 \\
 \text{(D)} \quad \left. \begin{array}{l}
 \text{Finde} \quad \sup(\langle w, x \rangle + \langle 0, y \rangle) \\
 \text{unter} \quad \sum_{i=1}^n x_i A_i + \sum_{i=n+1}^{n+m} y_i A_i \leq I_k
 \end{array} \right\} \text{ in der Variablen } (x, y) \in \mathbb{R}^{n+m}.
 \end{array}$$

Sobald (P) einen zulässigen Punkt besitzt, gibt es nach Satz 11.3 keine Dualitätslücke, da I_k positiv definit ist. Damit gilt nun

$$\begin{aligned}
 R^\circ &= \{w \in \mathbb{R}^n \mid \langle w, x \rangle \leq 1 \text{ für alle } x \in R\} \\
 &= \{w \in \mathbb{R}^n \mid \sup\{\langle w, u \rangle \mid (u, v) \in \mathbb{R}^{n+m} \text{ mit } \mathcal{M}(u, v) \geq 0\} \leq 1\} \\
 &= \{w \in \mathbb{R}^n \mid \inf\{\text{tr}(X) \mid \langle A_i, X \rangle = w_i \text{ für } i \leq n, \langle A_i, X \rangle = 0 \text{ für } i > n \text{ und } X \geq 0\} \leq 1\} \\
 &= \text{pr}_w(R')
 \end{aligned}$$

wobei

$$R' = \{(w, X) \in \mathbb{R}^n \times \text{Sym}_k \mid \langle A_i, X \rangle = w_i \text{ für } i \leq n, \langle A_i, X \rangle = 0 \text{ für } i > n, \text{tr}(X) \leq 1 \text{ und } X \geq 0\}.$$

Wir zeigen, dass R' und damit auch R ein spektraler Schatten ist. Sei dazu

$$L = \left\{ X' = \begin{bmatrix} X & b \\ b^T & c \end{bmatrix} \in \text{Sym}_{k+1} \mid \text{tr}(X') = 1 \text{ und } \langle A_i, X \rangle = 0 \text{ für } i = n + 1, \dots, n + m \right\}$$

und sei $\varphi: \text{Sym}_{k+1} \rightarrow \mathbb{R}^n \times \text{Sym}_k$ die lineare Abbildung $\varphi(X') = ((\langle A_1, X \rangle, \dots, \langle A_n, X \rangle), X)$, wobei wir $X' \in \text{Sym}_{k+1}$ wie oben als Blockmatrix mit linkem oberem $k \times k$ -Block X schreiben. Dann gilt $R' = \varphi(L \cap S_{k+1}^+)$, so dass R' ein spektraler Schatten ist. \square

Korollar 14.10. *Der Abschluss eines spektralen Schattens ist ein spektraler Schatten.*

Beweis. Sei $S \subset \mathbb{R}^n$ ein spektraler Schatten. Durch Verschieben können wir $0 \in S$ erreichen. Nach dem Bidualsatz 8.4 gilt dann $\text{clos}(S) = (S^\circ)^\circ$ und die Behauptung folgt aus 14.9. \square

Übung 14.5. Gib ein Beispiel eines nicht-abgeschlossenen spektralen Schattens.

Tatsächlich sind keine notwendigen Bedingungen für die Darstellbarkeit einer Teilmenge von \mathbb{R}^n als spektraler Schatten bekannt außer den offensichtlichen: Spektrale Schatten sind konvex und semialgebraisch (d.h. in einem bestimmten Sinn durch polynomiale Ungleichungen beschrieben).

Vermutung 14.11. *Jede konvexe semialgebraische Menge ist ein spektraler Schatten.*

Diese Vermutung stammt von Helton und Nie und wird durch eine Reihe von Resultaten aus den letzten zehn Jahren nahegelegt (vor allem die Sätze von Helton und Nie selbst in [5], [6]), auf die wir im nächsten Abschnitt noch zu sprechen kommen werden.

15. DIE LASSERRE-RELAXIERUNG

In diesem Abschnitt beschreiben wir in groben Umrissen ein allgemeines Approximationsverfahren für konvexe semialgebraische Mengen durch spektrale Schatten.

Definition 15.1. Eine Teilmenge $S \subset \mathbb{R}^n$ heißt *basisch abgeschlossen* (genauer oft basisch-abgeschlossen semialgebraisch), wenn es Polynome $h_1, \dots, h_r \in \mathbb{R}[x]$ in n Veränderlichen $x = (x_1, \dots, x_n)$ gibt derart, dass

$$S = \{u \in \mathbb{R}^n \mid h_1(u) \geq 0, \dots, h_r(u) \geq 0\}.$$

Eine abgeschlossene semialgebraische Menge ist eine endliche Vereinigung von basisch abgeschlossenen Mengen.¹ Wir betrachten aber nur basisch abgeschlossene Mengen.

Beispiele 15.2.

- (1) Viele natürliche Beispiele abgeschlossener konvexer Mengen sind per Definition basisch abgeschlossen (etwa Polyeder, Kugeln, Ellipsoide...).
- (2) Jedes Spektraeder ist basisch abgeschlossen. Denn ist $S = \mathcal{S}(\mathcal{M})$ ein Spektraeder, gegeben durch ein lineares Matrixpolynom \mathcal{M} , dann ist S die basisch abgeschlossene Menge die durch alle symmetrischen Minoren von \mathcal{M} beschrieben wird.
- (3) Jede starr konvexe Menge ist basisch abgeschlossen. Das ist nicht ganz leicht zu sehen, folgt aber etwa aus den Ergebnissen von Renegar [13].

Übung 15.1. Sei $S_1 = \{u \in \mathbb{R}^2 \mid u_1^2 + u_2^2 \leq 1\}$ und $S_2 = [0, 1] \times [-1, 1]$. Dann sind S_1 und S_2 basisch abgeschlossen. Zeige, dass jedoch $S_1 \cup S_2$ nicht basisch abgeschlossen ist. (Um diese Aufgabe exakt zu lösen, braucht man ein wenig reelle algebraische Geometrie).

Die Lasserre-Relaxierung, die wir jetzt vorstellen, stellt ein Verfahren da, um die konvexe Hülle einer basisch abgeschlossenen Menge durch einen spektralen Schatten zu approximieren. Sie geht auf Ideen von Lasserre aus der konvexen Optimierung zurück. Unsere Darstellung stützt sich wesentlich auf [14].

Notation 15.3. Es seien $h_1, \dots, h_r \in \mathbb{R}[x]$.

- (1) Wir schreiben

$$\Sigma = \{g_1^2 + \dots + g_k^2 \mid k \geq 1, g_1, \dots, g_k \in \mathbb{R}[x]\}$$

für den konvexen Kegel der *Quadratsummen* in $\mathbb{R}[x]$. Ist $\mathbb{R}[x]^2$ die Menge aller Quadrate in $\mathbb{R}[x]$, so gilt also $\Sigma = \text{cone}(\mathbb{R}[x]^2)$.

- (2) Es sei

$$M = \{s_0 + s_1 h_1 + \dots + s_r h_r \mid s_0, \dots, s_r \in \Sigma\}$$

der von h_1, \dots, h_r erzeugte *quadratische Modul*².

¹Das ist nicht die Definition, sondern eine Folgerung aus dem sogenannten Endlichkeitssatz; siehe etwa Thm. 2.4.1 im Buch von Prestel und Delzell [12].

²Ein quadratischer Modul ist ein „Modul“ über Σ , kein Modul über einem Ring wie sonst üblich.

Zur Vereinfachung der Schreibweise setzen wir $h_0 = 1$. Für jedes $d \geq 0$ sei nun

$$M_d = \left\{ \sum_{i=0}^r s_i h_i \mid \deg(s_i h_i) \leq d \text{ und } s_i \in \Sigma \text{ für } i = 0, \dots, r \right\}.$$

Es ist also $M_d = \text{cone}(g^2 h_i \mid g \in \mathbb{R}[x] \text{ mit } 2 \deg(g) + \deg(h_i) \leq d, i = 0, \dots, r)$.

(3) Für $d \geq 0$ sei $\mathbb{R}[x]_d$ der endlich-dimensionale Unterraum der Polynome vom Grad höchstens d in $\mathbb{R}[x]$. Es sei $\mathbb{R}[x]_d^* = \{\ell \mid \ell: \mathbb{R}[x]_d \rightarrow \mathbb{R} \text{ linear}\}$ sein Dualraum.

(4) Schließlich schreiben wir

$$M_d^* = \left\{ \ell \in \mathbb{R}[x]_d^* \mid \ell(f) \geq 0 \text{ für alle } f \in M_d \text{ und } \ell(1) = 1 \right\}.$$

Der Witz an der Definition von M ist, dass alle Elemente von M auf der Menge S nur nicht-negative Werte annehmen, d.h. es gilt $f(u) \geq 0$ für alle $f \in M$ und $u \in S$.

Bemerkung 15.4. Ein wichtiger Punkt an der Definition von M_d ist, dass die Inklusion $M_d \subset M \cap \mathbb{R}[x]_d$ i.a. strikt ist. Das liegt daran, dass ein Polynom vom Grad d auch als Summe von Polynomen von höherem Grad als d dargestellt werden kann. Sei z.B. $r = n = 1$ und $h_1 = 1 - x^2$. Dann ist $x + 2 = (x + 1)^2 + h_1$ ein Element von $M \cap \mathbb{R}[x]_1$ aber nicht in $M_1 = \mathbb{R}$. Es kann sogar passieren, dass $M \cap \mathbb{R}[x]_d$ in gar keinem M_e für $e \geq d$ enthalten ist.

Übung 15.2. Zeige: Sind $g_1, \dots, g_k \in \mathbb{R}[x]$, so gilt $\deg(g_1^2 + \dots + g_k^2) = 2 \max\{\deg(g_i) \mid i = 1, \dots, k\}$. Im Fall $r = 0$, also $M = \Sigma$, kann das Phänomen der „Gradkürzungen“ wie in der vorangehenden Bemerkung also nicht auftreten.

Das zentrale Lemma besteht in der folgenden Beobachtung:

Lemma 15.5. Für jedes $d \geq 0$ ist die Teilmenge M_d^* von $\mathbb{R}[x]_d^*$ ein Spektraeder.

Beweis. Es sei $d \geq 0$ und schreibe

$$V = \left\{ (p_0, \dots, p_r) \in \mathbb{R}[x]^{r+1} \mid 2 \deg(p_i) + \deg(h_i) \leq d \right\}.$$

Dies ist ein endlich-dimensionaler Vektorraum. Für $\ell \in \mathbb{R}[x]_d^*$ setze

$$\Phi_\ell: \begin{cases} V \times V & \rightarrow \mathbb{R} \\ ((p_0, \dots, p_r), (q_0, \dots, q_r)) & \mapsto \sum_{i=0}^r \ell(p_i q_i h_i) \end{cases}$$

eine symmetrische Bilinearform auf dem endlich-dimensionalen Vektorraum V . Dadurch ist eine lineare Abbildung

$$\Phi: \begin{cases} \mathbb{R}[x]_d^* & \rightarrow \text{Sym}(V) \\ \ell & \mapsto \Phi_\ell \end{cases}$$

gegeben, wobei $\text{Sym}(V)$ der Vektorraum der symmetrischen Bilinearformen auf V ist. Behaupte nun, dass $\ell \in \mathbb{R}[x]_d^*$ mit $\ell(1) = 1$ genau dann in M_d^* liegt, wenn Φ_ℓ positiv semidefinit ist. Denn ist $\ell \in M_d^*$, so gilt $\Phi_\ell(p, p) = \sum_{i=0}^r \ell(p_i^2 h_i) \geq 0$ für alle $p = (p_0, \dots, p_r) \in V$, da jeder Summand in M_d liegt. Ist umgekehrt Φ_ℓ positiv semidefinit und $g \in \mathbb{R}[x]$ mit $g^2 h_i \in \mathbb{R}[x]_d$, so gilt $\ell(g^2 h_i) = \Phi_\ell((0, \dots, g, \dots, 0), (0, \dots, g, \dots, 0)) \geq 0$, also $\ell \in M_d^*$.

Damit ist $M_d^* = \Phi^{-1}(\text{Sym}^+(V)) \cap \{\ell \in \mathbb{R}[x]_d^* \mid \ell(1) = 1\}$, wobei $\text{Sym}^+(V)$ die positiv semidefiniten Bilinearformen auf V bezeichnet. Wir erhalten eine Beschreibung durch Matrizen, wie in unserer Definition von Spektraeder, durch Wahl einer Basis der Länge $N = \dim(V)$ von V , so dass $\text{Sym}(V) \cong \text{Sym}_N$ und $\text{Sym}^+(V) \cong S_N^+$. \square

Bemerkung 15.6. Die Abbildung Φ im Beweis ist für gerades d injektiv.

Proposition 15.7. *Es sei S eine basisch abgeschlossene Menge beschrieben durch h_1, \dots, h_r und sei $M_d^* \subset \mathbb{R}[x]_d^*$ für $d \geq 0$ wie oben. Es sei*

$$\pi: \begin{cases} \mathbb{R}[x]_d^* & \rightarrow & \mathbb{R}^n \\ \ell & \mapsto & (\ell(x_1), \dots, \ell(x_n)). \end{cases}$$

Dann gilt $\text{conv}(S) \subset \pi(M_d^*)$.

Beweis. Sei $u \in S$ und sei

$$\ell_u: \begin{cases} \mathbb{R}[x]_d & \rightarrow & \mathbb{R} \\ f & \mapsto & f(u). \end{cases}$$

Dann ist $u = \pi(\ell_u)$ und da alle Polynome in M_d auf S nicht-negativ sind, gilt $\ell_u \in M_d^*$. \square

Definition 15.8. Die Menge $\pi(M_d^*)$ in der obigen Proposition heißt die *Lasserre-Relaxierung vom Grad d* von S bezüglich h_1, \dots, h_r und wird mit L_d bezeichnet.

Für $d_1 \leq d_2$ gilt $M_{d_1} \subset M_{d_2}$ und deshalb $M_{d_2}^* \subset M_{d_1}^*$ und damit $L_{d_2} \subset L_{d_1}$. Die Folge der Lasserre-Relaxierungen vom Grad d für wachsendes d stellt also eine geschachtelte Folge von Obermengen von $\text{conv}(S)$ dar.

Definition 15.9. Es seien $h_1, \dots, h_r \in \mathbb{R}[x]$ und $S \subset \mathbb{R}^n$ wie zuvor.

- (1) Die Lasserre-Relaxierung L_{d_0} von S vom Grad d_0 bezüglich h_1, \dots, h_r heißt *exakt*, wenn $\text{clos}(L_d) = \text{clos}(\text{conv}(S))$ gilt.
- (2) Die Folge der Lasserre-Relaxierungen $(L_d)_{d \geq 0}$ von S bezüglich h_1, \dots, h_r heißt *konvergent*, wenn gilt:

$$\text{clos}(\text{conv}(S)) = \bigcap_{d \geq 0} L_d.$$

Proposition 15.10. *Es seien $h_1, \dots, h_r \in \mathbb{R}[x]$, $d \geq 1$, M_d und $S \subset \mathbb{R}^n$ wie zuvor. Falls jedes Polynom $m \in \mathbb{R}[x]$ mit $\deg(m) = 1$ und $m(u) \geq 0$ für alle $u \in S$ in M_d enthalten ist, so ist die Lasserre-Relaxierung vom Grad d von S exakt.*

Beweis. Es sei $A = \text{clos}(\text{conv}(S))$. Falls $A \neq \text{clos}(L_d)$, so gibt es $u_0 \in L_d$ mit $u_0 \notin A$. Nach Satz 5.4 existiert dann eine Hyperebene H in \mathbb{R}^n , die u_0 und A strikt trennt, etwa $H = \{u \in \mathbb{R}^n \mid \langle c, u \rangle = \beta\}$ für $c = (\gamma_1, \dots, \gamma_n)^T \in \mathbb{R}^n$, $\beta \in \mathbb{R}$ mit $\langle c, v \rangle > \beta$ für alle $v \in A$ und $\langle c, u_0 \rangle < \beta$. Setze $m = \langle c, x \rangle - \beta = \sum_{i=1}^n \gamma_i x_i - \beta \in \mathbb{R}[x]$. Wegen $u \in L_d$ gibt es $\ell \in M_d^*$ mit $u_0 = (\ell(x_1), \dots, \ell(x_n))$. Nun ist $\ell(m) = \ell(\sum_{i=1}^n \gamma_i x_i - \beta) = (\sum_{i=1}^n \gamma_i \ell(x_i)) - \beta = m(\ell(x_1), \dots, \ell(x_n)) = m(u_0) < 0$ (beachte $\ell(1) = 1$, also $\ell(\beta) = \beta$). Es folgt $m \notin M_d$. \square

Bemerkung 15.11. Falls S nicht-leeres Inneres besitzt, so gilt auch die Umkehrung; siehe etwa [10, Prop. 3.1].

Satz 15.12. *Es seien $h_1, \dots, h_r \in \mathbb{R}[x]$, M und $S \subset \mathbb{R}^n$ wie zuvor. Falls S kompakt ist, so ist die Folge der Lasserre-Relaxierungen von S bzgl. h_1, \dots, h_r konvergent, falls eine der folgenden beiden Bedingungen erfüllt ist:*

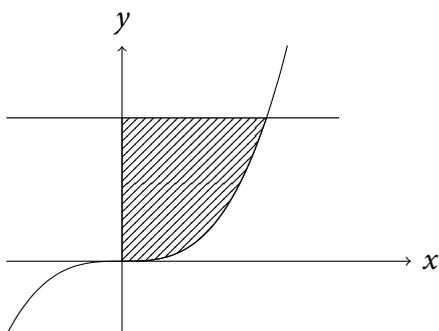
- (a) *Es gibt ein $r \in \mathbb{R}$ mit $r^2 - \sum_{i=1}^n x_i^2 \in M$.*
- (b) *M ist abgeschlossen unter Produkten, d.h. es gilt $fg \in M$ für alle $f, g \in M$.³*

Beweisskizze. Unter einer der beiden Voraussetzungen (a) oder (b) enthält der quadratische Modul M alle Polynome, die auf S nur strikt positive Werte annehmen (unter der Voraussetzung (a) sagt das der Positivstellensatz von Putinar, unter der Voraussetzung (b) der von Schmüdgen; siehe etwa [12]). Ist also $u_0 \in \mathbb{R}^n \setminus \text{conv}(S)$, dann gibt es (wie im

³Ein quadratischer Modul mit dieser Eigenschaft wird *Präordnung* genannt.

Beweis der vorangehenden Proposition) ein Polynom $m \in \mathbb{R}[x]$ vom Grad 1 mit $m(u) > 0$ für alle $u \in \text{conv}(S)$ und $m(u_0) < 0$. Wegen $M = \bigcup_{d=1}^{\infty} M_d$ und $m \in M$ gibt es dann ein d_0 mit $m \in M_{d_0}$. Es folgt $u_0 \notin L_{d_0}$ und insgesamt $\text{conv}(S) = \bigcap_{d \geq 1} L_d$, wie behauptet. \square

Beispiel 15.13. Zur Illustration diskutieren wir ein Beispiel im Detail. Es seien $h_1 = y - x^3$, $h_2 = x$, $h_3 = 1 - x$, $h_4 = y$, $h_5 = 1 - y$ in Variablen x, y . Offenbar ist S bereits selbst konvex, und wir behaupten, dass $L_3 = S$ gilt. Wir benutzen Prop. 15.10 und zeigen, dass M_3 alle Polynome $m \in \mathbb{R}[x, y]$ vom Grad 1 enthält, die auf S nicht-negativ sind.



Sei $m \in \mathbb{R}[x]$ ein solches Polynom. Falls $m(u) > 0$ für alle $u \in S$, so nimmt u auf S sein Minimum ε an (da S kompakt ist) und wegen $\varepsilon \in M_3$ genügt es dann $m - \varepsilon \in M_3$ zu zeigen. Wir müssen also nur solche m betrachten, die auf S eine Nullstelle haben, die also eine Stützgerade von S beschreiben. Nach dem Farkasschen Lemma (Aufgabe 10.3) ist jedes $m \in \mathbb{R}[x]_1$, das auf der Box $[0, 1] \times [0, 1]$ nicht-negativ ist, in $\text{cone}(x, 1 - x, y, 1 - y) \subset M_3$ enthalten. (Dabei ist $h_4 = y$ redundant, denn es gilt $y = x^2 \cdot x + y - x^3 = h_1 + x^2 h_2$.)

Es bleibt nur der Fall zu betrachten, dass m eine Tangente an die Kurve $y = x^3$ für $x \in (0, 1)$ beschreibt. Die Tangente an diese Kurve im Punkt (a, a^3) wird durch das Polynom $m_a(x, y) = y - 3a^2x + 2a^3$ beschrieben. Durch etwas Probieren und quadratische Ergänzung finden wir

$$\begin{aligned} m_a &= y - 3a^2x + 2a^3 = x^3 - 3a^2x + 2a^3 + (y - x^3) \\ &= (x - a)^2x + 2a(x - a)^2 + (y - x^3) \\ &= 2a(x - a)^2 + h_1 + (x - a)^2h_2 \in M_3. \end{aligned}$$

Bemerkung 15.14. Man kann zeigen, dass die Lasserre-Relaxierung für die Menge in Beispiel 3.6 (also $S \cup [-1, 0] \times [0, 1]$ mit S wie oben) beschrieben durch $h_1 = y - x^3$, $h_2 = x + 1$, $h_3 = y$, $h_4 = 1 - y$ für kein $d \geq 0$ exakt ist, was im wesentlichen an dem nicht-exponierten Punkt $(0, 0)$ liegt. Trotzdem ist auch diese Menge ein spektraler Schatten (siehe [10]).

Übung 15.3. Bestimme L_1 und L_2 in Beispiel 15.13.

Helton und Nie haben in [5], [6] eine Serie von Resultaten bewiesen, die hinreichende Bedingungen für die Darstellbarkeit einer abgeschlossenen konvexen semialgebraischen Teilmenge als spektraler Schatten formulieren. Diese Bedingungen sind recht verschiedener Natur (von algebraisch bis differential-geometrisch), aber die Beweise reduzieren letzten Endes immer auf die Exaktheit einer Lasserre-Relaxierung (global oder lokal). Eine Kostprobe (siehe dazu auch [14] Thm. 4.25):

Satz 15.15 (Helton-Nie). *Es sei $S = \{u \in \mathbb{R}^n \mid h_1(u) \geq 0, \dots, h_r(u) \geq 0\}$ eine kompakte basisch abgeschlossene Menge und M der von h_1, \dots, h_r erzeugte quadratische Modul wie oben. Falls M unter Multiplikation abgeschlossen ist und die Funktionen h_1, \dots, h_r auf der Menge $\text{conv}(S)$ strikt konkav sind, dann gibt es $d \geq 0$ derart, dass die Lasserre-Relaxierung vom Grad d exakt ist.* \square

ZITIERTE LITERATUR

- [1] A. Barvinok. *A course in convexity*, Bd. 54 von *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.
- [2] J. deLoera, B. Sturmfels und C. Vinzant. The central curve in linear programming, 2010. Preprint.
URL <http://arxiv.org/abs/1012.3978>
- [3] W. Forst und D. Hoffmann. *Optimization—theory and practice*. Springer Undergraduate Texts in Mathematics and Technology. Springer, New York, 2010. doi:10.1007/978-0-387-78977-4.
URL <http://dx.doi.org/10.1007/978-0-387-78977-4>
- [4] A. Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.
- [5] J. W. Helton und J. Nie. Sufficient and necessary conditions for semidefinite representability of convex hulls and sets. *SIAM J. Optim.*, 20(2) (2009) 759–791. doi:10.1137/07070526X.
URL <http://dx.doi.org/10.1137/07070526X>
- [6] ——. Semidefinite representation of convex sets. *Math. Program.*, 122(1, Ser. A) (2010) 21–64. doi:10.1007/s10107-008-0240-y.
URL <http://dx.doi.org/10.1007/s10107-008-0240-y>
- [7] J. W. Helton und V. Vinnikov. Linear matrix inequality representation of sets. *Comm. Pure Appl. Math.*, 60(5) (2007) 654–674. doi:10.1002/cpa.20155.
- [8] A. Nemirovski. Advances in convex optimization: conic programming. In *International Congress of Mathematicians. Vol. I*, S. 413–444. Eur. Math. Soc., Zürich, 2007.
- [9] T. Netzer. On semidefinite representations of non-closed sets. *Linear Algebra Appl.*, 432(12) (2010) 3072–3078. doi:10.1016/j.laa.2010.02.005.
URL <http://dx.doi.org/10.1016/j.laa.2010.02.005>
- [10] T. Netzer, D. Plaumann und M. Schweighofer. Exposed faces of semidefinitely representable sets. *SIAM J. Optim.*, 20(4) (2010) 1944–1955. doi:10.1137/090750196.
URL <http://dx.doi.org/10.1137/090750196>
- [11] T. Netzer und R. Sinn. A note on the convex hull of finitely many projections of spectrahedra, 2009. Online Note.
URL <http://arxiv.org/abs/0908.3386>
- [12] A. Prestel und C. N. Delzell. *Positive polynomials*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2001. From Hilbert’s 17th problem to real algebra.
- [13] J. Renegar. Hyperbolic programs, and their derivative relaxations. *Found. Comput. Math.*, 6(1) (2006) 59–79. doi:10.1007/s10208-004-0136-z.
URL <http://dx.doi.org/10.1007/s10208-004-0136-z>
- [14] R. Sinn. Spectrahedra and a relaxation of convex, semialgebraic sets, 2010. Diplomarbeit, Universität Konstanz.
URL <http://www.math.uni-konstanz.de/sinn/thesis.pdf>