Mitschrift der Vorlesung Ganzzahlige Lineare Optimierung

26.10.2010 - 10.2.2011

Inhaltsverzeichnis

Kapitel 1

Lineare Programme

1.1 Einleitung

Definition 1.1. *Seien*

- *A eine ganzzahlige* $m \times n$*-Matrix mit den Zeilenvektoren* a_i^T
- $\{M_1, M_2\}$ *eine Partition von* $\underline{m}$
- $\{N_1, N_2\}$ *eine Partition von* $\underline{n}$
- $x \in \mathbb{R}^n, b, c \in \mathbb{Z}^n$

Dann heißt das Problem $\min(c^T x = \sum_{i=1}^m c_i x)$ *mit den Randbedingungen*

$$
\begin{aligned}
a_i^T x &= b_i & \forall i \in M_1 \\
a_i^T x &\geq b_i & \forall i \in M_2 \\
x_j &\geq 0 & \forall j \in N_1 \\
x_j &\gtreqless 0 & \forall j \in N_2
\end{aligned}
$$

allgemeines Lineares Programm *oder kurz allgemeines LP-Problem.*
Für $M_1 = N_2 = \emptyset$ *ist es in* ***kanonischer Form*** *.*
Für $M_2 = N_2 = \emptyset$ *ist es in* ***standard Form*** *.*

allgemeines Lineares Programm

kanonische Form

standard Form

Satz 1.2. *Alle drei Formen des LP-Problems sind äquivalent (d.h. können ineinander überführt werden).*

Beweis. Da LP-Probleme in kanonischer und standard Form automatisch auch allgemeine LP-Probleme sind, reicht es zu zeigen, dass jedes allgemeine LP-Problem in ein Problem in kanonischer und standard Form überführt werden kann.

1. *allgemeine in kanonische Form*
   Die Umwandlung erfolgt durch Ersetzen von Bedingungen durch äquivalente Bedingungen. Dabei muss das Gleichungssystem (Matrix) um zusätzliche Gleichungen (Zeilen) und Variablen (Spalten) erweitert werden. Alle Randbedingungen der Form $a_i^T x = b_i$ sind äquivalent zu $a_i^T x \geq b_i \wedge -a_i^T x \geq$

$-b_i$. Die Matrix erweitert sich folglich um $|M_1|$ Zeilen. Für alle Randbedingungen der Form $x_j \gtrless 0$ definiere zwei Variablen:

$$x_j^+ : = \begin{cases} x_j & \text{für } x_j > 0 \\ 0 & \text{sonst} \end{cases}$$
$$x_j^- : = \begin{cases} -x_j & \text{für } x_j < 0 \\ 0 & \text{sonst} \end{cases}$$

Folglich gilt $x_j = x_j^+ - x_j^-$. Die Bedingung $x_j \gtrless 0$ ist folglich äquivalent zu $x_j^+ \geq 0 \wedge x_j^- \geq 0$. Die Matrix erweiter sich also nochmals um $|N_1|$ Spalten (da wir eine Variable mehr haben) und der Vektor x um ebensoviele Komponenten. Das entstandene Problem ist in kanonischer Form.

2. *allgemeine in standard Form*
   siehe Übung

□

1.2 Standard Form

Wir betrachten zunächst ein Problem in standard Form unter folgenden Vorraussetzungen:

(a) $rang(A) = m$

(b) die Menge der zulässigen Punkte $F = \{x | Ax = b, x \geq 0\}$ ist nicht leer

(c) die Menge der Werte der Zielfunktion $\{c^T | x \in F\}$ ist nach unten beschränkt

Basis

Basislösung

Definition 1.3. *Eine **Basis** B von A ist eine Auswahl von linear unabhängigen Spaltenvektoren aus A. Alternativ entspricht B einer regulären $m \times n$-Matrix. Die **Basislösung** zu B ist ein Vektor $x \in \mathbb{R}^n$ mit*

$$x_{j_k} = \begin{cases} t_k & \textit{wenn } A_{j_k} \in B \\ 0 & \textit{sonst} \end{cases}$$
$$t = B^{-1} b$$

Eine Basislösung x zu einer Matrix A können wir wie folgt berechnen:

1. Wähle eine Basis B von A
2. Setze alle Komponenten von x, die nicht zu den gewählten Spalten aus B gehören auf 0.
3. Löse das resultierende Gleichungssystem, um die restlichen Komponenten von x zu bestimmen.

Lemma 1.4. *Es sei $x = (x_1, \dots, x_n)$ eine Basislösung, $\alpha = \max\limits_{i \in \underline{m}|, j \in \underline{n}|} \{|a_{i,j}|\}$ und $\beta = \max\limits_{j \in \underline{n}|} \{|b_j|\}$. Dann gilt $|x_j| \leq m! \alpha^{m-1} \beta$ und $x_j \in \mathbb{Q}$.*

Beweis. Für eine Nicht-Basiskomponente gilt die Aussage, da diese per Definition 0 sind. Die Basiskomponente x_j ist die Summe von m Produkten von Elementen von B^{-1} und b. Nach Definition der Inversen gilt:

$$B^{-1} = \frac{1}{\det B}\operatorname{Adj}(B) = \frac{1}{\det B}\begin{pmatrix} \mathcal{B}_{1,1} & \dots & \mathcal{B}_{1,m} \\ \vdots & & \vdots \\ \mathcal{B}_{m,1} & \dots & \mathcal{B}_{m,m} \end{pmatrix}$$

wobei $\mathcal{B}_{i,j}$ das Produkt von $(-1)^{i+j}$ und der Determinanten der Matrix ist, die durch Streichen der i-tem Zeile und j-ten Spalte aus B entsteht. Aufgrund des Entwicklungssatzes für Determinanten ist $\mathcal{B}_{i,j}$ die Summe von $(m-1)!$ Produkten von $m-1$ Elementen aus A. Somit gilt:

$$\begin{aligned}
|\mathcal{B}_{i,j}| &\le (m-1)!\alpha^{m-1} && \text{da alle Elemente von } A \le \alpha \text{ sind} \\
|x_j| &= \left|\sum_{i\in \underline{m}|} B^{-1}_{i,j} b_i\right| && \text{nach Definition} \\
&= \frac{1}{\det B}\left|\sum_{i\in \underline{m}|} \mathcal{B}_{i,j} b_i\right| \\
&\le \frac{1}{\det B}\left|\sum_{i\in \underline{m}|} (m-1)!\alpha^{m-1} b_i\right| \\
&= \frac{1}{\det B}\left|m\cdot(m-1)!\alpha^{m-1}\cdot\beta\right| \\
&= \frac{1}{\det B} m\cdot(m-1)!\alpha^{m-1}\cdot\beta && \text{da alle Faktoren} \ge 0 \text{ sind}
\end{aligned}$$

Da $\det B$ ganzzahlig ist, folgt $\det B \ge 1$. □

Satz 1.5. *Unter der Vorraussetzungen a und b existert mind. eine Basislösung.*

Beweis. **Annahme:** Es existiert eine Lösung $x \in F$ mit $t > m$ Nicht-Null-Komponenten und es gibt keine Lösung $x' \in F$ mit weniger Nicht-Null-Komponenten.

Wir können O.B.d.A. annehmen, dass die ersten t Komponenten von x größer als 0 sind und die restlichen Komponenten gleich 0 sind (wenn nicht können wir dies durch Vertauschung erreichen). Es gilt also $x_1, \dots, x_t > 0$ und $x_{t+1}, \dots, x_n = 0$. Hierraus folgt:

$$b = A_1x_1 + \ldots + A_nx_n = A_1x_1 + \ldots + A_tx_t \tag{1.1}$$

Es sei nun r der Rang der Matrix $[A1, \dots, A_t]$. Wenn $r = 0$ wäre, so wäre $\vec{0}$ eine zulässige Basislösung mit weniger Nicht-Null-Komponenten als x. Damit die Annahme stimmen kann muss also $0 < r \le m < t$ gelten. Mit dem Gauß-Jordan-Verfahren lassen sich nun die ersten r Zeilen und Spalten in eine reguläre

Form überführen. Das Problem sieht nun wie folgt aus:

$$\begin{pmatrix} 1 & & 0 & -\overline{a}_{1,r+1} & \dots & -\overline{a}_{1,t} \\ & \ddots & & \vdots & \ddots & \vdots \\ 0 & & 1 & -\overline{a}_{r,r+1} & \dots & -\overline{a}_{r,t} \\ ? & & & \dots & & ? \\ \vdots & & & & & \\ ? & & & & & \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ \vdots \end{pmatrix} = \begin{pmatrix} \overline{b}_1 \\ \vdots \\ \overline{b}_r \\ \vdots \end{pmatrix}$$

Es lässt sich nun (1.1) schreiben als:

$$x_j = \overline{b_j} + \sum_{i=r+1}^{t} \overline{a_{j,i}} x_i \qquad \text{für } j = 1, \dots, r \tag{1.2}$$

Setze $\Theta = \min(x_{r+1}, \Theta_1)$ mit $\Theta_1 = \min(\frac{x_j}{\overline{a}_{j,r+1}}, j = 1, \dots, r, \overline{a}_{j,r+1} > 0)$. Konstruiere eine neue Lösung $\hat{x}$:

$$\hat{x}_j = \begin{cases} \overline{b}_j + \sum_{i=r+1}^{t} \overline{a}_{j,i} \hat{x}_i & \text{für } j < r+1 \\ x_j - \Theta & \text{für } j = r+1 \\ x_j & \text{für } j > r+1 \end{cases} \tag{1.3}$$

Diese erfüllt das Gleichungssystem. Dann gilt für $j < r$:

$$\hat{x}_j = \overline{b}_j + \sum_{i=r+1}^{t} \overline{a}_{j,i} \hat{x}_i \qquad j < r+1 \tag{1.4}$$

$$= \overline{b}_j + \hat{a}_{j,r+1}(x_{r+1} - \Theta) + \sum_{i=r+1}^{t} \overline{a}_{j,i} \hat{x}_i \qquad j < r+1 \tag{1.5}$$

$$= \underbrace{\overline{b}_j + \sum_{i=r+1}^{t} \overline{a}_{j,i} \hat{x}_i}_{x_j} - \Theta \overline{a}_{j,r+1} \tag{1.6}$$

Ist $\Theta = x_{r+1}$, so ist $\hat{x}_{r+1} = 0$. Ist $\Theta = \Theta_1 = \frac{x_k}{\overline{a}_{k,r+1}}$. mit $k \le r$, so ist $\hat{x}_k = x_k - \Theta \overline{a}_{k,r+1} = 0$. Für die Zulässigkeit ist noch zu zeigen: $\hat{x}_j \ge 0 \forall j \le r+1$. Sehen wir uns $\hat{x}_{r+1}$ an: Nach Definition: $\hat{x}_{r+1} = x_{r+1} - \Theta \ge \Theta - \Theta = 0$

$$\hat{x}_j = x_j - \Theta \overline{a}_{j,r+1} \tag{1.7}$$

$$\ge \begin{cases} x_j + \Theta |\overline{a}_{j,r+1}| > 0 & \overline{a}_{j,r+1} \le 0 \\ x_j - \frac{x_j}{\overline{a}_{j,r+1}} \overline{a}_{j,r+1} = 0 & \overline{a}_{j,r+1} > 0 \end{cases} \tag{1.8}$$

Insgesamt ist $\hat{x}$ eine zulässige Lösung mit einer Nicht-Nullkomponente weniger als x, was ein Widerspruch zur Annahme ist. □

Es gibt also eine Lösung $x \in F$ mit $t \le m$ Nicht-Nullkomponentent. O.B.d.A sind die zugehörigen Spalten linear unabhängig (ansonsten Arg. oben mit Elimination von Variablen bzw. Gleich 0 setzen wiederholen). Falls $t < m$ ist, erweitere die Spalten zu einer Basis für x. (Austauschsatz von Steinitz)

Satz 1.6. *Unter der Vorraussetzungen a,b und c ist das Lineare Programm* $\min(c^T x), Ax = b, x \geq 0$ *äquivalent zu* $\min(c^T x), Ax = b, x \geq 0, x \leq M$ *wobei* $M = (m+1)!\alpha^m \beta$ *mit* $\alpha = max(|a_{i,j}|, |c_j|), \beta = max(|b_i|, |z|)$ *und z die Größte untere Schranke von* $\{c^T x | Ax = b, x \geq 0\}$ *ist.*

> In diesem Fall ist äquivalent zu verstehen als Gleichheit der Lösungsmengen. Es ist also jede Lösung des ersten Problems auch eine Lösung des zweiten Problems und umgekehrt.

1.3 Geometrie von Linearen Programmen

Ein **linearer Teilraum** S des $\mathbb{R}^d$ ist eine Teilmenge von $\mathbb{R}^b$, die bezüglich der Vektoraddition und Skalarmultiplikation abgeschlossen ist. Ein **affiner Teilraum** A des $\mathbb{R}^d$ ist ein linearer Teilraum S, verschoben um einen Vektor $u \in \mathbb{R}^d$: $A = \{u + x | x \in S\}$. Die Dimension eines linearen Teilraumes S ist gleich der maximale Zahl von linear unabhängigen Vektoren in S. Ebenso bei affinen Teilräumen. Die Dimension irgendeiner Menge X ist die kleinste Dimension eines affinen Teilraums, der X enthält. Äquivalent können wir einen affinen bzw. linearen Teilraum des $\mathbb{R}^d$ wie folgt darstellen: $A = \{x \in \mathbb{R}^d | a_i^T x = b_i, 1 \leq i \leq m\}$ bzw. $S = \{x \in \mathbb{R}^d | a_i^T x = 0, 1 \leq i \leq m\}$. Zum Beispiel hat eine Kante die Dimension 1 und ein Menge von k Punkten höchstens die Dimension k-1. Ein affiner Teilraum des $\mathbb{R}^d$ der Dimension $d-1$ wird als **Hyperebene** bezeichnet. Alternativ ist dies eine Menge von Punkten $H = \{x \in \mathbb{R}^d | a^T x = b\}$ mit $a \neq 0$. Eine Hyperebene definiert zwei **abgeschlossene Halbräume** $H^+ = \{x \in \mathbb{R}^d | a^T x \geq b\}$ und $H^- = \{x \in \mathbb{R}^d | a^T x \leq b\}$. Der Durchschnitt von endl. vielen Halbräumen wird bezeichnet als ein **Polyeder** . Ein Polyeder heißt **Polytop** , wenn es beschränkt ist.

linearer Teilraum
affiner Teilraum
Hyperebene
abgeschlossene Halbräume
Polyeder
Polytop

Es sei P ein Polytop der Dimension d im $\mathbb{R}^d$ und es sein $H \in \mathbb{R}^d$ eine Hyperebene. So ist $F = P \cap H$ eine Seitenfläche von P, wenn H mindestens einen Punkt gemeinsam mit P hat und wenn P in höchstens einem der beiden Halbräumen H^+ und H^- liegt. Es gibt drei wichtige Fälle:

1. eine **Facette** , d.h. eine Seitenfläche der Dimension $d-1$. Facette
2. eine **Ecke** , d.h. eine Seitenfläche der Dimension 0. Ecke
3. eine **Kante** , d.h. eine Seitenfläche der Dimension 1. Kante

Beispiel: ein Würfel im $\mathbb{R}^3$ ist gegeben durch $P = \{(x_1, x_2, x_3) | 0 \leq x_i 1 \leq 1, i = 1, 2, 3\}$. Dieser Würfel hat 6 Facetten, 8 Ecken und 12 Kanten.

$F_1 = P \cap \{(x_1, x_2, x_3) | x_3 = 1\}$	Facette
$F_2 = P \cap \{(x_1, x_2, x_3) | x_1 - x_3 = 1\}$	Ecke
$F_3 = P \cap \{(x_1, x_2, x_3) | x_1 + x_2 + x_3 = 1\}$	Kante

1.3.1 Umwandlung

Satz 1.7. 1. *Jedes Polytop ist die konvexe Hülle seiner Ecken.*

2. *Ist V eine endliche Menge von Punkten/Vektoren, so ist die konvexe Hülle von V ein Polytop.*

konvexe Hülle

Definition 1.8. ***konvexe Hülle***

$$conv(V) = \Big\{ \sum_{i=1}^{k} \lambda_i v_i \Big| v_i \in V, \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1 \Big\}$$

Wir sind hauptsächlich interessiert an Polytopne, die im positiven Oktanten liegen. D.h. die ersten d-Halbräume sind definiert durch $x_j \geq 0 (j = 1, \ldots, d)$. Wegen 1.7 kann ein Polytop P auf verschiedene Weisen dargestellt werden:

a) als konvexe Hülle einer endlichen Punktemenge

b) als Durchschnitt von endlich vielen Halbräumen (der zusätzlich beschränkt ist)

c) als zulässiger Bereich eines LP-Problems (algebraische Darstellung)

Umwandlung LP-Problem in Durchschnitt von Halbräumen Es sei $F = \{x | Ax = b, x \geq 0\}$ der zulässige Bereich eines LP-Problems. Wir setzen vorraus, dass die Bedingungen a,b und c erfüllt sind. Da $rang(A) = m$ gilt, können wir O.B.d.A. annehmen, dass die Gleichungen $Ax = b$ in der folgenden Form vorliegen:

$$x_{i+n-m} = b_i - \sum_{j=1}^{n-m} a_{i,j} x_j \qquad i = 1, \ldots, m \tag{1.9}$$

Die Umformung erfolgt über Gauß-Jordan. Die Matrix enthält also einen Einheitsmatrix-Bereich im rechts-oberen Teil. Wegen der Vorraussetzung, dass $x_{i+n-m} \geq 0$ ist $Ax = b, x \geq 0$ äquivalent zu:

$$b_i - \sum_{j=1}^{n-m} a_{i,j} x_j \geq 0 \qquad i = 1, \ldots, m \tag{1.10}$$

$$x_j \geq 0 \qquad i = 1, \ldots, n-m \tag{1.11}$$

Das ist der Durchschnitt von n Halbräumen und ein Polytop im $\mathbb{R}^{n-m}$.

Umwandlung Durchschnitt von Halbräumen in LP-Problem Die n Halbräume, die P bestimmen seinen

$$h_{i,1} x_1 + \ldots + h_{i,n-m} x_{n-m} \leq g_i \qquad i = 1, \ldots, n$$

Nach unserer Vorraussetzung sind die ersten $n-m$ Ungleichungen von der Form $x_i \geq 0, i = 1, \ldots, n-m$. Verwende positive Schlupfvariablen $x_{n-m+1}, \ldots, x_n$ und erhalte:

$$h_{i,1} x_1 + \ldots + h_{i,n-m} x_{n-m} + x_i = g_i \qquad i = n-m+i, \ldots, n$$

Nun kann man dieses Problem als LP-Problem schreiben. Bilde dazu $A = (H, I)$ (ein $m \times n$-Matrix) und erhalte das folgende System:

$$
\begin{aligned}
A &= \begin{pmatrix} h_{1,1} & \dots & h_{1,n-m} & 1 & & 0 \\ \vdots & \ddots & \vdots & & \ddots & \\ h_{i,1} & \dots & h_{i,n-m} & 0 & & 1 \end{pmatrix} \\
Ax &= b \\
x &\geq 0 \\
b &= (g_{n-m+1}, \dots, g_n)^T
\end{aligned}
$$

Definition 1.9. *Der x-Vektor in den Halbraumdurchschnitten hatte nur die Dimension $n-m$, der x-Vektor des LP-Problems hatte die Dimension n. Jeder Punkt $\hat{x} \in (x_1, ..., x_{n-m})^T \in P$ kann transformiert werden zu einem Punkt aus F mit n Komponenten. Dazu muss man lediglich die Schlupfvariablen ausrechnen:*

$$x_i = g_i - \sum_{j=1}^{n-m} h_{i,j} x_j \qquad \text{für } i = n-m+1, \dots, n$$

Ungekehrt wird ein $x \in F$ durch weglassen der letzten m Komponenten zu einem Punkt $\hat{x} \in P$.

Satz 1.10. *Es sei P ein nicht-leeres Polytop, $F = \{x | Ax = b, x \geq 0\}$ der zugehörige zulässige Bereich eines LP und $\hat{x} = (x_1, \dots, x_{n-m})^T \in P$. Dann sind äquivalent:*

a) Der Punkt $\hat{x}$ ist eine Ecke von P.

b) Wenn $\hat{x} = \lambda \hat{x}^1 + (1-\lambda)\hat{x}^2$ mit $\hat{x}^1, \hat{x}^2 \in P$ und $0 < \lambda < 1$, dann gillt $\hat{x} = \hat{x}^1 = \hat{x}^2$.

c) Der zugeordnete Vektor $x \in \mathbb{R}^n$, definiert durch 1.9 ist eine zulässige Basislösung von F.

Beweis. **a $\Rightarrow$ b:** siehe Übung

b $\Rightarrow$ c: $\hat{x}$ habe die Eigenschaft b. Betrachte den zugeordneten Punkt $x \in F$ und die Menge $B = \{A_j | x_j > 0; j = 1, \dots, n\}$. Wir wollen zeigen, dass B eine linear unabhängige Menge von Spalten ist.
Annahme: B ist eine linear abhängige Menge von Spaltenvektoren.
Dann gibt es Zahlen λ_j (nicht alle 0) mit $\sum_{A_j \in B} \lambda_j A_j = 0$. Da $x \in F$ gilt, gilt $\sum_{A_j \in B} x_j A_j = b$ und $x_j \geq 0$. Dann gilt für jedes $\theta \in \mathbb{R}$:

$$\sum_{A_j \in B} (x_j \pm \theta \lambda_j) A_j = b$$

Da $x_j > 0$ für alle $A_j \in B$, gibt es ein $\theta > 0$ mit $x_j \pm \theta \lambda_j \geq 0 \forall A_j \in B$. Dann gibt es aber zwei Punkte $x^1, x^2 \in F$, definiert durch:

$$x_j^i = \begin{cases} x_j + (-1)^i \theta \lambda & A_j \in B \\ 0 & \text{sonst} \end{cases} \qquad i = 1, 2$$

Da gilt:

$$x = \frac{1}{2}x^1 + \frac{1}{2}x^2$$

Für die zugeordneten Punkte $\hat{x}^1, \hat{x}^2 \in P$ gilt $\hat{x} = \frac{1}{2}\hat{x}^1 + \frac{1}{2}\hat{x}^2$, da diese durch Weglassung von Komponenten entstehen. Jedoch gilt $\hat{x}^1 \neq \hat{x}^2$. Also ist B eine Menge von linear unabhängigen Spaltenvektoren mit $|B| \leq m$ und kann zu einer Basis erweitert werden.

c $\Rightarrow$ a: Wenn $y = (y_1, \ldots, y_n)^T$ eine zulässige Basislösung von $Ax = b, x \geq 0$ ist, so gibt es einen Kostenvektor c, so dass y der eindeutige Lösungsvektor des Systems $c^T x \leq c^T y, Ax = b, x \geq 0$ ist. Wir benutzen 1.9 um hierraus Bedingungen für Punkte aus dem Polytop P zu gewinnen:

$$\begin{aligned} c^T x &= \sum_{i=1}^{n} c_i x_i \\ &= \sum_{i=1}^{n-m} c_i x_i + \sum_{i=n-m+1}^{n} c_i x_i \\ &= \sum_{i=1}^{n-m} c_i x_i + \sum_{i=n-m+1}^{n} c_i \left(g_i - \sum_{j=1}^{n-m} h_{i,j} x_j \right) \end{aligned}$$

Fassen wir die Koeffizienten von x_j jeweils zusammen, so bekommen wir:

$$d_j = c_j - \sum_{i=n-m+1}^{n} c_i h_{i,j}$$

Es folgt dann, dass $\hat{y} = (y_1, \ldots, y_{n-m})$ der eindeutige Punkt im $\mathbb{R}^{n-m}$ ist mit

$$d^T \hat{x} \leq d^T \hat{y} \forall \hat{x} \in P$$

Also ist $\hat{x}$ eine Ecke von P.

> Die Menge der $\hat{x}$, für die $d^T \hat{x} = d^T \hat{y}$ gilt, bilden eine Hyperebene. Da der Schnittpunkt dieser Hypereben mit P exakt einen Punkt hat (nämlich $\hat{y}$) und P komplett auf einer Seite dieser Hyperebene liegt (nämlich $d^T \hat{x} \geq d^T \hat{y} \forall \hat{x} \in P$), ist $\hat{y}$ eine Ecke vom P.

$\square$

Index