Polyedertheorie

Prof. Dr. Martin Grötschel 3. Januar 2001

Der nachfolgende Text entspricht im wesentlichen Kapitel 2 des Vorlesungsskripts zur Vorlesung “Lineare Optimierung”, die von mir im Wintersemester 2000/2001 gehalten wurde. In diesem Kapitel ist die Notation zur Polyedertheorie zusammengefasst.

2 Polyeder

Das Hauptanliegen dieser Vorlesung ist die Behandlung von Problemen der linearen Programmierung. Man kann die Verfahren zur Lösung derartiger Probleme rein algebraisch darstellen als Manipulation von linearen Gleichungs- und Ungleichungssystemen. Die meisten Schritte dieser Algorithmen haben jedoch auch eine geometrische Bedeutung; und wenn man erst einmal die geometrischen Prinzipien und Regeln verstanden hat, die hinter diesen Schritten liegen, ist es viel einfacher, die Algorithmen und ihre Korrektheitsbeweise zu verstehen. Wir wollen daher in dieser Vorlesung einen geometrischen Zugang wählen und zunächst die Lösungsmengen linearer Programme studieren. Speziell wollen wir unser Augenmerk auf geometrische Sachverhalte und ihre algebraische Charakterisierung legen. Dies wird uns helfen, aus geometrischen Einsichten algebraische Verfahren abzuleiten.

(2.1) Definition.

(a) *Eine Teilmenge $G \subseteq \mathbb{K}^n$ heißt* **Hyperebene**, *falls es einen Vektor $a \in \mathbb{K}^n \setminus \{0\}$ und $\alpha \in \mathbb{K}$ gibt mit*

$$G = \{x \in \mathbb{K}^n \mid a^T x = \alpha\}.$$

Der Vektor a heißt **Normalenvektor** *zu G.*

(b) *Eine Teilmenge $H \subseteq \mathbb{K}^n$ heißt* **Halbraum**, *falls es einen Vektor $a \in \mathbb{K}^n \setminus \{0\}$ und $\alpha \in \mathbb{K}$ gibt mit*

$$H = \{x \in \mathbb{K}^n \mid a^T x \leq \alpha\}.$$

Wir nennen a den **Normalenvektor** *zu H. Die Hyperebene $G = \{x \in \mathbb{K}^n \mid a^T x = \alpha\}$ heißt die zum Halbraum H gehörende Hyperebene oder die H berandende Hyperebene, und H heißt der zu G gehörende Halbraum.*

(c) *Eine Teilmenge $P \subseteq \mathbb{K}^n$ heißt* **Polyeder**, *falls es ein $m \in \mathbb{Z}_+$, eine Matrix $A \in \mathbb{K}^{(m,n)}$ und einen Vektor $b \in \mathbb{K}^m$ gibt mit*

$$P = \{x \in \mathbb{K}^n \mid Ax \leq b\}.$$

Um zu betonen, daß P durch A und b definiert ist, schreiben wir auch

$$P = P(A, b) := \{x \in \mathbb{K}^n \mid Ax \leq b\}.$$

(d) *Ein Polyeder P heißt* **Polytop**, *wenn es beschränkt ist, d. h. wenn es ein $B \in \mathbb{K}$, $B > 0$ gibt mit $P \subseteq \{x \in \mathbb{K}^n \mid \|x\| \leq B\}$.*

□

Polyeder können wir natürlich auch in folgender Form schreiben

$$P(A, b) = \bigcap_{i=1}^{m} \{x \in \mathbb{K}^n \mid A_{i}.x \leq b_i\}.$$

Halbräume sind offensichtlich Polyeder. Aber auch die leere Menge ist ein Polyeder, denn $\emptyset = \{x \mid 0^T x \leq -1\}$, und der gesamte Raum ist ein Polyeder, denn $\mathbb{K}^n = \{x \mid 0^T x \leq 0\}$. Sind alle Zeilenvektoren $A_{i}.$ von A vom Nullvektor verschieden, so sind die bei der obigen Durchschnittsbildung beteiligten Mengen Halbräume. Ist ein Zeilenvektor von A der Nullvektor, sagen wir $A_{1}. = 0^T$, so ist $\{x \in \mathbb{K}^n \mid A_{1}.x \leq b_1\}$ entweder leer (falls $b_1 < 0$) oder der gesamte Raum $\mathbb{K}^n$ (falls $b_1 \geq 0$). Das heißt, entweder ist $P(A, b)$ leer oder die Mengen $\{x \mid A_{i}. \leq b_i\}$ mit $A_{i}. = 0$ können bei der obigen Durchschnittsbildung weggelassen werden. Daraus folgt

Jedes Polyeder $P \neq \mathbb{K}^n$ ist Durchschnitt von endlich vielen Halbräumen.

Gilt $P = P(A, b)$, so nennen wir das Ungleichungssystem $Ax \leq b$ ein **P definierendes System** (von linearen Ungleichungen). Sind $\alpha > 0$ und $1 \leq i < j \leq m$, so gilt offensichtlich

$$P(A, b) = P(A, b) \cap \{x \mid \alpha A_{i}.x \leq \alpha b_i\} \cap \{x \mid (A_{i}. + A_{j}.)x \leq b_i + b_j\}.$$

Daraus folgt, daß A und b zwar $P = P(A, b)$ eindeutig bestimmen, daß aber P unendlich viele Darstellungen der Form $P(D, d)$ hat.

(2.2) Beispiel. Wir betrachten das Ungleichungssystem

$$\begin{array}{lrcr}
(1) & 2x_1 \qquad\quad & \leq & 5 \\
(2) & -\ 2x_2 & \leq & 1 \\
(3) & -x_1 - \ \ x_2 & \leq & -1 \\
(4) & 2x_1 + 9x_2 & \leq & 23 \\
(5) & 6x_1 - 2x_2 & \leq & 13
\end{array}$$

Hieraus erhalten wir die folgende Matrix A und den Vektor b:

$$A = \begin{pmatrix} 2 & 0 \\ 0 & -2 \\ -1 & -1 \\ 2 & 9 \\ 6 & -2 \end{pmatrix} \qquad b = \begin{pmatrix} 5 \\ 1 \\ -1 \\ 23 \\ 13 \end{pmatrix}$$

Das Polyeder $P = P(A, b)$ ist die Lösungsmenge des obigen Ungleichungssystems (1),. . .,(5) und ist in Abbildung 2.1 graphisch dargestellt.

□

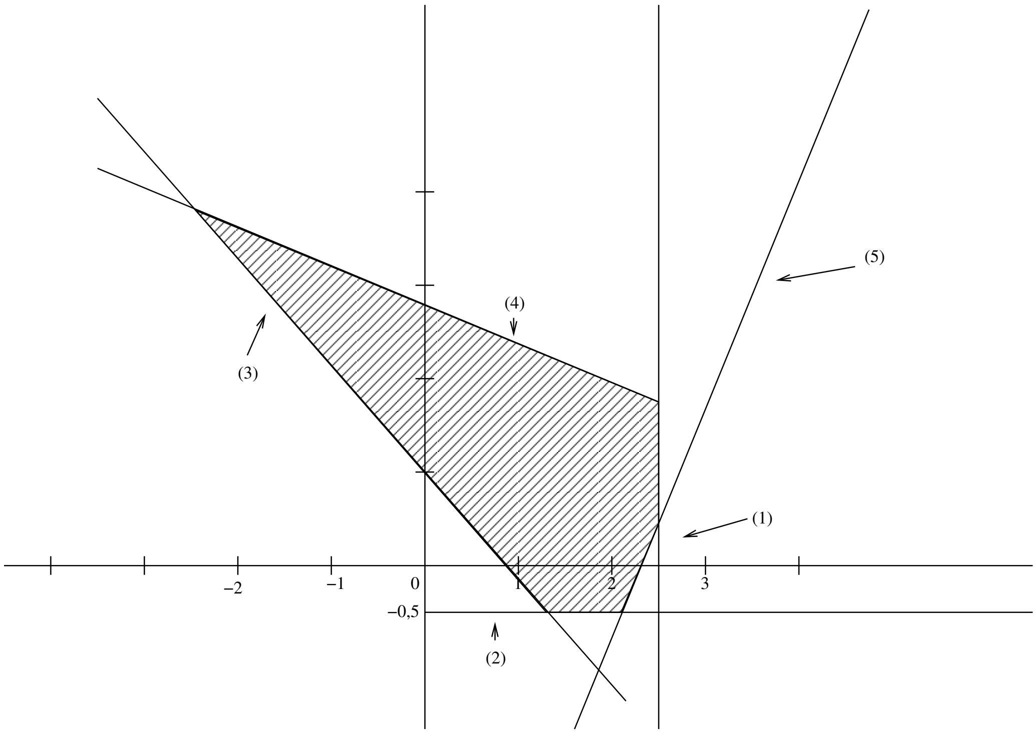

Abb. 2.1

Die Mengen zulässiger Lösungen linearer Programme treten nicht immer in der Form $Ax \leq b$ auf. Häufig gibt es auch Gleichungen und vorzeichenbeschränkte Variable. Vorzeichenbeschränkungen sind natürlich auch lineare Ungleichungssysteme, und ein Gleichungssytem $Dx = d$ kann in der Form von zwei Ungleichungssystemen $Dx \leq d$ und $-Dx \leq -d$ geschrieben werden. Allgemeiner gilt

(2.3) Bemerkung. *Die Lösungsmenge des Systems*

$$\begin{array}{ll} Bx + Cy & = c \\ Dx + Ey & \leq d \\ x & \geq 0 \\ x \in \mathbb{K}^p, & y \in \mathbb{K}^q \end{array}$$

ist ein Polyeder.

Beweis : Setze $n := p + q$ und

$$A := \begin{pmatrix} B & C \\ -B & -C \\ D & E \\ -I & 0 \end{pmatrix}, \qquad b := \begin{pmatrix} c \\ -c \\ d \\ 0 \end{pmatrix},$$

dann ist $P(A, b)$ die Lösungsmenge des vorgegebenen Gleichungs- und Ungleichungssystems.

$\square$

Ein spezieller Polyedertyp wird uns häufig begegnen, weswegen wir für ihn eine besondere Bezeichnung wählen wollen. Für $A \in \mathbb{K}^{(m,n)}$, $b \in \mathbb{K}^m$ setzen wir

$$P^=(A, b) := \{x \in \mathbb{K}^n \mid Ax = b, x \geq 0\}.$$

Nicht alle Polyeder können in der Form $P^=(A, b)$ dargestellt werden, z. B. nicht $P = \{x \in \mathbb{K} \mid x \leq 1\}$.

Wir werden später viele Sätze über Polyeder P beweisen, deren Aussagen **darstellungsabhängig** sind, d. h. die Art und Weise, wie P gegeben ist, geht explizit in die Satzaussage ein. So werden sich z. B. die Charakterisierungen gewisser Polyedereigenschaften von $P(A, b)$ (zumindest formal) von den entsprechenden Charakterisierungen von $P^=(A, b)$ unterscheiden. Darstellungsabhängige Sätze wollen wir jedoch nur einmal beweisen (normalerweise für Darstellungen, bei denen die Resultate besonders einprägsam oder einfach sind), deshalb werden wir uns nun Transformationsregeln überlegen, die angeben, wie man von einer Darstellungsweise zu einer anderen und wieder zurück kommt.

(2.4) Transformationen.

Regel I: Einführung von Schlupfvariablen

Gegeben seien $a \in \mathbb{K}^n$, $\alpha \in \mathbb{K}$. Wir schreiben die Ungleichung

(i) $a^T x \leq \alpha$

in der Form einer Gleichung und einer Vorzeichenbeschränkung

(ii) $a^T x + y = \alpha, \quad y \geq 0.$

y ist eine neue Variable, genannt **Schlupfvariable**.

Es gilt:

$$\begin{array}{lll} x & \text{erfüllt (i)} & \Longrightarrow \binom{x}{y} \text{ erfüllt (ii) für } y = \alpha - a^T x. \\ \binom{x}{y} & \text{erfüllt (ii)} & \Longrightarrow x \text{ erfüllt (i).} \end{array}$$

Allgemein: Ein Ungleichungssystem $Ax \leq b$ kann durch Einführung eines Schlupfvariablenvektors y transformiert werden in ein Gleichungssystem mit Vorzeichenbedingung $Ax + y = b$, $y \geq 0$. Zwischen den Lösungsmengen dieser beiden Systeme bestehen die oben angegebenen Beziehungen. $P(A, b)$ und $\{\binom{x}{y} \in \mathbb{K}^{n+m} \mid Ax + Iy = b, y \geq 0\}$ sind jedoch zwei durchaus verschiedene Polyeder in verschiedenen Vektorräumen.

Regel II: Einführung von vorzeichenbeschränkten Variablen

Ist x eine (eindimensionale) nicht vorzeichenbeschränkte Variable, so können wir zwei vorzeichenbeschränkte Variablen x^+ und x^- einführen, um x darzustellen. Wir setzen

$$x := x^+ - x^- \quad \text{mit} \quad x^+ \geq 0,\ x^- \geq 0.$$

□

Mit den Regeln I und II aus (2.4) können wir z. B. jedem Polyeder $P(A,b) \subseteq \mathbb{K}^n$ ein Polyeder $P^=(D,d) \subseteq \mathbb{K}^{2n+m}$ wie folgt zuordnen. Wir setzen

$$D := (A, -A, I_m), \ d := b,$$

d. h. es gilt

$$P^=(D,d) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{K}^{2n+m} \mid Ax - Ay + z = b, \ x, y, z \geq 0 \right\}.$$

Es ist üblich, die oben definierten Polyeder $P(A,b)$ und $P^=(D,d)$ **äquivalent** zu nennen. Hierbei hat "Äquivalenz" folgende Bedeutung.

Für $x \in \mathbb{K}^n$ sei

$$\begin{aligned} x^+ &:= (x_1^+, \ldots, x_n^+)^T \quad \text{mit} \quad x_i^+ = \max\{0, x_i\}, \\ x^- &:= (x_1^-, \ldots, x_n^-)^T \quad \text{mit} \quad x_i^- = \max\{0, -x_i\}, \end{aligned}$$

dann gilt:

$$x \in P(A,b) \quad \Longrightarrow \quad \begin{pmatrix} x^+ \\ x^- \\ z \end{pmatrix} \in P^=(D,d) \ \text{ für } z = b - Ax$$

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} \in P^=(D,d) \quad \Longrightarrow \quad x := u - v \in P(A,b).$$

Besonders einfache Polyeder, auf die sich jedoch fast alle Aussagen der Polyedertheorie zurückführen lassen, sind polyedrische Kegel.

(2.5) Definition. *Ein Kegel $C \subseteq \mathbb{K}^n$ heißt* **polyedrisch** *genau dann, wenn C ein Polyeder ist.*

□

(2.6) Bemerkung. *Ein Kegel $C \subseteq \mathbb{K}^n$ ist genau dann polyedrisch, wenn es eine Matrix $A \in \mathbb{K}^{(m,n)}$ gibt mit*

$$C = P(A, 0).$$

Beweis : Gilt $C = P(A,0)$, so ist C ein Polyeder und offensichtlich ein Kegel.

Sei C ein polyedrischer Kegel, dann existieren nach Definition (2.1) (c) eine Matrix A und ein Vektor b mit $C = P(A,b)$. Da jeder Kegel den Nullvektor enthält, gilt $0 = A0 \leq b$.Angenommen, es existiert ein $\overline{x} \in C$ mit $A\overline{x} \not\leq 0$, d. h. es existiert eine Zeile von A, sagen wir $A_{i.}$, mit $t := A_{i.}\overline{x} > 0$. Da C ein Kegel ist, gilt $\lambda\overline{x} \in C$ für alle $\lambda \in \mathbb{K}_+$. Jedoch für $\overline{\lambda} := \frac{b_i}{t} + 1$ gilt einerseits $\overline{\lambda}\overline{x} \in C$ und andererseits $A_{i.}(\overline{\lambda}\overline{x}) = \overline{\lambda}t > b_i$, ein Widerspruch. Daraus folgt, für alle $x \in C$ gilt $Ax \leq 0$. Hieraus ergibt sich $C = P(A,0)$.

□

In der folgenden Tabelle 2.1 haben wir alle möglichen Transformationen aufgelistet. Sie soll als "Nachschlagewerk" dienen.

Transformation nach von	$By \le d$	$By = d$ $y \ge 0$	$By \le d$ $y \ge 0$
$Ax = b$ hierbei ist $x_i^+ = \max\{0, x_i\}$ $x_i^- = \max\{0, -x_i\}$	$\begin{pmatrix} A \\ -A \end{pmatrix} y \le \begin{pmatrix} b \\ -b \end{pmatrix}$ $y = x$	$(A, -A)\, y = b$ $y \ge 0$ $y = \begin{pmatrix} x^+ \\ x^- \end{pmatrix}$	$\begin{pmatrix} A & -A \\ -A & A \end{pmatrix} y \le \begin{pmatrix} b \\ -b \end{pmatrix}$ $y \ge 0$ $y = \begin{pmatrix} x^+ \\ x^- \end{pmatrix}$
$Ax \le b$	$Ay \le b$ $y = x$	$(A, -A, I)\, y = b$ $y \ge 0$ $y = \begin{pmatrix} x^+ \\ x^- \\ b - Ax \end{pmatrix}$	$(A, -A)\, y \le b$ $y \ge 0$ $y = \begin{pmatrix} x^+ \\ x^- \end{pmatrix}$
$Ax = b$ $x \ge 0$	$\begin{pmatrix} A \\ -A \\ -I \end{pmatrix} y \le \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$ $y = x$	$Ay = b$ $y \ge 0$ $y = x$	$\begin{pmatrix} A \\ -A \end{pmatrix} y \le \begin{pmatrix} b \\ -b \end{pmatrix}$ $y \ge 0$ $y = x$
$Ax \le b$ $x \ge 0$	$\begin{pmatrix} A \\ -I \end{pmatrix} y \le \begin{pmatrix} b \\ 0 \end{pmatrix}$ $y = x$	$(A, I)\, y = b$ $y \ge 0$ $y = \begin{pmatrix} x \\ b - Ax \end{pmatrix}$	$Ay \le b$ $y \ge 0$ $y = x$