

KAPITEL 1 — EINFÜHRUNG: STABILE MATCHINGS

F. VALLENTIN, A. GUNDERT

In diesem Kapitel werden wir ein erstes konkretes Problem des Operations Research kennenlernen. Es handelt sich um das Problem des stabilen Matchings, ein wichtiges Zuordnungsproblem, das viele Anwendungen in der nichtmonetären Ökonomie besitzt. Im Jahr 2012 bekamen Lloyd S. Shapley and Alvin E. Roth den Nobelpreis für Ökonomie für *the theory of stable allocations and the practice of market design*.

1. GRUNDBEGRIFFE

Definition 1.1. *Gegeben seien Mengen*

$$\mathcal{M} = \{A_1, \dots, A_n\} \quad \text{und} \quad \mathcal{F} = \{B_1, \dots, B_n\}$$

mit je n Elementen. Ein Matching ist eine Bijektion $\sigma: \mathcal{M} \rightarrow \mathcal{F}$.

Dieses hat die folgende Interpretation¹: Die n -elementige Menge \mathcal{M} stellt n Männer dar, die n -elementige Menge \mathcal{F} n Frauen. Ein Matching $\sigma: \mathcal{M} \rightarrow \mathcal{F}$ entspricht n „traditionellen“ Hochzeiten: Es handelt sich um monogame, heterosexuelle Ehen, wobei niemand ledig ist.

Desweiteren stellen wir uns vor, dass jeder Mann eine *Präferenzliste* für die n Frauen besitzt und umgekehrt jede Frau eine Präferenzliste für die n Männer besitzt. Beide Präferenzlisten werden durch zwei Matrizen mit jeweils n^2 Einträgen dargestellt.

Beispiel 1.2. *Hier ist ein einfaches Beispiel für ein Matching mit vier Männern $\mathcal{M} = \{A, B, C, D\}$ und vier Frauen $\mathcal{F} = \{a, b, c, d\}$.*

$A \bullet \text{-----} \bullet a$

$B \bullet \text{-----} \bullet b$

$C \bullet \text{-----} \bullet c$

$D \bullet \text{-----} \bullet d$

Folgende Tabellen sind Beispiele für Präferenzlisten:

Männer	Frauen
$A : c \ b \ d \ a$	$a : A \ B \ D \ C$
$B : b \ a \ c \ d$	$b : C \ A \ D \ B$
$C : b \ d \ a \ c$	$c : C \ B \ D \ A$
$D : c \ a \ d \ b$	$d : B \ A \ C \ D$

Date: 28. April 2014.

¹Den anachronistischen Charakter dieses Modells nehmen wir unkommentiert hin.

Dabei ist die Präferenz höher, je weiter links eine Person gelistet ist. Zum Beispiel bevorzugt Mann A Frau c gegenüber Frau b, Frau b gegenüber Frau d und diese wiederum gegenüber Frau a.

Notation 1.3. Wir schreiben zum Beispiel

$$c >_A b \iff \text{Mann A bevorzugt Frau c gegenüber Frau b}$$

und analog

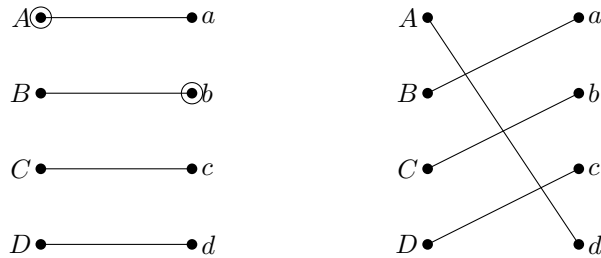
$$A >_a B \iff \text{Frau a bevorzugt Mann A gegenüber Mann B.}$$

Definition 1.4. Ein Matching $\sigma: \mathcal{M} \rightarrow \mathcal{F}$ heißt stabil, wenn für alle $M \in \mathcal{M}$ und alle $f \in \mathcal{F}$ mit $\sigma(M) \neq f$ stets

$$\sigma(M) >_M f \quad \text{oder} \quad \sigma^{-1}(f) >_f M$$

gilt. Ansonsten könnte Interesse an einem Seitensprung bestehen.

In unserem Beispiel:



ist nicht stabil

ist stabil

Die Bestimmung von stabilen Matchings hat viele Anwendungen bei Zuordnungsproblem, z.B. wird seit 1952 in den USA im Rahmen des National Resident Matching Program Medizinstudenten auf Krankenhäuser verteilt, wobei ein Verfahren angewendet wird, das auf dem Konzept der stabilen Matchings beruht. In realistischen Situationen muss man das Modell anpassen, da z.B. die Präferenztabellen nicht vollständig sind, oder die Mengen \mathcal{M} und \mathcal{F} unterschiedliche Kardinalität aufweisen.

Die Definition wirft zwei Fragen auf:

- (1) Wie findet man ein stabiles Matching?
- (2) Gibt es immer eins?

Die erste Frage kann man naiv beantworten: Teste einfach alle $n!$ Möglichkeiten. Das ist aber keine gute Idee, weil die Rechenzeit schnell sehr groß wird, wie die folgende Tabelle zeigt:

n	Rechenzeit (Annahme: pro Bijektion 10^{-9} Sekunden)
10	0,03 Sekunden
15	21 Minuten
20	77 Jahre
25	$4,9 \cdot 10^8$ Jahre ($\approx \frac{1}{10}$ Alter der Erde)

2. GALE-SHAPLEY ALGORITHMUS

Der Algorithmus von Gale und Shapley gibt eine deutlich bessere Antwort auf die erste Frage. Außerdem kann mit ihm auch die zweite Frage positiv beantworten.

Algorithmus 2.1. (in Pseudocode)

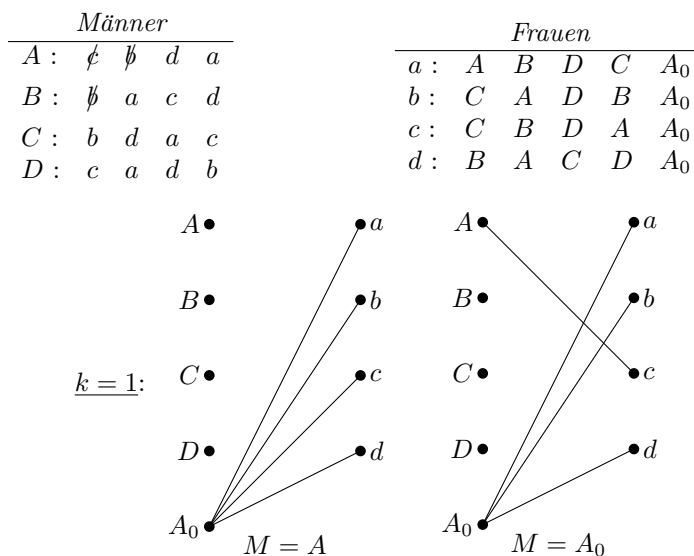
```

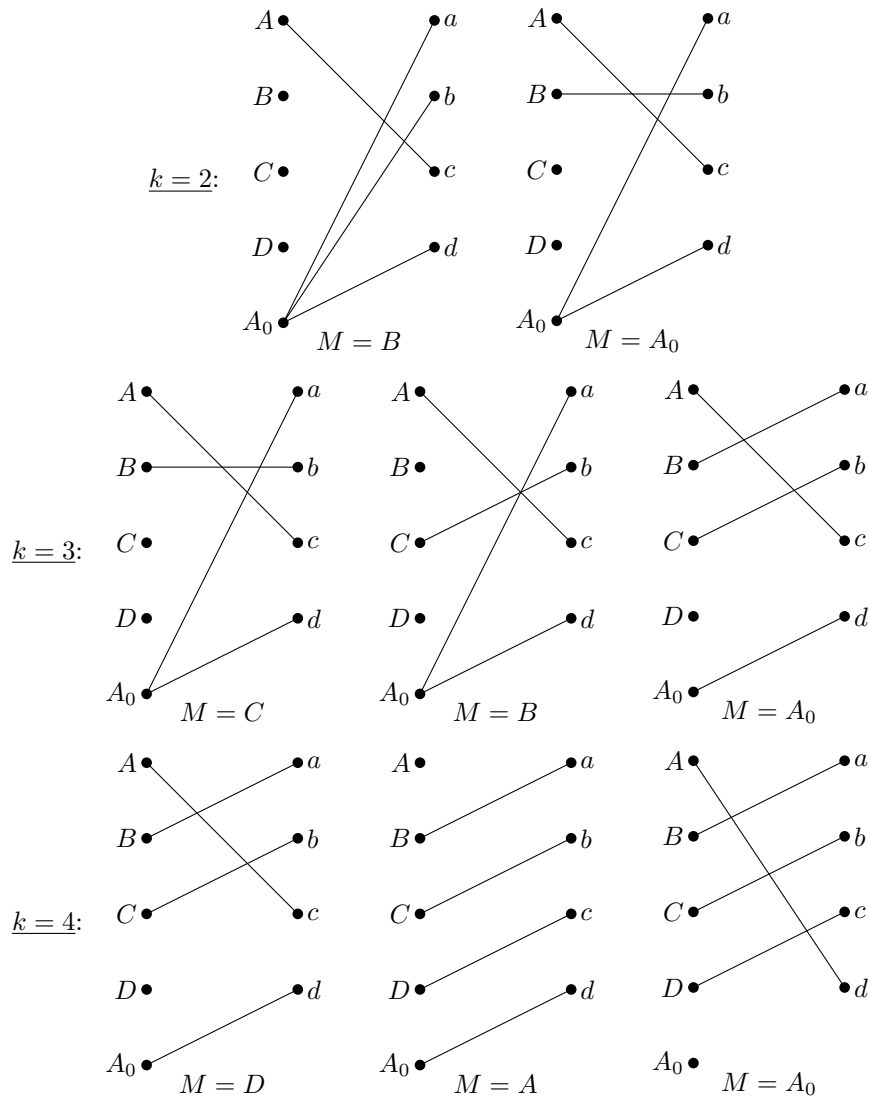
Verlobe Frauen  $a_1, \dots, a_n$  mit virtuellem Mann  $A_0$ , den jede Frau am
wenigsten bevorzugt. ;
for  $k = 1$  to  $n$  do
   $M = A_k$ 
  while  $M \neq A_0$  do
     $f =$  beste verbleibende Frau auf  $M$ 's Liste
     $M$  macht  $f$  Heiratsantrag
    if  $M >_f$  aktueller Verlobter von  $f$  then
      verlobe  $M$  und  $f$ 
       $M =$  bisheriger Verlobter von  $f$ 
    end
    if  $M \neq A_0$  then
      streiche  $f$  von  $M$ 's Liste
    end
  end
end

```

Algorithmus 1 : Gale-Shapley Algorithmus

In unserem Beispiel:





Das letzte Matching ist stabil, was auch folgender Satz zeigt.

Satz 2.2. *Der Gale-Shapley Algorithmus berechnet ein stabiles Matching. Insbesondere gibt es immer eines.*

Beweis. (1) Der Algorithmus ist wohldefiniert: Die Listen der Männer sind niemals leer. Denn angenommen die Liste von M wäre leer. Dann ist M von allen Frauen abgewiesen worden. Jede Frau ist aber höchstens mit einem Mann verlobt, die Situation der Frauen verschlechtert sich nie und M ist stets besser als A_0 . Es muss also neben M (und A_0) n „echte“ Männer geben. Widerspruch

(2) Das resultierende Matching ist stabil. Denn angenommen $\sigma(M) \neq f$ und $f >_M \sigma(M)$. Dann ist M im Laufe des Algorithmus von f abgewiesen worden und $\sigma^{-1}(f) >_f M$. Somit ist die Stabilität nicht verletzt. \square

Bemerkung 2.3. *Im Algorithmus werden (für $n \geq 2$) weniger als n^2 Heiratsanträge gemacht (siehe erstes Aufgabenblatt) \rightarrow Viel effizienter als $n!$.*

Satz 2.4. *Der Gale-Shapley Algorithmus berechnet das für die Männer beste stabile Matching, das heißt wenn f von M 's Liste entfernt wird, dann gibt es kein stabiles Matching σ mit $\sigma(M) = f$.*

Für den Beweis benötigen wir noch das folgende Lemma:

Lemma 2.5. *Falls „streiche f von M 's Liste “aufgerufen wird, gibt es einen Mann \tilde{M} mit $\tilde{M} >_f M$ und $f >_{\tilde{M}}$ alle Frauen auf \tilde{M} 's verbleibenden Liste.*

Beweis. „Streiche f von M 's Liste “passiert, wenn

- (1) M möchte f heiraten, aber es gibt einen Mann \tilde{M} , den aktuellen Verlobten von f mit $\tilde{M} >_f M$.
- (2) M und f waren verlobt, aber es gibt einen Mann \tilde{M} mit $\tilde{M} >_f M$, der f einen Heiratsantrag gemacht hat.

Es gibt also in beiden Situationen einen Mann \tilde{M} mit $\tilde{M} >_f M$ für den auch gilt: $f >_{\tilde{M}}$ alle Frauen auf \tilde{M} 's verbleibender Liste. \square

Beweis. (des Satzes): Wir zeigen: Für alle stabilen Matchings σ und $M \in \mathcal{M}$, $f \in \mathcal{F}$ gilt: $\sigma(M) = f \Rightarrow f$ wird nicht von M 's Liste gestrichen.

Angenommen das stimmt nicht. Dann gibt es ein stabiles Matching σ mit mindestens einem Paar (M, f) sodass

- $\sigma(M) = f$ und
- f wird von M 's Liste gestrichen.

Wähle unter diesen Paaren (M, f) dasjenige aus, bei dem die „Streichung“ am frühesten passiert (im Laufe des Algorithmus). Nach dem Lemma gibt es \tilde{M} mit $\tilde{M} >_f M$ und $f >_{\tilde{M}}$ alle Frauen auf \tilde{M} 's verbleibender Liste (zum Zeitpunkt der Streichung). Betrachte $\sigma(\tilde{M})$: Da σ stabil, gilt $\sigma(\tilde{M}) >_{\tilde{M}} f = \sigma(M)$. Das heißt $\sigma(\tilde{M})$ wurde zu einem früheren Zeitpunkt von \tilde{M} 's Liste gestrichen. Widerspruch zur Wahl von (M, f) . \square

Korollar 2.6. *Das stabile Matching, welches der Gale-Shapley Algorithmus berechnet, ist unabhängig von der Reihenfolge der Männer und Frauen.*

3. GEOMETRISCHE MODELLIERUNG

Der Gale-Shapley Algorithmus ist ein *kombinatorischer Algorithmus* (verwendet keine Zahlen).

Vorteil: + sehr schnell, effizient

Nachteil: - relativ kompliziert

- sehr problemspezifisch (kann aber angepasst werden, z.B. für m Männer, n Frauen, unvollständige Präferenzlisten)

Jetzt: Geometrische Modellierung mit linearen Ungleichungen.

Vorteil: + einfache Modellierung
+ leicht veränderbar

Nachteil: - im Allgemeinen nicht die schnellste Lösungsmethode

Codiere ein Matching $\sigma: \mathcal{M} \rightarrow \mathcal{F}$, $\mathcal{M} = \{A_1, \dots, A_n\}$, $\mathcal{F} = \{a_1, \dots, a_n\}$ als Permutationsmatrix:

$$X \in \{0, 1\}^{n \times n} \quad \text{mit} \quad X_{ij} = \begin{cases} 1, & \text{falls } \sigma(A_i) = a_j \\ 0, & \text{sonst.} \end{cases}$$

Satz 3.1. $X \in \{0, 1\}^{n \times n}$ codiert ein stabiles Matching $\Leftrightarrow X$ erfüllt folgende Ungleichungen:

- (1) $\sum_{j=1}^n X_{ij} \leq 1$ für alle $i = 1, \dots, n$ (Zeilensummen)
- (2) $\sum_{i=1}^n X_{ij} \leq 1$ für alle $j = 1, \dots, n$ (Spaltensummen)
- (3) $X_{ij} \geq 0$ für alle $i, j = 1, \dots, n$
- (4) $X_{ij} + \sum_{k: a_k > a_i a_j} X_{ik} + \sum_{k: A_k > a_j A_i} X_{kj} \geq 1$ für alle $i, j = 1, \dots, n$ (Stabilitätsbedingung)

Beweis. Klar nach Definition von stabilem Matching. □

Später in der Vorlesung:

Die Menge

$$\text{SMP} = \{X \in \mathbb{R}^{n \times n} : X \text{ erfüllt die Bedingungen (1)-(4)}\}$$

ist ein Polytop (Lösungsmenge eines Systems linearer Ungleichungen, die beschränkt ist). Zum Beispiel ein 5-Eck ist ein zweidimensionales Polytop oder ein Würfel ist ein dreidimensionales Polytop.

Vande Vate (1989):

Die Ecken des Polytops SMP codieren genau die stabilen Matchings.

Sei $C \in \mathbb{R}^{n \times n}$ eine Heiratsbewertungsmatrix. (C_{ij} gibt den Wert der Heirat zwischen A_i und a_j an). Dann erhält man das optimale stabile Matching bezüglich C mit dem linearen Optimierungsproblem:

$$\begin{aligned} \max \quad & \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} \\ & X \in \mathbb{R}^{n \times n} \\ & X \text{ erfüllt Bedingungen (1)-(4)} \end{aligned}$$

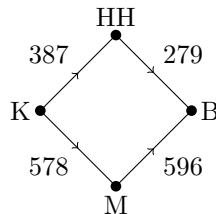
PROF. DR. F. VALLENTIN, DR. A. GUNDELT, MATHEMATISCHES INSTITUT, UNIVERSITÄT ZU KÖLN,
WEYERTAL 86-90, 50931 KÖLN, DEUTSCHLAND
E-mail address: frank.vallentin@uni-koeln.de, anna.gundert@uni-koeln.de

KAPITEL 2 — KÜRZESTE WEGE

F. VALLENTIN, A. GUNDERT

Das Ziel dieses Kapitels ist es kürzeste Wege in einem gegebenen Netzwerk zu verstehen und zu berechnen.

Ein einführendes Beispiel für ein Netzwerk zwischen den vier Städten Köln (K), Hamburg (HH), München (M) und Berlin (B) ist der folgende Graph:



Die Zahlen an den Kanten geben Entfernungen zwischen den Städten an. Ein Ziel könnte zum Beispiel sein, den kürzesten Weg von Köln nach Berlin zu finden, welcher in diesem Netzwerk auch leicht ohne jede Methodik zu erkennen ist. In größeren Netzwerken sind kürzeste Wege jedoch nicht immer per Auge zu sehen. Daher lernen wir systematische Verfahren kennen, kürzeste Wege zu berechnen. Dafür benötigen wir zunächst etwas Notation.

1. NICHTNEGATIVE KANTENLÄNGEN

Notation 1.1. *Wir nennen*

$D = (V, A)$ *einen gerichteten Graph*

wobei

V *eine endliche Menge von Knoten*

und

$A \subseteq \{(v, w) \in V \times V : v \neq w\}$ *eine Menge gerichteter Kanten ist.*

Auf den Kanten definieren wir Kantenlängen durch eine Längenfunktion

$$l: A \rightarrow \mathbb{Z}_+ = \{0, 1, \dots\}$$

$P = (v_0, a_1, v_1, a_2, v_2, \dots, a_m, v_m)$ *heißt Kantenfolge, falls $v_0, \dots, v_m \in V$, $a_1, \dots, a_m \in A$ und $a_i = (v_{i-1}, v_i)$. Wir nennen v_0 Startknoten und v_m Endknoten.*

Die Länge von P definieren wir als

$$l(P) = \sum_{i=1}^m l(a_i)$$

Wenn $|\{v_0, \dots, v_m\}| = m + 1$ (d.h. die Knoten sind alle paarweise verschieden), heißt P v_0 - v_m -Weg oder kurz Weg.

Date: 5. Mai 2014.

Für Knoten $s, t \in V$ ist der Abstand von s zu t definiert als

$$\text{dist}(s, t) := \min_{P \text{ ist } s\text{-}t\text{-Weg}} l(P).$$

Für eine Knotenmenge $U \subseteq V$ setze

$$\delta^+(U) := \{(v, w) \in A : v \in U, w \notin U\}$$

als die Menge der Kanten, die U verlassen und

$$\delta^-(U) := \{(v, w) \in A : v \notin U, w \in U\}$$

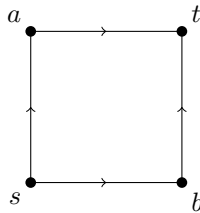
als die Menge der Kanten, die U betreten.

Eine Kantenmenge $A' \subseteq A$ heißt s - t -Schnitt, falls

$$A' = \delta^+(U)$$

für ein $U \subseteq V$ und $s \in U, t \notin U$.

Beispiel 1.2. Wir bestimmen in dem Graph



alle s - t -Schnitte. Alle Teilmengen der Knoten, die s enthalten, aber nicht t , sind

$$U_1 = \{s\}, U_2 = \{s, a\}, U_3 = \{s, b\}, U_4 = \{s, a, b\}.$$

Alle s - t -Schnitte haben die Form $\delta^+(U_i)$ für $i = 1, \dots, 4$. Also bekommen wir die vier s - t -Schnitte

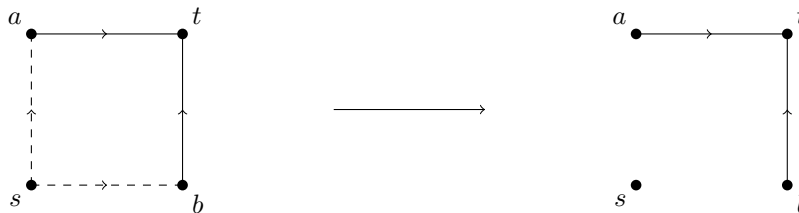
$$\delta^+(U_1) = \{(s, a), (s, b)\}$$

$$\delta^+(U_2) = \{(s, b), (a, t)\}$$

$$\delta^+(U_3) = \{(s, a), (b, t)\}$$

$$\delta^+(U_4) = \{(a, t), (b, t)\}$$

Noch mal zur Verdeutlichung: s - t -Schnitte sind Kantenmengen, die jedoch durch Knotenmengen bestimmt sind. Die Kanten in einem s - t -Schnitt trennen eine Teilmenge der Knoten, die s enthält, vom Rest der Knoten, welcher t enthält: Nach Entfernen der Kanten im Schnitt gibt es keinen Weg mehr von s nach t .



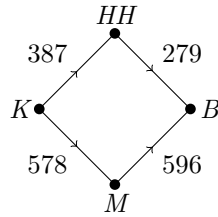
In diesem Beispiel haben wir die Kanten aus $\delta^+(U_1)$ (gestrichelt) entfernt. In dem resultierenden Graphen rechts gibt es keinen s - t -Weg mehr.

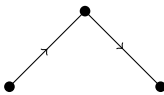
Satz 1.3. (*min-max-Charakterisierung kürzester Wege, Robacker 1965*)

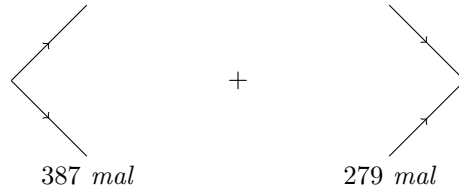
Es sei $D = (V, A)$ ein gerichteter Graph $s, t \in V$ zwei Knoten und $l: A \rightarrow \mathbb{Z}$ eine Längenfunktion. Dann gilt:

$$\begin{aligned} \text{dist}(s, t) &= \min_{P \text{ ist } s\text{-}t\text{-Weg}} l(P) \\ &= \max k, \text{ wobei} \\ &\quad k = \text{Anzahl von } s\text{-}t\text{-Schnitten } C_1, \dots, C_k \text{ (mit Vielfachheiten),} \\ &\quad \text{so dass für alle } a \in A \text{ gilt:} \\ &\quad |\{j = 1, \dots, k : a \in C_j\}| \leq l(a) \end{aligned}$$

Bemerkung 1.4. Die s - t -Schnitte geben ein Zertifikat, dass ein gegebener s - t -Weg minimal ist. In dem Beispiel



ist  der kürzeste K - B -Weg, weil es folgende K - B -Schnitte gibt:



Beweis. (Satz 1.3)

Falls es keinen s - t -Weg gibt, ist die Aussage trivial: Beide Seiten sind gleich $+\infty$. Im folgenden existiere ein s - t -Weg:

min \geq max: Sei $P = (v_0, a_1, v_1, \dots, a_m, v_m)$ ein s - t -Weg und seien C_1, \dots, C_k s - t -Schnitte mit oben genannter Eigenschaft. Dann gilt:

$$\begin{aligned} l(P) &= \sum_{i=1}^m l(a_i) \geq \sum_{i=1}^m |\{j = 1, \dots, k : a_i \in C_j\}| \\ &= \sum_{j=1}^k |C_j \cap \{a_1, \dots, a_m\}| \\ &\geq \sum_{j=1}^k 1 = k \end{aligned}$$

min \leq max: Definiere

$$U_i = \{v \in V : \text{dist}(s, v) < i\} \quad \text{für } i = 1, \dots, \text{dist}(s, t) = \min.$$

Definiere

$$C_i = \delta^+(U_i) = \{(v, w) \in A : v \in U_i, w \notin U_i\}.$$

Dies sind $\text{dist}(s, t)$ viele s - t -Schnitte. Es sei $a = (u, v) \in A$ eine Kante. Dann gilt:

$$l(a) = \text{dist}(s, v) - \text{dist}(s, u).$$

Andererseits enthalten nur die s - t -Schnitte

$$C_{\text{dist}(s, u)+1}, \dots, C_{\text{dist}(s, v)}$$

diese Kante a , was $\text{dist}(s, v) - \text{dist}(s, u)$ viele sind. \square

2. BELIEBIGE KANTENLÄNGEN

Manchmal sind auch negative Kantenlängen nützlich. Zum Beispiel wenn man einen längsten s - t -Weg finden will (also einen kürzesten negativen Weg).

Beispiel 2.1. Rucksackproblem

Wir haben einen Rucksack mit $2,5l$ Volumen und 5 nützliche Gegenstände.

Gegenstand i	Volumen a_i	Nutzen c_i
1 : Schlafsack	1,5l	4
2 : Taschenmesser	0,5l	4
3 : Kekse	1l	3
4 : Thermoskanne	1,5l	5
5 : Isomatte	1l	4

Aufgabe: Finde eine Auswahl von $1, \dots, 5$, so dass die Gegenstände in den Rucksack passen und ihr summierter Nutzen maximal ist.

Formulierung als kürzeste Wege Problem:

Wir erstellen einen Graphen $D = (V, A)$ mit Kantenlängen wie folgt:

$$\underline{\text{Knoten}}: (i, x) \in V \quad i = 0, 1, \dots, 6, \quad x = 0, \frac{1}{2}, 1, \dots, \frac{5}{2}$$

$$\underline{\text{Kanten}}: - ((i-1, x), (i, x)) \text{ mit Länge } 0 \quad i = 1, \dots, 5$$

$$- ((i-1, x), (i, x + a_i)) \text{ mit Länge } -c_i \quad i = 1, \dots, 5$$

$$- ((5, x), (6, \frac{5}{2})) \text{ mit Länge } 0 \quad \text{für alle } x$$

Dann: Kürzeste $(0, 0)$ - $(6, \frac{5}{2})$ -Wege in D geben eine optimale Auswahl.

```

Input : Gerichteter Graph  $D = (V, A)$ ,  $n = |V|$ ,  $s \in V$ ,  $l: A \rightarrow \mathbb{Z}$ 
Output : Funktionen  $d_0, \dots, d_n: V \rightarrow \mathbb{Z}$ ,  $g: V \setminus \{s\} \rightarrow V$ 
Setze  $d_0(s) = 0$ ,  $d_0(v) = \infty \forall v \in V \setminus \{s\}$ .
for  $k = 0$  to  $n - 1$  do
     $d_{k+1}(v) = d_k(v) \forall v \in V$ 
    for  $(u, v) \in A$  do
        if  $d_{k+1}(v) > d_k(u) + l(u, v)$  then
             $d_{k+1}(v) = d_k(u) + l(u, v)$ 
             $g(v) = u$ 
        end
    end
end
if  $d_n \neq d_{n-1}$  then
    Ausgabe: „Es gibt einen Kreis negativer Länge, der von  $s$  aus erreichbar ist.“.
end

```

Algorithmus 1 : Bellman-Ford Algorithmus

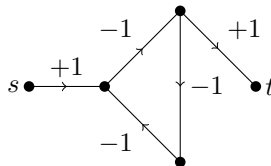
Satz 2.2. *Es gilt*

$$d_k(v) = \min\{l(P) : P \text{ ist } s\text{-}v\text{-Kantenfolge, die höchstens } k \text{ Kanten enthält}\}.$$

Beweis. (triviale) Induktion nach k . □

Mögliches Problem: gerichtete Kreise negativer Länge.

Beispiel 2.3. *In dem Graph*



existiert keine kürzeste s - t -Kantenfolge, weil der negative Kreis beliebig oft durchlaufen werden kann.

Definition 2.4. *Eine Kantenfolge $P = (v_0, a_1, v_1, \dots, a_m, v_m)$ heißt (gerichteter) Kreis, falls $v_0 = v_m$ und $|\{v_0, \dots, v_m\}| = m$.*

Satz 2.5. *Es sei $D = (V, A)$ ein gerichteter Graph mit Längenfunktion $l: A \rightarrow \mathbb{Z}$. Alle gerichteten Kreise in D haben nicht-negative Länge. Seien s, t durch (mindestens) eine Kantenfolge verbunden. Dann existiert eine kürzeste s - t -Kantenfolge, die ein Weg ist.*

Beweis. Klar ist, dass ein kürzester s - t -Weg P existiert. Angenommen es gibt eine s - t -Kantenfolge Q (die kein Weg ist) mit $l(P) > l(Q)$. Wähle ein solches Q mit minimaler Anzahl von Kanten. Da Q kein Weg ist, enthält Q einen Kreis C , der nach Voraussetzung nicht-negative Länge hat, also $l(C) \geq 0$. Sei Q' die Kantenfolge, die man aus Q erhält, indem man C löscht. Dann ist Q' ebenfalls eine s - t -Kantenfolge mit

$$l(Q') = l(Q) - l(C) \leq l(Q) < l(P),$$

aber mit weniger Kanten als Q . Widerspruch! □

Zurück zu Bellman-Ford:

- Laufzeit: proportional zu $|V| \cdot |A| \leq |V|^3$
- Falls D keine gerichteten Kreise negativer Länge enthält, dann gilt $d_{n-1}(v) = \text{dist}(s, v)$ und $v, g(v), g(g(v)), \dots, s$ ist die Umkehrung eines kürzesten s - t -Weges.

Satz 2.6. $d_n = d_{n-1} \iff$ alle von s aus erreichbaren Kreise haben nicht-negative Länge.

Beweis. Aufgabe 3.1. □

3. GEOMETRISCHE MODELLIERUNG, POTENTIALE

Definition 3.1. Eine Funktion $p: V \rightarrow \mathbb{R}$ heißt Potential für D, l , falls

$$\forall a = (u, v) \in A: \quad p(v) - p(u) \leq l(a)$$

Satz 3.2. Es existiert genau ein Potential $p: V \rightarrow \mathbb{R}$ für D, l , wenn für alle gerichteten Kreise C in D gilt, dass $l(C) \geq 0$.

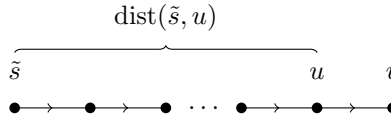
Beweis. „ \Rightarrow “: Sei $C = (v_0, a_1, v_1, \dots, a_m, v_m)$ mit $v_0 = v_m$ ein Kreis und sei p eine Potentialfunktion. Dann gilt:

$$l(C) = \sum_{i=1}^m l(a_i) \geq \sum_{i=1}^m (p(v_i) - p(v_{i-1})) = 0$$

„ \Leftarrow “: Füge zu $D = (V, A)$ einen neuen Knoten \tilde{s} und neue Kanten (\tilde{s}, t) für alle $t \in V$ hinzu, wobei $l(\tilde{s}, t) = 0$ definiert wird. Dann ist $p(t) = \text{dist}(\tilde{s}, t)$ eine Potentialfunktion. Denn

$$\begin{aligned} \forall a = (u, v) \in A: \quad p(v) - p(u) &= \text{dist}(\tilde{s}, v) - \text{dist}(\tilde{s}, u) \leq l(a) \\ &\Leftrightarrow \text{dist}(\tilde{s}, v) \leq \text{dist}(\tilde{s}, u) + l(a). \end{aligned}$$

Dies gilt, denn



ist eine \tilde{s} - v -Kantenfolge der Länge $\text{dist}(\tilde{s}, u) + l(a)$ und, weil alle gerichteten Kreise nicht-negative Länge haben, gibt es eine kürzeste \tilde{s} - v -Kantenfolge, die ein Weg ist, also:

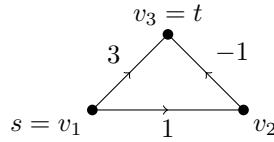
$$\text{dist}(\tilde{s}, v) \leq \text{dist}(\tilde{s}, u) + l(a).$$

□

Satz 3.3. (geometrische Modellierung kürzester Wege mit linearen Ungleichungen) Es sei $D = (V, A)$ ein gerichteter Graph, $l: A \rightarrow \mathbb{Z}$ eine Längenfunktion. Alle gerichteten Kreise haben nicht-negative Länge. Seien $s, t \in V$ und es gebe einen s - t -Weg. Dann gilt

$$\begin{aligned} \text{dist}(s, t) &= \min_{P \text{ ist } s\text{-}t\text{-Weg}} l(P) \\ &= \max_{\substack{p: V \rightarrow \mathbb{R} \\ p(v) - p(u) \leq l(a) \forall a = (u, v) \in A}} p(t) - p(s) \end{aligned}$$

Beispiel 3.4. Wir wenden den Satz auf den Graphen



an. Man erhält

$$\begin{aligned}
 \text{dist}(v_1, v_3) &= \max_{\substack{p_3 - p_1 \\ p_1, p_2, p_3 \in \mathbb{R} \\ p_2 - p_1 \leq 1 \\ p_3 - p_1 \leq 3 \\ p_3 - p_2 \leq -1}} p_3 - p_1 \\
 &= \max_{\substack{(-1, 0, 1)p \\ p \in \mathbb{R}^3 \\ (-1, 1, 0)p \leq 1 \\ (-1, 0, 1)p \leq 3 \\ (0, -1, 1)p \leq -1}} p
 \end{aligned}$$

Dabei schreiben wir $p = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$ als Vektor statt als Funktion, womit gemeint ist, dass jedem Funktionswert $p(v_i)$ genau ein Eintrag p_i zugeordnet wird. Auf diese Schreibweise wird bei geometrischer Modellierung von graphentheoretischen Problemen häufig zurückgegriffen.

Beweis. (Satz 3.3)

max \leq min: Sei $P = (s = v_0, a_1, v_1, \dots, a_m, v_m = t)$ eine s - t -Kantenfolge und $p: V \rightarrow \mathbb{R}$ ein Potential. Dann gilt

$$l(P) = \sum_{i=1}^m l(a_i) \geq \sum_{i=1}^m (p(v_i) - p(v_{i-1})) = p(t) - p(s)$$

max \geq min: Setze $p(v) = \text{dist}(s, v)$, falls eine s - v -Kantenfolge existiert und $p(v) = 0$ sonst. Analog wie im Beweis von Satz 3.2 überprüft man, dass p ein Potential ist. \square

PROF. DR. F. VALLENTIN, DR. A. GUNDERT, MATHEMATISCHES INSTITUT, UNIVERSITÄT ZU KÖLN,
WEYERTAL 86-90, 50931 KÖLN, DEUTSCHLAND
E-mail address: frank.vallentin@uni-koeln.de, anna.gundert@uni-koeln.de

KAPITEL 3 — MATCHINGS IN BIPARTITEN GRAPHEN

F. VALLENTIN, A. GUNDERT

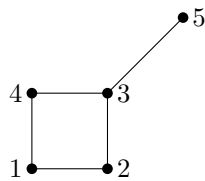
1. DEFINITIONEN

Notation 1.1. *Ähnlich wie im vorangegangenen Kapitel zunächst etwas Notation. Wir beschäftigen uns jetzt mit ungerichteten Graphen, das sind Graphen, in denen Kanten keine Orientierung besitzen. Die meisten Begriffe definiert man sehr ähnlich wie im gerichteten Fall.*

Wir nennen $G = (V, E)$ einen ungerichteten Graph, wobei

V eine endliche Menge von Knoten und

$E \subseteq \{\{u, v\} : u, v \in V, u \neq v\}$ eine Menge ungerichteter Kanten ist.



$$V = \{1, \dots, 5\}$$

$$E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 1\}, \{3, 5\}\}$$

Die Begriffe Kantenfolge, Weg, Kreis definieren wir genau wie bei gerichteten Graphen.

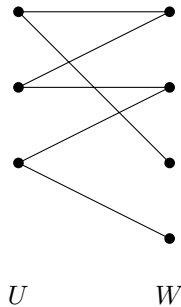
Sei $v \in V$ ein Knoten, dann heißt $w \in V$ Nachbar von v , falls $\{v, w\} \in E$.

Zwei Knoten $u, v \in V$ heißen wegzusammenhängend, falls es einen u - v -Weg gibt. Die Relation „wegzusammenhängend“ ist eine Äquivalenzrelation, ihre Äquivalenzklassen heißen Zusammenhangskomponenten. Falls G nur eine Zusammenhangskomponente besitzt, so heißt G zusammenhängend.

Der Graph $G = (V, E)$ heißt bipartit, falls zwei Teilmengen $U, W \subseteq V$ existieren, so dass

$$V = U \dot{\cup} W \quad (\text{das heißt } V = U \cup W \text{ und } U \cap W = \emptyset) \text{ und}$$

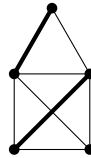
$$\forall e \in E: |e \cap U| = |e \cap W| = 1.$$



2. MATCHINGS MIT MAXIMALER KARDINALITÄT

Definition 2.1. Sei $G = (V, E)$ ein ungerichteter Graph. Ein Matching $M \subseteq E$ ist eine Teilmenge disjunkter Kanten, das heißt

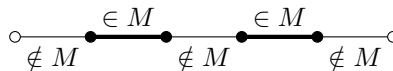
$$\forall e, f \in M : e \cap f = \emptyset.$$



Im folgenden fasse wir einen Weg $P = (v_0, e_1, v_1, \dots, e_m, v_m)$ als Teilmenge $P = \{e_1, \dots, e_m\} \subseteq E$ auf.

Definition 2.2. Sei $M \subseteq E$ ein Matching. Ein Weg $P \subseteq E$ heißt M -augmentierend, falls

- (1) seine Endknoten nicht von M überdeckt sind und
- (2) seine Kanten alternierend aus M und nicht aus M sind.



Klar: Falls P ein M -augmentierender Weg ist, dann ist

$$M' = M \Delta P = (M \setminus P) \cup (P \setminus M) \quad (\text{symmetrische Differenz})$$

ein Matching mit $|M'| = |M| + 1$.

Satz 2.3. Sei M ein Matching in G . Entweder ist M ein Matching mit maximaler Kardinalität, oder es gibt einen M -augmentierenden Weg.

Beweis. (1) Falls M maximale Kardinalität hat, folgt sofort, dass es keinen M -augmentierenden Weg geben kann (siehe oben).

- (2) Sei M' ein Matching mit $|M'| > |M|$. Betrachte die Zusammenhangskomponenten von $G' = (V, M \cup M')$. Da jeder Knoten von G' höchstens zwei Nachbarn hat, bestehen die Zusammenhangskomponenten von G' nur aus Wegen (eventuell der Länge 0) oder aus Kreisen (siehe Aufgabe 3.2). Da $|M'| > |M|$, muss es eine Zusammenhangskomponente von G' geben, die mehr Kanten aus M' als aus M enthält. Diese Zusammenhangskomponente ist ein M -augmentierender Weg.

□

Algorithmus 2.4. zur Bestimmung eines Matchings maximaler Kardinalität

```

M = ∅
while ∃ M-augmentierender Weg P do
    M = M Δ P.
end
    
```

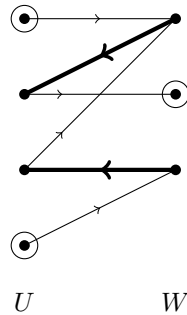
Wie findet man M-augmentierende Wege?

→ einfacher Algorithmus für bipartite Graphen (→ hier)

→ Edmonds „Blüten“-Algorithmus für allgemeine Graphen (→ Spezialvorlesung).

Sei $G = (V, E)$ ein bipartiter Graph mit Partition $V = U \dot{\cup} W$ und sei $M \subseteq E$ ein Matching. Definiere den gerichteten Residualgraph $D_M = (V, A_M)$ mit

$$A_M = \{(u, w) \in U \times W : e = \{u, w\} \in E \setminus M\} \cup \{(w, u) \in W \times U : e = \{u, w\} \in E \cap M\}$$



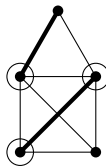
Seien $U_M \subseteq U$, $W_M \subseteq W$ die Knoten, die nicht von M überdeckt werden. Dann entspricht jeder gerichtete Weg von einem Knoten in U_M zu einem Knoten in W_M einem M-augmentierenden Weg und umgekehrt.

Algorithmische Umsetzung:

Finde kürzeste Wege zwischen je zwei Knoten in U_M und W_M mit Bellman-Ford, wobei $l: A_M \rightarrow \mathbb{Z}$, $l(a) = 1$ gesetzt wird (dies ist eine einfache, aber suboptimale Strategie; besser: verwende zum Beispiel Tiefensuche).

3. KÖNIGS MATCHING THEOREM

Definition 3.1. Sei $G = (V, E)$ ein ungerichteter Graph. Eine Teilmenge $C \subseteq V$ heißt Knotenüberdeckung, falls $\forall e \in E : |C \cap e| \geq 1$.



Satz 3.2. (*Königs Matching Theorem, 1931*)

Sei $G = (V, E)$ ein bipartiter Graph. Dann gilt

$$\begin{aligned}\nu(G) &= \max\{|M| : M \subseteq E \text{ Matching in } G\} \\ &= \min\{|C| : C \subseteq V \text{ Knotenüberdeckung von } G\} \\ &= \tau(G).\end{aligned}$$

Beweis. $\max \leq \min$: Sei M ein Matching und C eine Knotenüberdeckung. Dann gilt $|M| \leq |C|$, weil je mindestens einer der Endknoten der Kanten in M zu C gehören muss.

$\max \geq \min$: Sei $M \subseteq E$ ein Matching in G mit maximaler Kardinalität,

$$M = \{\{u_1, w_1\}, \dots, \{u_m, w_m\}\}, \quad u_i \in U, w_i \in W.$$

Betrachte den Residualgraph D_M und definiere $C = \{v_1, \dots, v_m\}$ durch

$$v_i = \begin{cases} w_i, & \text{falls es in } D_M \text{ einen gerichteten Weg von } U_M \text{ nach } w_i \text{ gibt.} \\ u_i, & \text{sonst.} \end{cases}$$

Wir zeigen nun: C ist eine Knotenüberdeckung mit $|C| = m$ (wobei Letzteres klar ist). Sei dafür $\{a, b\} \in E$ eine beliebige Kante. Zu zeigen ist: $a \in C$ oder $b \in C$.

Sei o.B.d.A. $a \in U, b \in W$.

Fall 1: $a \in U_M, b \in W_M$.

Dann wäre $\{a, b\}$ ein M -augmentierender Weg, aber das ist unmöglich, weil M ein maximales Matching ist.

Fall 2: $a \in U_M, b \notin W_M$.

Dann ist $(a, b) \in A_M \Rightarrow b \in C$ (jeweils nach Definition von A_M bzw. C).

Fall 3: $a \notin U_M, b \in W_M$.

Angenommen $a \notin C$. Weil $a \notin U_M$, gibt es ein $w_i \in C$ mit $(w_i, a) \in A_M$. Somit gibt es einen gerichteten Weg in D_M von U_M nach w_i . Dieser kann um die Kanten $(w_i, a), (a, b)$ verlängert werden, was einem M -augmentierenden Weg entspricht. Widerspruch zur Maximalität von M . Das heißt $a \in C$.

Fall 4: $a \notin U_M, b \notin W_M$.

Falls $a \in C$, ist nichts zu zeigen. Sei nun $a \notin C$. Dann gibt es ein w_i mit $\{a, w_i\} \in M$ und $w_i \in C$. Falls $w_i = b$, also $b \in C$, sind wir fertig. Falls $w_i \neq b$, dann gibt es einen gerichteten Weg von U_M nach w_i und dieser kann um die Kanten $(w_i, a), (a, b)$ erweitert werden, das heißt $b \in C$. \square

4. MATCHINGS MIT MAXIMALEM GEWICHT

Im folgenden beschäftigen wir uns mit Graphen, die gewichtete Kanten besitzen, das heißt jeder Kante wird ein Gewicht (eine Zahl) zugeordnet. Ziel wird es wieder sein, ein maximales Matching zu bestimmen. Aber diesmal wollen wir nicht, wie vorher, die Anzahl der Kanten, sondern das Gesamtgewicht maximieren. Nochmal zur Übersicht:

gegeben: $G = (V, E)$ ungerichteter Graph, $w: E \rightarrow \mathbb{R}$ Gewichtsfunktion.
 Für $M \subseteq E$ definiere $w(M) = \sum_{e \in M} w(e)$.

gesucht: $\nu_w(G) = \max\{w(M) : M \subseteq E \text{ Matching in } G\}$.

Definition 4.1. Ein Matching $M \subseteq E$ heißt extrem, falls für alle Matchings $M' \subseteq E$ mit $|M'| = |M|$ gilt, dass $w(M') \leq w(M)$.

Definition 4.2. Sei $M \subseteq E$ ein Matching in G . Definiere die Längenfunktion $l_M: E \rightarrow \mathbb{R}$ durch

$$l_M(e) = \begin{cases} w(e), & \text{falls } e \in M \\ -w(e), & \text{falls } e \notin M. \end{cases}$$

Für $P \subseteq E$ definiere $l_M(P) = \sum_{e \in P} l_M(e)$.

Satz 4.3. Sei $M \subseteq E$ ein extremes Matching und $P \subseteq E$ ein M -augmentierender Weg minimaler Länge. Dann ist auch $M' = M \Delta P$ ein extremes Matching.

Beweis. Sei N ein extremes Matching mit $|N| = |M| + 1$. Da $|N| > |M|$ ist, enthält der Graph $(V, M \cup N)$ eine Zusammenhangskomponente Q , die ein M -augmentierender Weg ist. Dann gilt $l_M(Q) \geq l_M(P)$. Es ist $N \Delta Q$ ein Matching mit $|N \Delta Q| = |M|$. Also $w(N \Delta Q) \leq w(M)$ und zusammen

$$w(N) = w(N \Delta Q) - l_M(Q) \leq w(M) - l_M(P) = w(M'),$$

das heißt M' ist extrem. (Man beachte, dass die Gleichheiten gelten, weil die Längenfunktion l_M den Kanten, die nicht in M liegen, den negativen Wert von w zuordnet.) \square

Algorithmus 4.4. Bestimmung eines Matchings mit maximalem Gewicht („ungarische Methode“, Egerváry 1931)

```

M0 = ∅ ; k = 1;
while ∃ Mk-1-augmentierender Weg do
    Wähle Mk-1-augmentierenden Weg P mit minimaler Länge
    Mk = Mk-1 Δ P
    k = k + 1
end
output: max{w(Mi) : i = 0, 1, ..., k - 1}.
    
```

Finde M_{k-1} -augmentierenden Weg mit minimaler Länge:

Betrachte wieder den Residualgraph D_M ($M = M_{k-1}$)

$$D_M = (V, A_M), A_M = \{(u, w) \in U \times W : \{u, w\} \in E \setminus M\} \cup \{(w, u) \in W \times U : \{u, w\} \in E \cap M\}$$

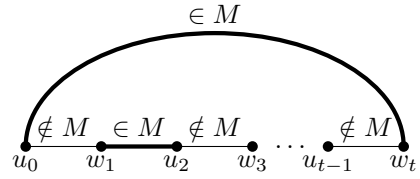
mit Längenfunktion $l_M: A_M \rightarrow \mathbb{R}$

$$l_M((a, b)) = l_M(\{a, b\}).$$

Finde einen kürzesten Weg von U_M nach W_M . Dies kann man mit Bellman-Ford machen, weil D_M keine gerichteten Kreise negativer Länge besitzt, wie der folgende Satz zeigt:

Satz 4.5. Sei $M \subseteq E$ ein extremes Matching. Dann besitzt D_M keine gerichteten Kreise negativer Länge.

Beweis. Angenommen $C \subseteq E$ entspricht einem gerichteten Kreis in D_M mit $l_M(C) < 0$. Dann haben wir die folgende Situation:



Dann ist $M' = M \Delta C$ ein Matching mit $|M'| = |M|$ und es gilt:

$$w(M') = w(M) - l_M(C) > w(M),$$

das heißt M ist nicht extrem. Widerspruch! □

PROF. DR. F. VALLENTIN, DR. A. GUNDELT, MATHEMATISCHES INSTITUT, UNIVERSITÄT ZU KÖLN,
WEYERTAL 86-90, 50931 KÖLN, DEUTSCHLAND
E-mail address: frank.vallentin@uni-koeln.de, anna.gundert@uni-koeln.de

KAPITEL 4 — FLÜSSE IN NETZWERKEN

F. VALLENTIN, A. GUNDERT

1. DAS MAX-FLOW-MIN-CUT THEOREM

Es sei $D = (V, A)$ ein gerichteter Graph, $s, t \in V$ zwei Knoten. Wir nennen s *Quelle* und t *Senke*.

Definition 1.1. Eine Funktion $f: A \rightarrow \mathbb{R}_{\geq 0}$ heißt *s-t-Fluss*, falls

$$\forall v \in V \setminus \{s, t\} : \sum_{a \in \delta^+(v)} f(a) = \sum_{a \in \delta^-(v)} f(a) \text{ gilt.}$$

(Flusserhaltungsgesetz)

Dabei schreiben wir kurz

$$\delta^+(v) = \delta^+(\{v\}) = \{(v, w) \in A : w \in V\},$$

$$\delta^-(v) = \delta^-(\{v\}) = \{(w, v) \in A : w \in V\}.$$

Der Wert eines s-t-Flusses ist

$$\text{value}(f) = \sum_{a \in \delta^+(s)} f(a) - \sum_{a \in \delta^-(s)} f(a)$$

Ein s-t-Fluss heißt beschränkt durch eine Kapazitätsfunktion $c: A \rightarrow \mathbb{R}_{\geq 0}$, falls

$$\forall a \in A : f(a) \leq c(a) \text{ gilt.}$$

(Notation: $f \leq c$).

Ziel: Finde einen maximalen s-t-Fluss.

Wir haben also das folgende Maximierungsproblem:

$$\begin{aligned} \max \quad & \text{value}(f) \\ & f: A \rightarrow \mathbb{R}_{\geq 0} \\ & f \text{ ist s-t-Fluss} \\ & f(a) \leq c(a) \quad \forall a \in A \end{aligned}$$

Lemma 1.2. Es seien ein gerichteter Graph $G = (V, A)$, Knoten $s, t \in V$ und eine Kapazitätsfunktion $c: A \rightarrow \mathbb{R}_{\geq 0}$ gegeben. Sei f ein s-t-Fluss, der durch c beschränkt ist. Sei $U \subseteq V$ mit $s \in U, t \notin U$, so dass $\delta^+(U)$ ein s-t-Schnitt ist. Dann gilt

$$(1) \quad \text{value}(f) \leq c(\delta^+(U)) = \sum_{a \in \delta^+(U)} c(a)$$

Es gilt Gleichheit in (1) genau dann, wenn

$$\begin{aligned} f(a) &= c(a) \quad \forall a \in \delta^+(U) \text{ und} \\ f(a) &= 0 \quad \forall a \in \delta^-(U) \end{aligned}$$

Beweis. Es gilt

$$\begin{aligned}
\text{value}(f) &= \sum_{a \in \delta^+(s)} f(a) - \sum_{a \in \delta^-(s)} f(a) \\
&= \sum_{v \in U} \left(\sum_{a \in \delta^+(v)} f(a) - \sum_{a \in \delta^-(v)} f(a) \right) \quad (\text{wg. Flusserhaltung}) \\
&= \sum_{a \in \delta^+(U)} f(a) - \sum_{a \in \delta^-(U)} f(a) \\
&\leq \sum_{a \in \delta^+(U)} c(a).
\end{aligned}$$

□

Definition 1.3. Es sei $D = (V, A)$ ein gerichteter Graph und $a = (u, v)$ eine Kante. Definiere

$$a^{-1} = (v, u) \text{ und } A^{-1} = \{a^{-1} : a \in A\}$$

Sei f ein s - t -Fluss und c eine Kapazitätsfunktion. Definiere den Residualgraph $D_f = (V, A_f)$ durch

$$A_f = \{a \in A : f(a) < c(a)\} \cup \{a^{-1} \in A^{-1} : f(a) > 0\}$$

Lemma 1.4. Sei f ein s - t -Fluss mit $f \leq c$. Angenommen D_f enthält keinen gerichteten s - t -Weg. Sei $U \subseteq V$ die Menge der Knoten, die in D_f von s aus erreichbar sind. Dann gilt

$$\text{value}(f) = c(\delta^+(U))$$

Insbesondere ist f maximal.

Beweis. Für $a \in \delta^+(U)$ gilt $a \notin A_f$, das heißt $f(a) = c(a)$. Für $a \in \delta^-(U)$ gilt $a^{-1} \notin A_f$, das heißt $f(a) = 0$. Also ist $\text{value}(f) = c(\delta^+(U))$ und f ist nach Lemma 1.2 maximal. □

Definition 1.5. Sei P ein gerichteter Weg in D_f . Definiere $\chi^P : A \rightarrow \mathbb{R}$ durch

$$\chi^P(a) = \begin{cases} 1, & \text{falls } P \text{ die Kante } a \text{ durchläuft} \\ -1, & \text{falls } P \text{ die Kante } a^{-1} \text{ durchläuft} \\ 0, & \text{sonst.} \end{cases}$$

Theorem 1.6. (Max-Flow = Min-Cut; Ford-Fulkerson, 1954)

Es seien ein gerichteter Graph $D = (V, A)$, zwei Knoten $s, t \in V$ und eine Kapazitätsfunktion $c : A \rightarrow \mathbb{R}_{\geq 0}$ gegeben. Dann gilt

$$\begin{array}{ll}
\max \text{value}(f) & = \min c(\delta^+(U)) \\
f \text{ ist } s\text{-}t\text{-Fluss} & U \subseteq V \\
f \leq c & s \in U, t \notin U
\end{array}$$

Beweis.

- max ≤ min: Wurde in Lemma 1.2 gezeigt.

- max \geq min: Sei f ein maximaler s - t -Fluss mit $f \leq c$. Zu zeigen ist

$$\text{value}(f) = c(\delta^+(U)) \text{ für ein } U \subseteq V \text{ mit } s \in U, t \notin U.$$

Betrachte D_f . Falls es keinen gerichteten s - t -Weg in D_f gibt, können wir Lemma 1.4 anwenden und sind fertig. Angenommen es gibt einen s - t -Weg P in D_f . Dann ist $f' = f + \varepsilon\chi^P$ für genügend kleines $\varepsilon > 0$ ebenfalls ein s - t -Fluss, der durch c beschränkt ist, mit

$$\text{value}(f') = \text{value}(f) + \varepsilon$$

Das ist ein Widerspruch zur Maximalität von f . □

Korollar 1.7. Falls c ganzzahlig ist, das heißt $c(a) \in \mathbb{Z} \forall a \in A$, dann gibt es einen ganzzahligen maximalen s - t -Fluss.

Beweis. Wähle $\varepsilon = 1$ im obigen Beweis. □

Algorithmus 1.8. (Ford-Fulkerson)

Input : Gerichteter Graph $D = (V, A)$, $s, t \in V$, $c: A \rightarrow \mathbb{R}_{\geq 0}$
Output : Maximaler s - t -Fluss f
 Setze $f = 0$
while \exists gerichteter s - t -Weg P in D_f **do**
 | $f = f + \varepsilon\chi^P$, wobei $\varepsilon > 0$ maximal gewählt, so dass $0 \leq f + \varepsilon\chi^P \leq c$ gilt
end

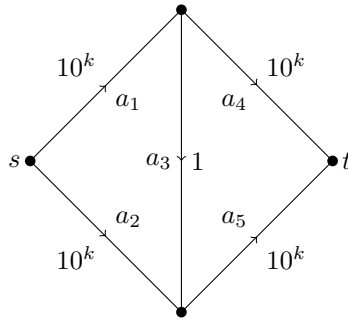
Satz 1.9. Falls $c(a) \in \mathbb{Q} \forall a \in A$, dann terminiert der Ford-Fulkerson Algorithmus in endlich vielen Schritten; sonst im Allgemeinen nicht.

Beweis. Wenn $c(a) \in \mathbb{Q} \forall a \in A$, dann existiert ein $K \in \mathbb{N}$ mit $Kc(a) \in \mathbb{N} \forall a \in A$. Das heißt in jeder Iteration ist ε ein Vielfaches von $\frac{1}{K}$ und der Wert des Flusses vergrößert sich mindestens um $\frac{1}{K}$. Da der Wert $c(\delta^+(s))$ nicht überschreiten kann, terminiert der Algorithmus nach höchstens $Kc(\delta^+(s))$ vielen Schritten. □

Es gibt Beispiele mit $c(a) \in \mathbb{R}$, so dass der Algorithmus nicht terminiert. (\rightarrow siehe Schrijver - CO - Buch, 10.4a)

2. VERBESSERUNG DES FORD-FULKERSON-ALGORITHMUS

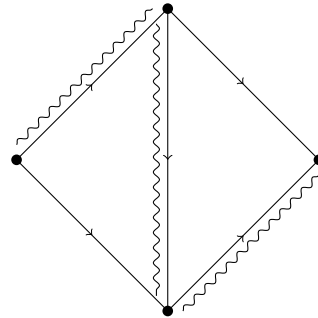
Beispiel 2.1. Die Anzahl der Iterationen im Ford-Fulkerson Algorithmus hängt von der Wahl der Wege ab, entlang derer der Flusswert erhöht wird. Dadurch, dass es keine Vorschrift gibt, nach der man solche Wege wählt, kann die Anzahl der Iterationen unnötig hoch werden, wie das folgende Beispiel zeigt.



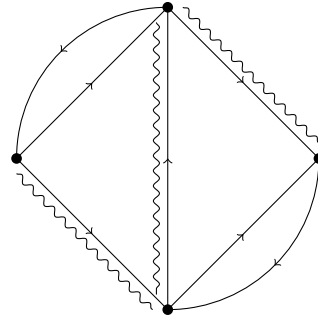
Offensichtlich ist $f = (f(a_1), f(a_2), f(a_3), f(a_4), f(a_5)) = (10^k, 10^k, 0, 10^k, 10^k)$ der s - t -Fluss mit maximalem Wert für dieses Netzwerk mit $\text{value}(f) = 2 \cdot 10^k$.

Aber: wählt man die Wege im Ford-Fulkerson Algorithmus sehr ungünstig, dann wird der Wert des Flusses in jeder Iteration nur um 1 erhöht. Man könnte also insgesamt bis zu $2 \cdot 10^k$ Iterationen durchführen, wie wir nun sehen werden (obwohl man im günstigsten Fall mit zwei Iterationen auskäme). Der Fluss wird immer entlang der Schlangenlinien verbessert:

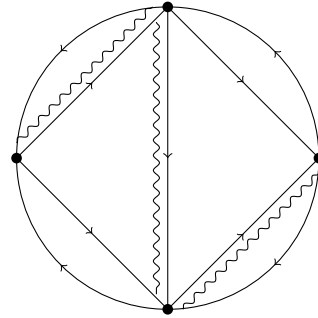
$$\underline{f_1 = (0, 0, 0, 0, 0)} \quad D_{f_1}:$$



$$\underline{f_2 = (1, 0, 1, 0, 1)} \quad D_{f_2}:$$



$$\underline{f_3 = (1, 1, 0, 1, 1)} \quad D_{f_3}:$$



$$f_4 = (2, 1, 1, 1, 2)$$

⋮

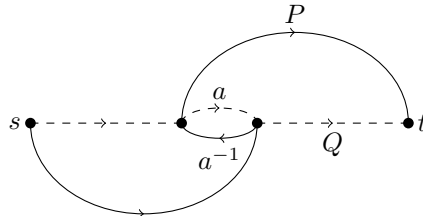
Und so weiter.

Problem: Die gewählten s - t -Wege in D_{f_k} waren nicht die kürzesten.

Lösung: (Dinits (1970), Edmonds-Karp (1972)) So etwas kann nicht passieren, wenn die gewählten s - t -Wege minimale Länge (bezogen auf die Anzahl der Kanten) haben.

Lemma 2.2. *Es sei $D = (V, A)$ ein gerichteter Graph, $s, t \in V$ zwei Knoten und $\mu(D) := \text{dist}(s, t)$. Sei $\alpha(D) \subseteq A$ die Menge von Kanten, die in einem kürzesten s - t -Weg vorkommen. Definiere $D' = (V, A \cup \alpha(D)^{-1})$. Dann gilt $\mu(D') = \mu(D)$ und $\alpha(D') = \alpha(D)$.*

Beweis. Es genügt zu zeigen, dass $\mu(D)$ und $\alpha(D)$ sich nicht verändern, wenn wir eine Kante $a^{-1} \in \alpha(D)^{-1}$ zu D hinzufügen. Angenommen das ist nicht so. Dann gibt es einen s - t -Weg $P \subseteq A \cup \{a^{-1}\}$, der durch a^{-1} geht und Länge höchstens $\mu(D)$ hat. Da $a \in \alpha(D)$, gibt es einen s - t -Weg $Q \subseteq A$, der durch a geht und Länge $\mu(D)$ hat. Wir haben also die folgende Situation:



Dann enthält $(P \cup Q) \setminus \{a, a^{-1}\}$ einen s - t -Weg der Länge $l < \mu(D)$. Widerspruch. \square

Satz 2.3. *Falls in jeder Iteration des Ford-Fulkerson Algorithmus ein kürzester s - t -Weg in D_f ausgewählt wird, beträgt die Anzahl der Iterationen höchstens $|V| \cdot |A|$. Insbesondere ist sie unabhängig von der Kapazitätsfunktion.*

Beweis. Falls wir einen Fluss f entlang eines kürzesten s - t -Weges P in D_f verbessern und einen neuen Fluss f' erhalten, dann ist die Kantenmenge von $D_{f'}$ in der Kantenmenge von $D' = (V, A_f \cup \alpha(D_f)^{-1})$ enthalten, deswegen gilt $\mu(D_{f'}) \geq \mu(D')$. Außerdem haben wir $\mu(D_f) = \mu(D')$ und $\alpha(D_f) = \alpha(D')$ nach Lemma 2.2.

Falls $\mu(D_{f'}) = \mu(D_f)$, dann gilt $\alpha(D_{f'}) \subseteq \alpha(D') = \alpha(D_f)$. Da wenigstens eine Kante aus P zu D_f aber nicht zu $D_{f'}$ gehört, haben wir strikte Inklusion $\alpha(D_{f'}) \subsetneq \alpha(D_f)$. Diese Situation ($\mu(D_{f'}) = \mu(D_f)$) kann maximal $|A|$ -mal vorkommen, danach muss $\mu(D_{f'}) > \mu(D_f)$ sein. Da $\mu(D_f)$ sich höchstens $|V|$ -mal vergrößern kann, folgt die Behauptung. \square

KAPITEL 5 — POLYEDERTHEORIE

F. VALLENTIN, A. GUNDERT

Viele Optimierungsprobleme des Operations Research lassen sich als *lineares Programm* formulieren:

$$\begin{array}{l} \text{gegeben: } c \in \mathbb{R}^n, \\ \quad a_1, \dots, a_m \in \mathbb{R}^n, \\ \quad b_1, \dots, b_m \in \mathbb{R} \\ \\ \text{gesucht: } \min c^T x \\ \quad x \in \mathbb{R}^n \\ \quad a_1^T x \leq b_1, \dots, a_m^T x \leq b_m. \end{array}$$

c definiert die *Zielfunktion* $x \mapsto c^T x$,

$a_j^T x \leq b_j$ sind lineare Ungleichungen, welche die *Nebenbedingungen* beschreiben.

Kurzform: $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ mit

$$A = \begin{bmatrix} - & a_1^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \quad \text{und} \quad Ax \leq b$$

Manchmal kommt noch eine Ganzzahligkeitsnebenbedingung $x \in \mathbb{Z}^n$ hinzu (\rightarrow spätere Kapitel).

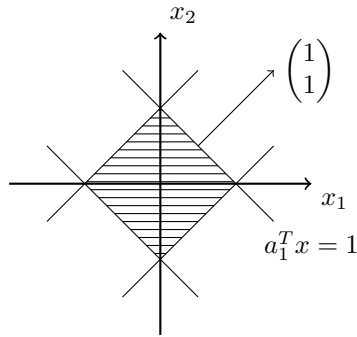
Dieses Kapitel: Wie sieht die Menge der *zulässigen Lösungen* (die Vektoren, die die Nebenbedingungen erfüllen)

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

geometrisch aus?

Beispiel 0.1.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$



Dies ist ein (konvexes) Polyeder.

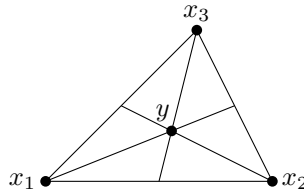
Polyedertheorie: Erweiterung der linearen Algebra über dem Körper \mathbb{R} durch lineare Ungleichungen.

1. KONVEXE MENGEN

Definition 1.1. Seien $x_1, \dots, x_N \in \mathbb{R}^n$ Punkte. Dann ist $y \in \mathbb{R}^n$ eine Konvexkombination von x_1, \dots, x_N , falls

$$y = \sum_{i=1}^N \alpha_i x_i \quad \text{mit } \alpha_1, \dots, \alpha_N \geq 0 \quad \text{und} \quad \sum_{i=1}^N \alpha_i = 1.$$

Beispiel 1.2. Der Schwerpunkt eines Dreiecks ist eine Konvexkombination der drei Eckpunkte.



$$y = \frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_3$$

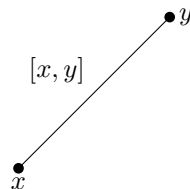
Definition 1.3. Eine Menge $C \subseteq \mathbb{R}^n$ heißt konvex, wenn sie unter der Bildung von Konvexkombinationen abgeschlossen ist, das heißt

$$\forall N \in \mathbb{N} \forall x_1, \dots, x_N \in C \forall \alpha_1, \dots, \alpha_N \geq 0, \sum_{i=1}^N \alpha_i = 1 \quad \Rightarrow \quad \sum_{i=1}^N \alpha_i x_i \in C.$$

Definition 1.4. Seien $x, y \in \mathbb{R}^n$ Punkte, dann ist

$$[x, y] = \{(1 - \alpha)x + \alpha y : \alpha \in [0, 1]\}$$

die Verbindungsstrecke zwischen x und y .



Satz 1.5. Eine Menge $C \subseteq \mathbb{R}^n$ ist konvex $\Leftrightarrow \forall x, y \in C : [x, y] \subseteq C$.

Beweis.

„ \Rightarrow “: Klar, denn dies ist ein Spezialfall der Definition mit $N = 2$.

„ \Leftarrow “: Zeige per Induktion nach N : Jede Konvexkombination von höchstens N Punkten aus C liegt wieder in C .

$N = 1$: \checkmark

$N \rightarrow N + 1$: Sei $y = \sum_{i=1}^{N+1} \alpha_i x_i$ mit $x_i \in C$, $\alpha_i \geq 0$ und $\sum_{i=1}^{N+1} \alpha_i = 1$.

1.Fall: $\alpha_{N+1} = 1$.

Dann ist $\alpha_1 = \dots = \alpha_N = 0$ und $y = x_{N+1} \in C$.

2.Fall: $\alpha_{N+1} \neq 1$.

Nach Induktionsvoraussetzung ist

$$y' = \sum_{i=1}^N \frac{\alpha_i}{1 - \alpha_{N+1}} x_i \in C, \quad \text{weil } \frac{\alpha_i}{1 - \alpha_{N+1}} \geq 0 \text{ und } \sum_{i=1}^N \frac{\alpha_i}{1 - \alpha_{N+1}} = 1$$

Desweiteren ist

$$y = (1 - \alpha_{N+1})y' + \alpha_{N+1}x_{N+1} \in [y', x_{N+1}] \subseteq C.$$

□

Beispiel 1.6.

a)  ist keine konvexe Menge.

b) Sei $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ eine Norm. Sei

$$K = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$$

die zugehörige Einheitskugel.

Behauptung: K ist konvex.

Beweis: Seien $x, y \in K$, $\alpha \in [0, 1]$. Dann ist

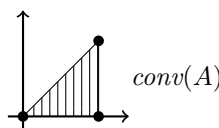
$$\|(1 - \alpha)x + \alpha y\| \leq (1 - \alpha)\|x\| + \alpha\|y\| \leq (1 - \alpha) \cdot 1 + \alpha \cdot 1 = 1.$$

Definition 1.7. Sei $A \subseteq \mathbb{R}^n$. Die konvexe Hülle von A ist

$$\text{conv } A = \bigcap_{\substack{B \supseteq A \\ B \text{ konvex}}} B.$$

Da der Durchschnitt von konvexen Mengen wieder konvex ist, handelt es sich bei $\text{conv } A$ um die inklusionsminimale konvexe Menge, die A enthält.

Beispiel 1.8. $A = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$



Satz 1.9. Sei $A \subseteq \mathbb{R}^n$. Dann gilt

$$\text{conv } A = \{y \in \mathbb{R}^n : \exists N \in \mathbb{N} \exists x_1, \dots, x_N \in A \exists \alpha_1, \dots, \alpha_N \geq 0, \sum_{i=1}^N \alpha_i = 1, y = \sum_{i=1}^N \alpha_i x_i\}.$$

Das heißt, $\text{conv } A$ besteht aus allen Konvexkombinationen der Menge A .

Beweis.

„ \subseteq “: trivial.

„ \supseteq “: Sei B konvex mit $B \supseteq A$. Da B abgeschlossen unter der Bildung von Konvexkombinationen ist, muss für jede Wahl von $N \in \mathbb{N}$, $x_1, \dots, x_N \in A$, $\alpha_1, \dots, \alpha_N \geq 0$, $\sum_{i=1}^N \alpha_i = 1$ stets $y = \sum_{i=1}^N \alpha_i x_i \in B$ gelten.

□

Definition 1.10. Eine Menge $P \subseteq \mathbb{R}^n$ heißt (konvexes) Polytop, falls es eine endliche Menge $A = \{x_1, \dots, x_N\}$ gibt mit

$$\begin{aligned} P &= \text{conv } A \\ &= \left\{ \sum_{i=1}^N \alpha_i x_i : \alpha_i \geq 0, \sum_{i=1}^N \alpha_i = 1 \right\} \end{aligned}$$

Sei $y \in \text{conv } A$, das heißt

$$y = \sum_{i=1}^N \alpha_i x_i \text{ für } N \in \mathbb{N}, x_1, \dots, x_N \in A, \alpha_1, \dots, \alpha_N \geq 0, \sum_{i=1}^N \alpha_i = 1.$$

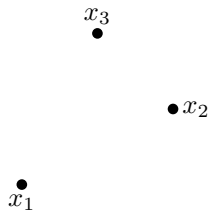
Frage: Wie groß muss N höchstens sein, damit man jedes y in $\text{conv } A$ darstellen kann?

Definition 1.11. Die Punkte $x_1, \dots, x_N \in \mathbb{R}^n$ sind affin unabhängig, falls

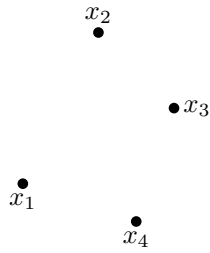
$$\forall \alpha_1, \dots, \alpha_N : \sum_{i=1}^N \alpha_i = 0, \sum_{i=1}^N \alpha_i x_i = 0 \Rightarrow \alpha_1 = \dots = \alpha_N = 0.$$

In anderen Worten: Die Vektoren $\begin{bmatrix} 1 \\ x_1 \end{bmatrix}, \dots, \begin{bmatrix} 1 \\ x_N \end{bmatrix} \in \mathbb{R}^{n+1}$ sind linear unabhängig.

Beispiel 1.12. Drei Punkte in der Ebene, die nicht alle auf der selben Gerade liegen, sind affin unabhängig.



Vier Punkte in der Ebene sind hingegen nie affin unabhängig.



Definition 1.13. Sei $A \subseteq \mathbb{R}^n$. Die Dimension von A ist

$$\dim A = \max\{N - 1 : \exists x_1, \dots, x_N \in A \text{ affin unabhängig}\}.$$

Es ist immer $\dim A \leq n$, weil $\text{rank} \begin{bmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_N \end{bmatrix} \leq n + 1$.

Satz 1.14. (Caratheodory)

Sei $A \subseteq \mathbb{R}^n$ und $y \in \text{conv } A$. Dann existieren $x_1, \dots, x_N \in A$ affin unabhängig mit $y \in \text{conv} \{x_1, \dots, x_N\}$. Insbesondere ist $N \leq \dim(A) + 1 \leq n + 1$.

Beweis. Sei $y \in \text{conv } A$ mit

$$y = \sum_{i=1}^N \alpha_i x_i, \quad x_i \in A, \alpha_i \geq 0, \sum_{i=1}^N \alpha_i = 1.$$

Angenommen N ist minimal und x_1, \dots, x_N sind nicht affin unabhängig. Dann gibt es $\beta_1, \dots, \beta_N \in \mathbb{R}$ mit $\sum_{i=1}^N \beta_i = 0$, $\sum_{i=1}^N \beta_i x_i = 0$ und $(\beta_1, \dots, \beta_N) \neq (0, \dots, 0)$. Sei oBdA. $\beta_1 > 0$ und $\frac{\alpha_1}{\beta_1}$ so klein wie möglich. Betrachte die Darstellung

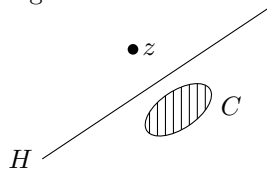
$$y = \sum_{i=1}^N \alpha_i x_i - \frac{\alpha_1}{\beta_1} \sum_{i=1}^N \beta_i x_i = \sum_{i=2}^N \gamma_i x_i, \quad \gamma_i = \alpha_i - \frac{\alpha_1}{\beta_1} \beta_i.$$

Da $\gamma_i \geq 0$ (klar, wenn $\beta_i \leq 0$; sonst wegen der Minimalität von $\frac{\alpha_1}{\beta_1}$) und $\sum_{i=2}^N \gamma_i = 1$, ist die Darstellung ein Widerspruch zur Minimalität von N . \square

2. TRENN- UND STÜTZHYPEREBENEN

Sei $C \subseteq \mathbb{R}^n$ eine abgeschlossene, konvexe Menge.

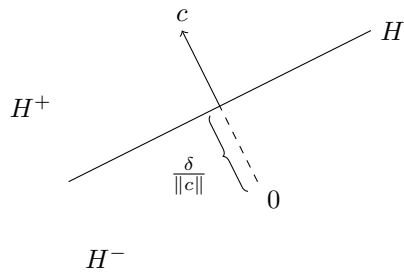
Fundamentale Eigenschaft: Jeder Punkt $z \notin C$ kann durch eine Hyperebene von C getrennt werden.



(Wichtige Verallgemeinerung in der Funktionalanalysis: Satz von Hahn-Banach)

Definition 2.1. Eine Teilmenge $H \subseteq \mathbb{R}^n$ heißt (affine) Hyperebene, falls es einen Vektor $c \in \mathbb{R}^n \setminus \{0\}$ und ein $\delta \in \mathbb{R}$ gibt mit

$$H = \{x \in \mathbb{R}^n : c^T x = \delta\}$$



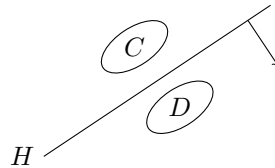
Die abgeschlossenen und konvexen Mengen

$$H^+ = \{x \in \mathbb{R}^n : c^T x \geq \delta\}$$

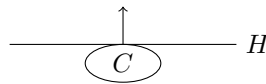
$$H^- = \{x \in \mathbb{R}^n : c^T x \leq \delta\}$$

heißen Halbräume

Definition 2.2. Seien $C, D \subseteq \mathbb{R}^n$. Eine Hyperebene H heißt Trennhyperebene von C und D , falls $C \subseteq H^-$ und $D \subseteq H^+$ (oder umgekehrt).

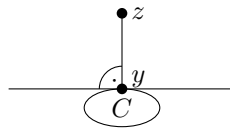


Definition 2.3. Sei $C \subseteq \mathbb{R}^n$. Eine Hyperebene H heißt Stützhyperebene von C , falls $C \subseteq H^-$ und $C \cap H \neq \emptyset$.



Frage: Wie findet man Trenn- bzw. Stützhyperebenen?

Antwort: Verwende *metrische Projektion*
 (= beste Approximation von $z \notin C$ in C
 = orthogonale Projektion, falls C affin linear ist).

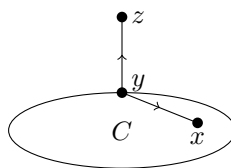


Lemma 2.4. Sei $C \subseteq \mathbb{R}^n$ eine abgeschlossene und konvexe Menge, die nicht leer ist. Sei $z \notin C$. Dann

$$\exists! y \in C : \|y - z\| = \inf_{x \in C} \|x - z\| \quad (\text{Notation: } y = \pi_C(z))$$

und

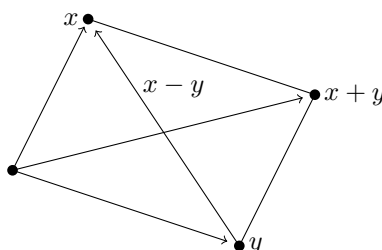
$$(z - y)^T (x - y) \leq 0 \quad \text{für alle } x \in C.$$



Beweis. (funktioniert auch in allgemeinen Hilberträumen)

Zur Erinnerung: Parallelogrammgleichung

Für $x, y \in \mathbb{R}^n$ gilt $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$.



Existenz von y : Sei $d = \inf_{x \in C} \|x - z\|$. OBdA ist $z = 0$ (durch Translation). Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge aus C mit $\lim_{n \rightarrow \infty} \|x_n\| = d$. Wir zeigen, dass $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge ist. Nach der Parallelogrammgleichung ist

$$\underbrace{\left\| \frac{x_n + x_m}{2} \right\|^2}_{\substack{\in C \\ \Rightarrow \geq d^2}} + \underbrace{\left\| \frac{x_n - x_m}{2} \right\|^2}_{\Rightarrow \rightarrow 0} = \underbrace{\frac{1}{2} \|x_n\|^2 + \frac{1}{2} \|x_m\|^2}_{\rightarrow d^2}$$

Der zweite Summand auf der linken Seite muss gegen 0 konvergieren, da die rechte Seite gegen d^2 konvergiert und der erste Summand auf der linken Seite sicher nicht kleiner als d^2 ist (denn $d = \inf_{x \in C} \|x - z\|$). Also ist $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge und weil \mathbb{R}^n vollständig ist, konvergiert $(x_n)_{n \in \mathbb{N}}$. Weil C abgeschlossen ist, liegt der Grenzwert $\lim_{n \rightarrow \infty} x_n$ ebenfalls in C . Dieser Grenzwert ist ein y mit der gewünschten Eigenschaft.

Eindeutigkeit von y : Angenommen es existieren $y, y' \in C$ mit $\|y\| = \|y'\| = \inf_{x \in C} \|x\|$ und $y \neq y'$. Dann ist

$$\left\| \frac{y + y'}{2} \right\|^2 < \left\| \frac{y + y'}{2} \right\|^2 + \left\| \frac{y - y'}{2} \right\|^2 = \frac{1}{2} \|y\|^2 + \frac{1}{2} \|y'\|^2 = d^2,$$

was ein Widerspruch zur Minimalität von d ist.

Zusatz: $(z - y)^T(x - y) \leq 0$ für alle $x \in C$.

Sei $\alpha \in [0, 1]$. Dann ist

$$\begin{aligned} \|z - y\|^2 &\leq \|z - ((1 - \alpha)y + \alpha x)\|^2 \\ &= \|z - y + \alpha(y - x)\|^2 \\ &= \|z - y\|^2 + 2\alpha(z - y)^T(y - x) + \alpha^2\|y - x\|^2 \end{aligned}$$

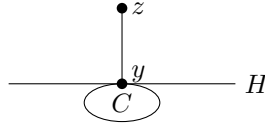
Demnach gilt für $\alpha \in (0, 1]$

$$(z - y)^T(x - y) \leq \frac{\alpha}{2}\|y - x\|^2$$

und die Behauptung folgt, da die rechte Seite beliebig klein ≥ 0 werden kann. \square

Theorem 2.5. Sei $C \in \mathbb{R}^n$ eine nichtleere, abgeschlossene und konvexe Menge. Sei $z \notin C$ ein Punkt außerhalb von C . Dann gibt es eine Trennhyperebene von $\{z\}$ und C .

Beweis. Definiere die Hyperebene $H = \{x \in \mathbb{R}^n : c^T x = \delta\}$ mit $c = z - y$ und $\delta = c^T y$, wobei $y = \pi_C(z)$.



Behauptung: a) $z \in H^+$
b) $C \subseteq H^-$.

zu a)

$$\begin{aligned} c^T z &\geq \delta \\ \Leftrightarrow (z - y)^T z &\geq (z - y)^T y \\ \Leftrightarrow (z - y)^T(z - y) &\geq 0. \quad \checkmark \end{aligned}$$

zu b) Für $x \in C$ ist

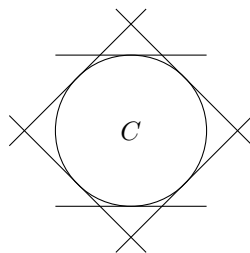
$$\begin{aligned} c^T x &\leq \delta \\ \Leftrightarrow (z - y)^T x &\leq (z - y)^T y \\ \Leftrightarrow (z - y)^T(x - y) &\leq 0. \quad \checkmark \end{aligned}$$

\square

Bemerkung 2.6. Die im Beweis konstruierte Hyperebene ist eine Stützhyperebene von C . Falls $\delta \in (c^T y, c^T z)$ gewählt wäre, wäre H eine strikte Trennhyperebene von $\{z\}$ und C , das heißt $z \in H^+$ und $z \notin H$, $C \subseteq H^-$ und $C \cap H = \emptyset$.

Korollar 2.7. Sei $C \subseteq \mathbb{R}^n$ eine nichtleere, abgeschlossene und konvexe Menge. Dann gilt

$$C = \bigcap_{H \text{ Stützhyperebene von } C} H^-$$



Beweis. → Aufgabe 7.1 □

Definition 2.8. Eine Menge $P \subseteq \mathbb{R}^n$ heißt Polyeder, falls es eine Matrix $A \in \mathbb{R}^{m \times n}$ und einen Vektor $b \in \mathbb{R}^m$ gibt, so dass

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

gilt.

Ein Polyeder entspricht also der Lösungsmenge eines Systems linearer Ungleichungen, was der Durchschnitt endlich vieler Halbräume ist (im Korollar 2.7 genügen endlich viele Stützhyperebenen, um ein Polyeder vollständig zu beschreiben).

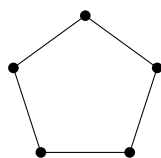
Jetzt: Verhältnis zwischen Polyedern und Polytopen (konvexe Hülle endlich vieler Punkte).

3. EXTREMPUNKTE UND ECKEN

Definition 3.1. Sei $C \subseteq \mathbb{R}^n$ eine konvexe Menge. Ein Punkt $z \in C$ heißt Extrempunkt von C , falls für alle $x, y \in C$ mit $z = \alpha x + (1 - \alpha)y$ und $\alpha \in (0, 1)$ stets $x = z = y$ folgt.

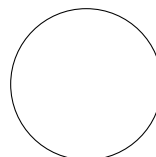
Beispiel 3.2. :

a)



5 Extrempunkte

b)



∞ -viele Extrempunkte

Sprechweise: Extrempunkte von Polyedern heißen *Ecken*.

Notation 3.3. Sei $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ ein Polyeder und $z \in P$. Mit A_z bezeichnen wir die Teilmatrix von A , die die Zeilen enthält mit $a_i^T z = b_i$ („aktive Ungleichungen“).

Satz 3.4. Sei $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ ein Polyeder und $z \in P$. Dann gilt z ist eine Ecke von $P \Leftrightarrow \text{rang } A_z = n$.

Beweis.

„ \Rightarrow “: Angenommen $\text{rang } A_z < n$. Dann existiert ein $c \in \mathbb{R}^n$, $c \neq 0$ mit $A_z c = 0$. Da $a_i^T z < b_i$ für alle Zeilen von A gilt, die nicht zu A_z gehören, gibt es ein $\delta > 0$ mit

$$a_i^T(z + \delta c) \leq b_i \quad \text{und} \quad a_i^T(z - \delta c) \leq b_i.$$

Also: $A(z + \delta c) \leq b$ und $A(z - \delta c) \leq b$ und $z + \delta c, z - \delta c \in P$. Das heißt $z = \frac{1}{2}(z + \delta c) + \frac{1}{2}(z - \delta c)$ und somit ist z keine Ecke von P .

„ \Leftarrow “: Es sei $z = \alpha x + (1 - \alpha)y$ mit $x, y \in P, \alpha \in (0, 1)$. Für eine Zeile a_i^T von A_z gilt

$$\begin{aligned} b_i &= a_i^T z \\ &= a_i^T (\alpha x + (1 - \alpha)y) \\ &= \alpha a_i^T x + (1 - \alpha) a_i^T y \\ &\leq \alpha b_i + (1 - \alpha) b_i = b_i \end{aligned}$$

Also $a_i^T x = a_i^T y = b_i$, da $\alpha \in (0, 1)$. Da $\text{rang } A_z = n$, hat das System $a_i^T w = b_i$, das aus den Zeilen von A_z besteht, eine eindeutige Lösung. Das heißt $x = z = y$ und z ist eine Ecke von P . □

Korollar 3.5. Ein Polyeder $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ mit $A \in \mathbb{R}^{m \times n}$ hat höchstens $\binom{m}{n}$ Ecken.

Theorem 3.6. (Minkowski)

Ein beschränktes Polyeder ist ein Polytop

Beweis. Seien x_1, \dots, x_t die Ecken des beschränkten Polyeders

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

Behauptung: $P = \text{conv}\{x_1, \dots, x_t\}$.

Beweis:

„ \supseteq “: klar.

„ \subseteq “: Sei $z \in P$. Wir zeigen: $z \in \text{conv}\{x_1, \dots, x_t\}$ per Induktion nach $n - \text{rang}(A_z)$.

Induktionsanfang: $n - \text{rang}(A_z) = 0$.

Das heißt $\text{rang}(A_z) = n \Rightarrow z$ ist eine Ecke von P , das heißt $z \in \{x_1, \dots, x_t\}$.

Induktionsschluss: $n - \text{rang}(A_z) > 0$.

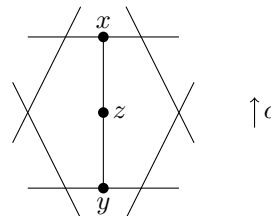
Das heißt $\text{rang}(A_z) < n \Rightarrow \exists c \neq 0 : A_z c = 0$. Sei

$$\mu_0 = \max\{\mu : z + \mu c \in P\}$$

$$\nu_0 = \min\{\nu : z + \nu c \in P\}.$$

μ_0 und ν_0 existieren, weil P kompakt ist. Definiere

$$x = z + \mu_0 c, \quad y = z + \nu_0 c$$



Es gilt

$$(*) \quad \mu_0 = \min \left\{ \frac{b_i - a_i^T z}{a_i^T c} : a_i^T \text{ Zeile von } A, a_i^T c > 0 \right\},$$

weil $a_i^T(z + \mu c) \leq b_i \Leftrightarrow \mu \leq \frac{b_i - a_i^T z}{a_i^T c}$ und μ_0 ist das größte solche μ . In $(*)$ sei i_0 eine Zeile von A , in der das Minimum angenommen wird. Dann ist

$$(1) \quad A_z x = A_z z + \mu_0 A_z c = A_z z$$

$$(2) \quad a_{i_0}^T x = a_{i_0}^T(z + \mu_0 c) = b_{i_0}.$$

Somit enthält A_x alle Zeilen von A_z und außerdem die Zeile $a_{i_0}^T$. Da $A_z c = 0$, aber $a_{i_0}^T c \neq 0$, gilt $\text{rang}(A_x) > \text{rang}(A_z)$. Nach Induktionsvoraussetzung ist $x \in \text{conv}\{x_1, \dots, x_t\}$. Genauso folgt $y \in \text{conv}\{x_1, \dots, x_t\}$. Da $z \in [x, y]$ ist, folgt die Behauptung $z \in \text{conv}\{x_1, \dots, x_t\}$. □

Definition 3.7. Sei $A \subseteq \mathbb{R}^n$. Dann heißt

$$A^* = \{y \in \mathbb{R}^n : x^T y \leq 1 \forall x \in A\}$$

die Polare von A .

Lemma 3.8.

- a) Für $\alpha > 0$ gilt $(\alpha A)^* = \frac{1}{\alpha} A^*$, wobei $\beta B = \{\beta b : b \in B\}$ mit $\beta \in \mathbb{R}$, $B \subseteq \mathbb{R}^n$.
- b) $A \subseteq B \Rightarrow A^* \supseteq B^*$.
- c) $(B_n)^* = B_n$, wobei $B_n = \{x \in \mathbb{R}^n : x^T x \leq 1\}$ die n -dimensionale Einheitskugel ist.
- d) $P = \text{conv}\{x_1, \dots, x_t\} \Rightarrow P^* = \{y \in \mathbb{R}^n : x_1^T y \leq 1, \dots, x_t^T y \leq 1\}$.

Beweis. a), b) \rightarrow Aufgabe 7.4.

c) „ \subseteq “: Sei $y \in (B_n)^*$. Falls $y = 0$, dann $y \in B_n$. Falls $y \neq 0$, dann $y^T \frac{1}{\|y\|} y \leq 1$. Also $\|y\| \leq 1$, das heißt $y \in B_n$.

„ \supseteq “: Seien $x, y \in B_n$. Dann gilt (nach Cauchy-Schwarz): $x^T y \leq \|x\| \|y\| \leq 1$, also $x \in (B_n)^*$.

d) „ \subseteq “: klar.

„ \supseteq “: $y \in \mathbb{R}^n$ erfülle $x_1^T y \leq 1, \dots, x_t^T y \leq 1$. Sei $z \in P$. Dann ist

$$z = \sum_{i=1}^t \alpha_i x_i \text{ mit } \alpha_i \geq 0, \sum_{i=1}^t \alpha_i = 1.$$

Es ist

$$z^T y = \left(\sum_{i=1}^t \alpha_i x_i \right)^T y = \sum_{i=1}^t \alpha_i x_i^T y \leq \sum_{i=1}^t \alpha_i = 1$$

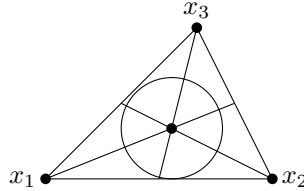
□

Theorem 3.9. (Weyl)

Ein Polytop ist ein beschränktes Polyeder.

Beweis. Sei $P = \text{conv}\{x_1, \dots, x_t\}$ ein Polytop im \mathbb{R}^n . Es sei
 oBdA. $0 \in P$ (durch Translation),
 oBdA. $\dim P = n$ (durch Einschränkung auf den Span von P),
 oBdA. x_1, \dots, x_{n+1} affin unabhängig (durch Ummumerierung).

Für den Schwerpunkt $x_0 = \frac{1}{n+1}(x_1 + \dots + x_{n+1})$ gibt es ein $r > 0$, so dass $B(x_0, r) \subseteq P$, wobei $B(x_0, r) = \{y \in \mathbb{R}^n : \|y - x_0\| \leq r\}$.



Sei oBdA. $x_0 = 0$ nach Translation. Nach (a) - (c) von Lemma 3.8 gilt $P^* \subseteq B(x_0, \frac{1}{r})$. Nach (d) ist P^* ein beschränktes Polyeder und nach dem Theorem von Minkowski ist P^* somit ein Polytop. Das heißt

$$P^* = \text{conv}\{y_1, \dots, y_s\} \text{ für } y_1, \dots, y_s \in \mathbb{R}^n.$$

Behauptung: $P = \{x \in \mathbb{R}^n : y_1^T x \leq 1, \dots, y_s^T x \leq 1\}$.

Beweis:

„ \subseteq “: Sei $x \in P$. Dann gilt $y_i^T x \leq 1$, weil $y_i \in P^*$.

„ \supseteq “: Sei $x \in \mathbb{R}^n$ mit $y_1^T x \leq 1, \dots, y_s^T x \leq 1$. Angenommen $x \notin P$. Dann gibt es eine Trennhyperebene von $\{x\}$ und P , das heißt

$$\exists c \neq 0, \delta \in \mathbb{R} : \quad c^T x > \delta \text{ und } c^T x' \leq \delta \quad \forall x' \in P.$$

Da $0 \in P$, ist $\delta > 0$. Wähle $\delta = 1$ (durch Skalieren von c). Dann ist $c \in P^*$ und somit $c = \sum_{j=1}^s \alpha_j y_j$ mit $\alpha_j \geq 0$ und $\sum_{j=1}^s \alpha_j = 1$. Desweiteren gilt

$$1 < c^T x = \left(\sum_{j=1}^s \alpha_j y_j \right)^T x \leq \sum_{j=1}^s \alpha_j = 1.$$

Das ist ein Widerspruch. Somit gilt $x \in P$. □

Korollar 3.10. (*Theorem von Minkowski-Weyl*)

Sei $P \subseteq \mathbb{R}^n$. Dann gilt: P ist ein beschränktes Polyeder $\Leftrightarrow P$ ist ein Polytop.

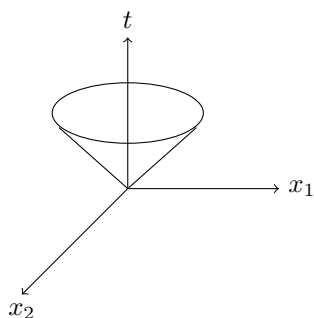
Als nächstes beschäftigen wir uns mit der Frage, wie allgemeine, eventuell unbeschränkte, Polyeder aussehen.

Definition 3.11. Eine Menge $C \subseteq \mathbb{R}^n$ heißt konvexer Kegel, falls

$$\forall \lambda, \mu \geq 0 \quad \forall x, y \in C : \quad \lambda x + \mu y \in C.$$

(Insbesondere ist C tatsächlich konvex.)

Beispiel 3.12. $\mathcal{L}^{n+1} = \{(x, t) \in \mathbb{R}^{n+1} : \|x\| \leq t\}$ ist ein konvexer Kegel („ice cream cone“).



Definition 3.13. (konische Hülle)

Sei $A \subseteq \mathbb{R}^n$. Die konische Hülle von A ist

$$\text{cone } A = \bigcap_{\substack{B \supseteq A \\ B \text{ konvexer Kegel}}} B$$

Satz 3.14. Sei $A \subseteq \mathbb{R}^n$, dann gilt

$$\text{cone } A = \{y \in \mathbb{R}^n : \exists N \in \mathbb{N}, \lambda_1, \dots, \lambda_N \geq 0, x_1, \dots, x_N \in A : y = \sum_{i=1}^N \lambda_i x_i\}$$

Definition 3.15. Ein konvexer Kegel $C \subseteq \mathbb{R}^n$ heißt endlich erzeugt, falls es $y_1, \dots, y_t \in \mathbb{R}^n$ gibt mit

$$C = \text{cone}\{y_1, \dots, y_t\}.$$

Satz 3.16. (Minkowski-Weyl für Kegel)

Ein konvexer Kegel $C \subseteq \mathbb{R}^n$ ist endlich erzeugt \Leftrightarrow

$$\exists A \in \mathbb{R}^{m \times n} : C = \{x \in \mathbb{R}^n : Ax \leq 0\}.$$

Beweis. Siehe zum Beispiel Chapter 7 in Schrijver - Theory of linear and integer programming. □

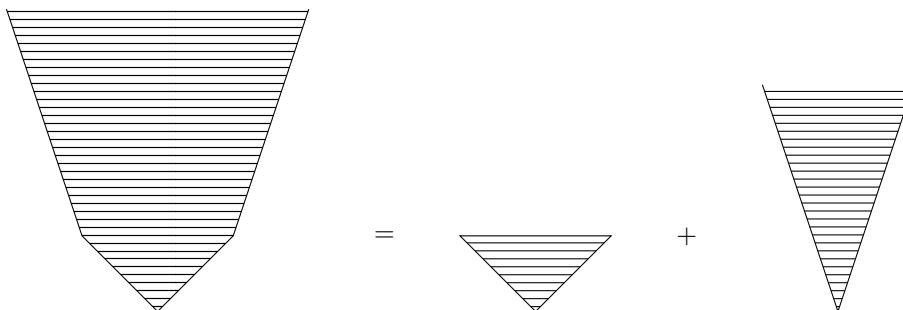
Theorem 3.17. Sei $P \subseteq \mathbb{R}^n$. P ist ein Polyeder \Leftrightarrow

$$\exists x_1, \dots, x_s, y_1, \dots, y_t \in \mathbb{R}^n : P = \text{conv}\{x_1, \dots, x_s\} + \text{cone}\{y_1, \dots, y_t\},$$

wobei $A + B = \{a + b : a \in A, b \in B\}$.

Beweis. Siehe zum Beispiel Kapitel 7 in Schrijver. □

Beispiel 3.18.



4. LEMMA VON FARKAS

Ein Lösbarkeitskriterium für lineare Gleichungssysteme

$$Ax = b \text{ mit } A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$$

ist durch den folgenden Satz gegeben.

Satz 4.1. $\exists x \in \mathbb{R}^n : Ax = b \Leftrightarrow \nexists y \in \mathbb{R}^m : y^T A = 0 \text{ und } y^T b \neq 0.$

Beweis.

„ \Rightarrow “: Angenommen $\exists x \in \mathbb{R}^n : Ax = b$ und $\exists y \in \mathbb{R}^m : y^T A = 0, y^T b \neq 0$. Dann

$$\underbrace{y^T A}_{=0} x = \underbrace{y^T b}_{\neq 0}.$$

Widerspruch.

„ \Leftarrow “: Angenommen $Ax = b$ hat keine Lösung. Dann gilt $\text{rang}[A|b] > \text{rang } A$.
Wähle $y \in \ker A^T$ mit $y^T b \neq 0$.

□

Wir erhalten ein analoges Resultat für Systeme linearer Ungleichungen.

Satz 4.2. (*Farkas Lemma*)

Seien $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. Dann

$\exists x \geq 0 : Ax = b \Leftrightarrow \nexists y \in \mathbb{R}^m : y^T A \geq 0 \text{ und } y^T b < 0.$

Beweis.

„ \Rightarrow “: Angenommen $\exists x \geq 0 : Ax = b$ und $\exists y \in \mathbb{R}^m : y^T A \geq 0, y^T b < 0$. Dann

$$y^T Ax = y^T b < 0, \text{ aber } \underbrace{(y^T A)}_{\geq 0} \underbrace{x}_{\geq 0} \geq 0.$$

Widerspruch.

„ \Leftarrow “: Angenommen $\nexists x \geq 0 : Ax = b$. Seien $a_1, \dots, a_n \in \mathbb{R}^m$ die Spaltenvektoren von A . Nach Voraussetzung ist $b \notin \text{cone}\{a_1, \dots, a_n\}$. Also gibt es eine Trennhyperebene von $\{b\}$ und $\text{cone}\{a_1, \dots, a_n\}$. Das heißt

$$\exists y \neq 0 : y^T b < 0 \text{ und } y^T x \geq 0 \forall x \in \text{cone}\{a_1, \dots, a_n\}.$$

□

Interpretation: Theorie der Alternativen.

Entweder: $\exists x \geq 0 : Ax = b$

oder: $\exists y : y^T A \geq 0, y^T b < 0.$

Korollar 4.3. (*Variante von Farkas Lemma*)

$\exists x \in \mathbb{R}^n : Ax \leq b \Leftrightarrow \nexists y \geq 0 : y^T A = 0, y^T b < 0.$

Beweis. Betrachte die Matrix

$$A' = [A \mid -A \mid I_m] \in \mathbb{R}^{m \times (n+n+m)},$$

wobei I_m die $m \times m$ Einheitsmatrix ist.

$$\text{Sei } x' = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^{n+n+m} \text{ mit } x_1, x_2 \in \mathbb{R}^n, x_3 \in \mathbb{R}^m.$$

$$\begin{aligned} \exists x' \geq 0 : A'x' = b &\Leftrightarrow \exists x_1, x_2, x_3 \geq 0 : Ax_1 - Ax_2 + x_3 = b \\ &\Leftrightarrow \exists x_1, x_2, x_3 \geq 0 : A(x_1 - x_2) + x_3 = b \\ &\Leftrightarrow \exists x \in \mathbb{R}^n : Ax \leq b. \end{aligned}$$

Wende Farkas Lemma auf $A'x' = b$ an:

$$\exists x' \geq 0 : A'x' = b \Leftrightarrow \nexists y \in \mathbb{R}^m : y^T[A] - A[I_m] \geq 0 \text{ und } y^T b < 0.$$

Dabei

$$\begin{aligned} y^T[A] - A[I_m] \geq 0 \in \mathbb{R}^{n+n+m} &\Leftrightarrow y^T A \geq 0, y^T(-A) \geq 0, y^T I_m \geq 0 \\ &\Leftrightarrow y^T A = 0, y \geq 0 \end{aligned}$$

□

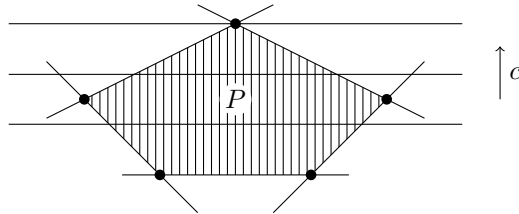
5. LINEARE PROGRAMMIERUNG

Ein *lineares Programm* (LP) in Standardform ist

$$\begin{aligned} \text{(LP)} \quad &\max c^T x \\ &x \in \mathbb{R}^n \\ &Ax \leq b, \end{aligned}$$

wobei $c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ gegeben sind.

Geometrische Interpretation: Maximiere die lineare Funktion $x \mapsto c^T x$ über dem Polyeder $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. Das heißt man schiebt eine zu c orthogonale Hyperebene so weit in Richtung c , dass sie P gerade noch schneidet.



Sprechweise:

x heißt *gültige/zulässige Lösung* $\Leftrightarrow x \in P$.

x heißt *optimale Lösung* $\Leftrightarrow x \in P$ und $c^T x = \max\{c^T y : y \in P\}$.

LP heißt *unbeschränkt* $\Leftrightarrow \sup\{c^T x : x \in P\} = +\infty$.

LP heißt *ungültig/unzulässig* $\Leftrightarrow \nexists x \in P$.

Satz 5.1. Falls die Menge der zulässigen Lösungen P ein nicht-leeres, beschränktes Polyeder ist, dann gibt es eine Ecke von P , die eine optimale Lösung von LP ist.

Beweis. Seien x_1, \dots, x_t die Ecken von P . Nach dem Theorem von Minkowski ist $P = \text{conv}\{x_1, \dots, x_t\}$. Angenommen keine Ecke ist eine optimale Lösung. Das heißt (und weil P kompakt ist):

$$\exists y \in P : c^T y > c^T x_i \text{ für alle } i = 1, \dots, t.$$

Schreibe $y = \sum_{i=1}^t \alpha_i x_i$ mit $\alpha_i \geq 0, \sum_{i=1}^t \alpha_i = 1$. Dann ist

$$c^T y = c^T \left(\sum_{i=1}^t \alpha_i x_i \right) = \sum_{i=1}^t \alpha_i c^T x_i < \sum_{i=1}^t \alpha_i c^T y = c^T y.$$

Widerspruch. □

Daraus ergibt sich ein offensichtlicher (aber eventuell sehr langsamer) Algorithmus zur Lösung von (LP), falls P beschränkt ist: Bestimme alle Ecken x_1, \dots, x_t von P und finde einen Index i_0 mit $c^T x_{i_0} \geq c^T x_i \forall i = 1, \dots, t$.

Problem: t (die Anzahl der Ecken) kann sehr groß werden, zum Beispiel beim Würfel:

$$C_n = \{x \in \mathbb{R}^n : -1 \leq x_i \leq 1, i = 1, \dots, n\}.$$

$2n$ Ungleichungen, aber 2^n Ecken.

Bessere Algorithmen lernen wir in späteren Kapiteln kennen. In diesem Kapitel beschäftigen wir uns zunächst mit Optimalitätsbedingungen und Dualitätstheorie.

Definition 5.2. Seien $A \in \mathbb{R}^{m \times n}, c \in \mathbb{R}^n, b \in \mathbb{R}^m$ gegeben. Dies definiert das primale LP

$$\begin{aligned} (PLP) \quad & p^* = \max c^T x \\ & x \in \mathbb{R}^n \\ & Ax \leq b \end{aligned}$$

und das duale LP

$$\begin{aligned} (DLP) \quad & d^* = \min y^T b \\ & y \in \mathbb{R}^m \\ & y \geq 0 \\ & y^T A = c^T. \end{aligned}$$

Satz 5.3. (schwache Dualität)

Sei x zulässig für (PLP) und y zulässig für (DLP), dann ist

$$c^T x \leq y^T b.$$

Insbesondere ist $p^* \leq d^*$.

Beweis. Es gilt

$$y^T b - c^T x \geq y^T (Ax) - (y^T A)x = 0$$

□

Satz 5.4. (Optimalitätsbedingung)

Seien x, y zulässig. Dann ist x optimal für (PLP) und y optimal für (DLP) $\Leftrightarrow y^T (Ax - b) = 0$.

Insbesondere gelten die Komplementaritätsbedingungen:

$$\begin{aligned} y_j \neq 0 &\Rightarrow (Ax - b)_j = 0 \\ (Ax - b)_j \neq 0 &\Rightarrow y_j = 0, \quad \text{für alle } j = 1, \dots, m. \end{aligned}$$

Beweis. \rightarrow Blatt 9. □

Theorem 5.5. (starke Dualität, von Neumann (1947))

Falls (PLP) und (DLP) beide zulässige Lösungen besitzen, dann gilt $p^* = d^*$.

Beweis.

$p^* \leq d^*$: ✓ schwache Dualität.

$p^* \geq d^*$: 1. Behauptung: $\exists x_0 : Ax_0 \leq b$ und $c^T x_0 = p^*$.

Beweis: Da (DLP) zulässig ist, folgt aus der schwachen Dualität:

$$p^* = \sup\{c^T x : Ax \leq b\} < +\infty.$$

Zu zeigen: $\exists x_0 : Ax_0 \leq b$ und $c^T x_0 \geq p^*$. Angenommen ein solches x_0 gibt es nicht. Das heißt $\nexists x_0 : \begin{bmatrix} A \\ -c^T x \end{bmatrix} x_0 \leq \begin{pmatrix} b \\ -p^* \end{pmatrix}$. Nach Farkas Lemma

$$\exists \begin{pmatrix} y \\ \lambda \end{pmatrix} \geq 0 : (y^T \lambda) \begin{bmatrix} A \\ -c^T x \end{bmatrix} = 0 \text{ und } (y^T \lambda) \begin{pmatrix} b \\ -p^* \end{pmatrix} < 0.$$

Dann

$$\begin{aligned} \lambda p^* &= \sup\{\lambda c^T x : Ax \leq b\} \\ &= \sup\{y^T Ax : Ax \leq b\} \\ &\leq y^T b \\ &< \lambda p^*. \end{aligned}$$

Widerspruch.

2. Behauptung: $\exists y_0 \geq 0 : y_0^T A = c^T$ und $y_0^T b \leq p^*$.

Beweis: Angenommen ein solches y_0 gibt es nicht. Dann hat

$$(y_0^T \lambda) \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} = (c^T \ p^*)$$

keine Lösung mit $(y_0^T \lambda) \geq (0 \ 0)$. Nach Farkas Lemma

$$\exists \begin{pmatrix} z \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \text{ mit } \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \begin{pmatrix} z \\ \mu \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ und } (c^T \ p^*) \begin{pmatrix} z \\ \mu \end{pmatrix} < 0.$$

1. Fall: $\mu = 0$

Dann ist $Az \geq 0$ und $c^T z < 0$. Nach Voraussetzung hat $Ax \leq b$ eine Lösung x_0 . Dann ist für großes τ :

$$A(x_0 - \tau z) \leq b \text{ und } c^T(x_0 - \tau z) > p^*.$$

Widerspruch.

2. Fall: $\mu > 0$.

OBdA. sei $\mu = 1$ (durch Skalieren). Dann

$$Az + b \geq 0 \text{ und } c^T z + p^* < 0.$$

Also

$$A(-z) \leq b \text{ und } c^T(-z) > p^*.$$

Widerspruch.

□

Korollar 5.6. *Es gilt*

$$\max\{c^T x : x \geq 0, Ax \leq b\} = \min\{y^T b : y^T A \geq c^T\},$$

falls beide Mengen gültiger Lösungen nicht leer sind.

Beweis. Setze $\tilde{A} = \begin{bmatrix} A \\ -A \\ -I \end{bmatrix}$, $\tilde{b} = \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$. Dann

$$\begin{aligned} & \max\{c^T x : x \geq 0, Ax = b\} \\ &= \max\{c^T x : \tilde{A}x \leq \tilde{b}\} \\ &= \min\{z^T \tilde{b} : z \geq 0, z^T \tilde{A} = c^T\} \\ &= \min\{u^T b - v^T b : u, v, w \geq 0, u^T A - v^T A - w^T = c^T\} \quad \left(\text{mit } z = \begin{pmatrix} u \\ v \\ w \end{pmatrix}\right) \\ &= \min\{y^T b : y^T A \geq c^T\} \quad (\text{mit } y = u - v). \end{aligned}$$

□

PROF. DR. F. VALLENTIN, DR. A. GUNDELT, MATHEMATISCHES INSTITUT, UNIVERSITÄT ZU KÖLN,
WEYERTAL 86–90, 50931 KÖLN, DEUTSCHLAND
E-mail address: frank.vallentin@uni-koeln.de, anna.gundert@uni-koeln.de

KAPITEL 6 — GANZZAHLIGE OPTIMIERUNG UND VOLLSTÄNDIG UNIMODULARE MATRIZEN

F. VALLENTIN, A. GUNDERT

1. GANZZAHLIGE LINEARE PROGRAMME

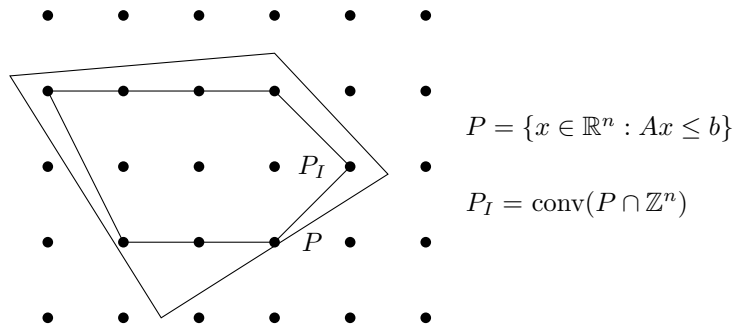
Viele Optimierungsprobleme des Operations Research lassen sich als *ganzzahlige lineare Programme* formulieren.

Definition 1.1. Ein ganzzahliges LP (in Standardform) ist von der Form:

$$\begin{aligned} \max \quad & c^T x \\ & x \in \mathbb{Z}^n \quad (\text{„}x \text{ ganzzahlig“}) \\ & Ax \leq b, \end{aligned}$$

wobei $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Ein ganzzahliges lineares Programm nennen wir auch *ILP* oder *IP* (= „integer linear program“).

Wir maximieren also eine lineare Funktion über den ganzzahligen Punkten in einem Polyeder.



Beispiel 1.2. Es sei $G = (V, E)$ ein ungerichteter Graph. Mit \mathbb{R}^V bezeichnen wir den Vektorraum der Funktionen von V nach \mathbb{R} , also

$$\begin{aligned} \mathbb{R}^V &= \{f: V \rightarrow \mathbb{R} : f \text{ Funktion}\} \quad \text{und analog} \\ \mathbb{R}^E &= \{f: E \rightarrow \mathbb{R} : f \text{ Funktion}\} \\ \mathbb{R}^{V \times E} &= \{f: V \times E \rightarrow \mathbb{R} : f \text{ Funktion}\} \end{aligned}$$

Wir definieren die Inzidenzmatrix $A \in \mathbb{R}^{V \times E}$ durch

$$A_{v,e} = \begin{cases} 1, & \text{falls } v \in e, \\ 0, & \text{sonst.} \end{cases}$$

Dann

$$\begin{aligned} \nu(G) &= \max\{|M| : M \subseteq E \text{ Matching in } G\} \\ &= \max\left\{\sum_{e \in E} x_e : x_e \geq 0 \forall e \in E, x_e \in \mathbb{Z} \forall e \in E, \sum_{e:v \in e} x_e \leq 1 \forall v \in V\right\} \\ &= \max\left\{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}^T x : x \geq 0, Ax \leq \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, x \in \mathbb{Z}^E\right\} \end{aligned}$$

Wir werden sehen, dass für G bipartit

$$\nu(G) = \max\left\{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}^T x : x \geq 0, Ax \leq \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, x \in \mathbb{R}^E\right\}$$

gilt. Im allgemeinen, nicht-bipartiten Fall kann aber „ $<$ “ gelten.

Zum Beispiel für $G = \text{Dreieck}$, $x_e = \frac{1}{2} \forall e \in E$.

Dualitätsbeziehung:

Es gilt

$$\begin{aligned} &\max\{c^T x : Ax \leq b, x \in \mathbb{Z}^n\} \\ &\leq \max\{c^T x : Ax \leq b, x \in \mathbb{R}^n\} \\ &= \min\{y^T b : y \geq 0, y^T A = c^T, y \in \mathbb{R}^m\} \\ &\leq \min\{y^T b : y \geq 0, y^T A = c^T, y \in \mathbb{Z}^m\}, \end{aligned}$$

falls (PLP) und (DLP) beide zulässige Lösungen besitzen. Im Allgemeinen gilt „ $<$ “, wie das obige Beispiel zeigt.

Kein Wunder: LP kann *effizient* gelöst werden (in polynomieller Zeit), ILP dagegen *wahrscheinlich nicht* (Äquivalent zum $P \neq NP$ -Problem, für dessen Lösung 1.000.000 Dollar ausgeschrieben sind).

Frage: In welcher Situation gilt „ $=$ “?

Definition 1.3. Ein Polytop $P \subseteq \mathbb{R}^n$ heißt ganzzahlig, falls jede Ecke z von P ganzzahlig ist, das heißt $z \in \mathbb{Z}^n$.

Falls P ganzzahlig ist, dann gilt also

$$\max\{c^T x : x \in P, x \in \mathbb{Z}^n\} = \max\{c^T x : x \in P, x \in \mathbb{R}^n\}$$

Wir suchen also Bedingungen an A und b , damit $P = \{x : Ax \leq b\}$ ganzzahlig ist.

2. VOLLSTÄNDIG-UNIMODULARE MATRIZEN

Definition 2.1. Eine Matrix $A \in \mathbb{R}^{m \times n}$ heißt vollständig-unimodular (VU), falls jeder ihrer Minoren (Determinanten quadratischer Teilmatrizen) gleich 0, -1 oder $+1$ ist. Insbesondere gilt für die Einträge: $A_{ij} \in \{0, -1, +1\}$.

Satz 2.2. Sei $A \in \mathbb{R}^{m \times n}$ VU, $b \in \mathbb{Z}^m$. Dann ist jede Ecke des Polyeders $P = \{x : Ax \leq b\}$ ganzzahlig.

Beweis. Sei z eine Ecke von P . Dann hat die Teilmatrix A_z vollen Rang n und enthält darum eine $n \times n$ Teilmatrix A' mit Rang n . OBdA sei A' in den ersten n Zeilen von A enthalten. Setze $b' = (b_1, \dots, b_n)^T \in \mathbb{Z}^n$, dann gilt $A'z = b'$. Nach der Cramerschen Regel ist $z_i = \frac{\det A'_i}{\det A'}$. Da $|\det A'| = 1$ und A'_i ganzzahlig ist, ist also $z_i \in \mathbb{Z}$. \square

Bemerkung 2.3. Nicht jedes Polyeder hat eine Ecke.

Definition 2.4. Ein Polyeder $P \subseteq \mathbb{R}^n$ heißt ganzzahlig, falls für alle $c \in \mathbb{R}^n$, für die $\max\{c^T x : x \in P\}$ endlich ist, das Maximum an einem ganzzahligen Vektor angenommen wird.

Satz 2.5. Sei $A \in \mathbb{R}^{m \times n}$ VU, $b \in \mathbb{Z}^m$. Dann ist $P = \{x : Ax \leq b\}$ ein ganzzahliges Polyeder.

Beweis. Sei $c \in \mathbb{R}^n$ gegeben und x^* eine optimale Lösung von $\max\{c^T x : Ax \leq b\}$. Wähle $d', d'' \in \mathbb{Z}^n$ mit $d' \leq x^* \leq d''$. Dann ist $Q = \{x \in \mathbb{R}^n : Ax \leq b, d' \leq x \leq d''\}$ ein Polytop. Weiter ist

$$Q = \left\{ x \in \mathbb{R}^n : \begin{bmatrix} A \\ -I \\ I \end{bmatrix} x \leq \begin{pmatrix} b \\ -d' \\ d'' \end{pmatrix} \right\}$$

und die Matrix $\begin{bmatrix} A \\ -I \\ I \end{bmatrix}$ ist VU. Nach dem vorherigen Satz ist Q also ein ganzzahliges

Polytop und außerdem wird $\max\{c^T x : x \in Q\}$ an einer Ecke \tilde{x} angenommen. Da $x^* \in Q$, ist $c^T \tilde{x} \geq c^T x^*$, und weil $A\tilde{x} \leq b$, ist \tilde{x} optimale Lösung von $\max\{c^T x : Ax \leq b\}$. \square

Korollar 2.6. Sei $A \in \mathbb{R}^{m \times n}$ VU, $b \in \mathbb{Z}^m$, $c \in \mathbb{Z}^n$. Dann haben die beiden linearen Programme

$$\max\{c^T x : Ax \leq b\} = \min\{y^T b : y \geq 0, y^T A = c^T\}$$

ganzzahlige Lösungen, falls die Optima endlich sind.

Beweis. Folgt aus dem vorhergehenden Satz, weil die Matrix

$$\begin{bmatrix} -I \\ A^T \\ -A^T \end{bmatrix} \in \mathbb{Z}^{(2n+m) \times m} \text{ VU ist.}$$

\square

Ziel: Charakterisierung von vollständig-unimodularen Matrizen mit Hilfe von Ganzzahligkeit von Polyedern.

Satz 2.7. (Hoffman, Kruskal 1956)

Sei $A \in \mathbb{Z}^{m \times n}$. Dann gilt

A ist VU \Leftrightarrow für alle $b \in \mathbb{Z}^m$ ist $P = \{x : Ax \leq b, x \geq 0\}$ ganzzahlig.

Beweis.

„ \Rightarrow “: Weil $P \subseteq \{x : x \geq 0\}$, wird $\max\{c^T x : x \in P\}$ in einer Ecke von P angenommen, falls das Maximum endlich ist (\rightarrow Aufgabe 9.1). Sei z eine Ecke von P . Definiere $B = \begin{bmatrix} A \\ -I \end{bmatrix}$. Betrachte

$$Bx = \begin{bmatrix} A \\ -I \end{bmatrix} x \leq \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

Dann enthält Bz eine reguläre Teilmatrix $B' \in \mathbb{Z}^{n \times n}$. Sei $b' \in \mathbb{Z}^n$ der entsprechende Teilvektor von $\begin{pmatrix} b \\ 0 \end{pmatrix}$. Dann ist $z = (B')^{-1}b'$ ganzzahlig aufgrund der Cramerschen Regel und weil A VU ist.

„ \Leftarrow “: Sei A' eine reguläre $k \times k$ -Teilmatrix von A . Zu zeigen: $|\det A'| = 1$.
 OBdA sei A' die obere linke Teilmatrix von A . Betrachte $B \in \mathbb{Z}^{m \times m}$ bestehend aus den ersten k und den letzten $m - k$ Spalten von $[A \ I_m]$, also

$$[A \ I_m] = \left[\begin{array}{c|c} A' & 0 \\ \hline 0 & I_{m-k} \end{array} \right] \text{ und } B = \left[\begin{array}{c|c} A' & 0 \\ \hline 0 & I_{m-k} \end{array} \right]$$

Dann ist $|\det B| = |\det A'|$.

Ziel: Zeige $B^{-1} \in \mathbb{Z}^{m \times m}$.

Denn: Weil dann $\det B^{-1} \in \mathbb{Z}$ und $\det B \cdot \det B^{-1} = 1$, muss $|\det B| = 1$ sein.

Sei $i \in \{1, \dots, m\}$. Wir zeigen, dass die i -te Spalte von B^{-1} ganzzahlig ist: Wähle $y \in \mathbb{Z}^m$ mit $y + B^{-1}e_i \geq 0$ und setze $z = y + B^{-1}e_i$. Dann ist $Bz = By + e_i \in \mathbb{Z}^m$. Setze $b = Bz$. Füge in z hinter den ersten k Komponenten $n - k + k$ Nullen ein und nenne diesen Vektor z' . Dann gilt nach Konstruktion

$$[A \ I]z' = Bz = b.$$

Da $z' \geq 0$, liegt der Vektor z'' , der aus den ersten n Komponenten von z' besteht, in $P = \{x : Ax \leq b, x \geq 0\}$. Das heißt es gilt

$$\begin{bmatrix} A \\ -I \end{bmatrix} z'' \leq \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

Die ersten k und die letzten $n - k$ dieser Ungleichungen sind mit Gleichheit erfüllt. Da die entsprechenden Zeilen von $\begin{bmatrix} A \\ -I \end{bmatrix}$ linear unabhängig sind, ist also z'' eine Ecke von P . Nach Voraussetzung ist $z'' \in \mathbb{Z}^n$. Auf den ersten n Komponenten stimmen z' und z'' überein. Die letzten m Komponenten von z' sind durch $b - Az''$ gegeben und also auch ganzzahlig. Also ist $z \in \mathbb{Z}^m$ und somit auch $B^{-1}e_i = z - y \in \mathbb{Z}^m$. □

3. VOLLSTÄNDIG-UNIMODULARE MATRIZEN UND BIPARTITE GRAPHEN

Satz 3.1. Sei $G = (V, E)$ ein ungerichteter Graph. Dann gilt
 G ist bipartit \Leftrightarrow Die Inzidenzmatrix $A \in \mathbb{R}^{V \times E}$ ist VU.

Beweis.

„ \Rightarrow “: Sei B eine $t \times t$ Teilmatrix von A .

Zeige: $\det B \in \{-1, 0, +1\}$ per Induktion nach t .

$t = 1$: \checkmark

$t > 1$:

1. Fall: B enthält eine Nullspalte. Dann ist $\det B = 0$.
2. Fall: B enthält eine Spalte, die genau eine 1 enthält. Dann ist B (nach eventueller Permutation von Zeilen und Spalten) von der Form

$$B = \begin{bmatrix} 1 & b^T \\ 0 & B' \end{bmatrix}, \quad b \in \mathbb{R}^{t-1}, \quad B' \in \mathbb{R}^{(t-1) \times (t-1)}.$$

Nach Induktionsvoraussetzung ist $\det B' \in \{-1, 0, +1\}$. Also auch $\det B \in \{-1, 0, +1\}$.

3. Fall: Jede Spalte von B enthält genau zwei Einsen. Da G bipartit ist, kann man (nach eventueller Permutation der Zeilen) B als

$$B = \begin{bmatrix} B' \\ B'' \end{bmatrix}$$

schreiben, wobei jede Spalte von B' genau eine 1 enthält und jede Spalte von B'' genau eine 1 enthält. Aufaddieren aller Zeilen von B' ergibt den Vektor $(1, \dots, 1)$. Genauso wie Aufaddieren aller Zeilen von B'' . Das heißt die Zeilen von B sind linear abhängig, also $\det B = 0$.

„ \Leftarrow “: Sei A VU. Angenommen G ist nicht bipartit. Dann enthält G einen Kreis ungerader Länge mit Knoten v_1, \dots, v_k und Kanten e_1, \dots, e_k . Die entsprechende Teilmatrix von A (mit Zeilen v_1, \dots, v_k und Spalten e_1, \dots, e_k) ist von der Form:

	e_1	e_2	e_3	\dots	e_{k-1}	e_k
v_1	1	0	0	\dots	0	1
v_2	1	1	0	\dots	0	0
v_3	0	1	1	\dots	0	0
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
v_{k-1}	0	0	0	\dots	1	0
v_k	0	0	0	\dots	1	1

Zum Beispiel durch Entwicklung nach der k -ten Spalte sieht man, dass die Determinante dieser Matrix gleich $(+1) \cdot 1 + (+1) \cdot 1 = 2$ ist (k ist ungerade!). Also kann A nicht VU sein. □

Korollar 3.2. (*Matching-Theorem von König; siehe Kapitel 3.3*)

Sei $G = (V, E)$ ein bipartiter Graph. Dann

$$\begin{aligned} \nu(G) &= \max\{|M| : M \subseteq E \text{ Matching in } G\} \\ &= \min\{|U| : U \subseteq V \text{ Knotenüberdeckung in } G\} \\ &= \tau(G). \end{aligned}$$

Beweis.

$$\begin{aligned}
\nu(G) &= \max\left\{\sum_{e \in E} x_e : x_e \geq 0 \forall e \in E, x_e \in \mathbb{Z}, \sum_{e: v \in e} x_e \leq 1 \forall v \in V\right\} \\
&= \max\{1^T x : x \in \mathbb{Z}^E, x \geq 0, Ax \leq 1\} \\
&= \max\{1^T x : x \in \mathbb{R}^E, x \geq 0, Ax \leq 1\} \quad (\text{nach Hoffman-Kruskal, weil } A \text{ VU ist}) \\
&= \min\{1^T y : y \in \mathbb{R}^V, y \geq 0, y^T A \geq 1\} \quad (\text{LP Dualitat}) \\
&= \min\{1^T y : y \in \mathbb{Z}^V, y \geq 0, y^T A \geq 1\} \quad (\text{Hoffman-Kruskal}) \\
&= \min\left\{\sum_{v \in V} y_v : y_v \in \mathbb{Z}, y_v \geq 0 \forall v \in V, y_u + y_v \geq 1 \forall \{u, v\} \in E\right\} \\
&= \tau(G).
\end{aligned}$$

Ein Vektor $y \in \mathbb{Z}^V$, der das Minimum erzielt, hat nur 0/1- Komponenten. Hierbei ist $y_v = 1 \Leftrightarrow v$ ist Knoten in einer minimalen Knotenuberdeckung. \square

Korollar 3.3. Sei $G = (V, E)$ ein bipartiter Graph. Dann

$$\text{conv}\{\chi^M : M \subseteq E \text{ Matching in } G\} = \{x \in \mathbb{R}^E : x \geq 0, Ax \leq 1\},$$

wobei A die Inzidenzmatrix von G ist.

Beweis.

„ \subseteq “: \checkmark

„ \supseteq “: $Q = \{x \in \mathbb{R}^E : x \geq 0, Ax \leq 1\}$ ist beschrankt, also ein Polytop und somit ist Q die konvexe Hulle seiner Ecken. Da A VU ist, sind alle Ecken von Q ganzzahlig. Und jeder ganzzahlige Vektor in Q ist Inzidenzvektor eines Matchings in G . \square

Korollar 3.4. (Theorem von Egervary (1931))

Sei $G = (V, E)$ ein bipartiter Graph und $w : E \rightarrow \mathbb{Z}$ eine Gewichtsfunktion. Dann ist

$$\begin{aligned}
\nu_W(G) &= \max\{w(M) : M \subseteq E \text{ Matching in } G\} \\
&= \min\left\{\sum_{v \in V} y_v : y \in \mathbb{Z}^V, y \geq 0, y_u + y_v \geq w(\{u, v\}) \forall \{u, v\} \in E\right\}.
\end{aligned}$$

Beweis. Weil die Inzidenzmatrix $A \in \mathbb{R}^{V \times E}$ von G VU ist, gilt

$$\max\{w^T x : x \geq 0, Ax \leq 1\} = \min\{1^T y : y \geq 0, y^T A \geq w\},$$

wobei beide Probleme optimale ganzzahlige Losungen haben. \square

4. VOLLSTANDIG-UNIMODULARE MATRIZEN UND GERICHTETE GRAPHEN

Definition 4.1. Die Inzidenzmatrix eines gerichteten Graphen $D = (V, A)$ ist $M \in \mathbb{R}^{V \times A}$ mit

$$M_{v,a} = \begin{cases} +1, & \text{falls } a \in \delta^-(v) \quad \xrightarrow{a} \bullet v \\ -1, & \text{falls } a \in \delta^+(v) \quad v \bullet \xrightarrow{a} \\ 0, & \text{sonst.} \end{cases}$$

Jede Spalte von M enthält genau eine $+1$ und genau eine -1 .

Satz 4.2. M ist VU.

Beweis. Sei $B \in \mathbb{R}^{t \times t}$ eine Teilmatrix von M . Zeige per Induktion nach t , dass $\det B \in \{-1, 0, +1\}$.

$t = 1$: ✓

$t > 1$:

1.Fall: B hat eine Nullspalte $\Rightarrow \det B = 0$.

2.Fall: B hat eine Spalte, die genau einen Eintrag $\neq 0$ enthält.
Dann, nach eventueller Vertauschung von Zeilen und Spalten, ist

$$B = \begin{bmatrix} \pm 1 & b^T \\ 0 & B' \end{bmatrix}, \quad b \in \mathbb{R}^{t-1}, \quad B' \in \mathbb{R}^{(t-1) \times (t-1)}.$$

Nach Induktionsvoraussetzung ist $\det B' \in \{-1, 0, +1\}$,
also $\det B = (\pm 1) \cdot \det B' \in \{-1, 0, +1\}$.

3.Fall: Jede Spalte von B enthält zwei Einträge $\neq 0$. Aufaddieren sämtlicher Zeilen von B ergibt 0, das heißt $\det B = 0$.

□

Korollar 4.3. (*Max-flow-min-cut Theorem*)

Sei $D = (V, A)$ ein gerichteter Graph, $s, t \in V$ und $c: A \rightarrow \mathbb{R}_{\geq 0}$ eine Kapazitätsfunktion. Dann

$$\begin{array}{ll} \max \text{value}(f) & = \min c(\delta^+(U)) \\ f \text{ ist } s\text{-}t\text{-Fluss} & U \subseteq V \\ 0 \leq f \leq c & s \in U, t \notin U. \end{array}$$

Beweis.

$\max \leq \min$: ✓

$\max \geq \min$: Sei $M \in \mathbb{R}^{V \times A}$ die Inzidenzmatrix von D und sei M' die Matrix, die man durch Streichen der zu s und t gehörenden Zeilen aus M bekommt. Dann gilt

$$f \in \mathbb{R}^A \text{ ist } s\text{-}t\text{-Fluss} \Leftrightarrow M'f = 0.$$

Sei $w^T \in \mathbb{R}^A$ die zu t gehörende Zeile von M . Dann ist

$$\text{value}(f) = w^T f.$$

Also

$$\max \text{value}(f) = \max\{w^T f : 0 \leq f \leq c, M'f = 0\}$$

f ist s - t -Fluss

$$0 \leq f \leq c$$

$$\stackrel{(1)}{=} \min\{y^T c : y \geq 0, z \in \mathbb{R}^{V \setminus \{s, t\}}, y^T + z^T M' \geq w^T\}$$

$$= \min\{y^T c : (y^T \ z^T) \begin{bmatrix} I & I \\ 0 & M' \end{bmatrix} \geq (0 \ w^T)\},$$

wobei (1) nach LP Dualität gilt. Die Matrix $\begin{bmatrix} I & I \\ 0 & M' \end{bmatrix}$ ist VU und der Vektor $(0 \ w^T)$ ist ganzzahlig, also wird das Maximum an ganzzahligen Vektoren y^* , z^* angenommen. Erweitern wir z^* zu $z^{**} \in \mathbb{Z}^V$ durch $z_t^{**} = -1$ und $z_s^{**} = 0$ und $z_v^{**} = z_v^* \ \forall v \in V \setminus \{s, t\}$, so gilt $(y^*)^T + (z^{**})^T M \geq 0$, denn

$$\begin{aligned} [(y^*)^T + (z^{**})^T M]_a &= \underbrace{y_a^* + [(z^*)^T M']_a}_{\geq w_a} + \underbrace{z_s^{**}}_{=0} \cdot M_{s,a} + \underbrace{z_t^{**}}_{=-1} \cdot M_{t,a} \\ &\geq w_a + (-1) \cdot \underbrace{M_{t,a}}_{=w_a} = 0. \end{aligned}$$

Setze $U = \{v \in V : z_v^{**} \geq 0\}$. Dann ist $s \in U$, $t \notin U$.

Behauptung: $\sum_{a \in \delta^+(U)} c(a) \leq (y^*)^T c$ (= max value(f)).

Beweis: Es genügt zu zeigen, dass falls $a \in \delta^+(U)$, dann $y_a^* \geq 1$. Da $z^{**} \in \mathbb{Z}^V$, gilt für $a = (u, v) \in \delta^+(U)$:

$$z_u^{**} \geq 0, \quad z_v^{**} \leq -1 \text{ nach Definition von } U.$$

Wegen $[(y^*)^T + (z^{**})^T M]_a \geq 0$ gilt

$$\begin{aligned} y_a^* + z_v^{**} - z_u^{**} &\geq 0, \text{ das heißt} \\ y_a^* &\geq z_u^{**} - z_v^{**} \geq 1. \end{aligned}$$

□

KAPITEL 7 — DAS ELIMINATIONSVERFAHREN VON FOURIER UND MOTZKIN

F. VALLENTIN, A. GUNDERT

Ein algorithmisches Grundproblem in der Polyedertheorie ist zu entscheiden, ob ein Polyeder $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ nicht leer ist.

In der linearen Algebra hat man für ein ähnliches Problem, zu entscheiden, ob die Lösungsmenge $L = \{x \in \mathbb{R}^n : Ax = b\}$ eines linearen Gleichungssystems nicht leer ist, das Eliminationsverfahren von Gauß.

Wir lernen hier ein ähnliches Verfahren für unser Problem kennen.

1. FOURIER-MOTZKIN-ELIMINATION

Gegeben seien $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Ziel ist es, $x \in \mathbb{R}^n$ mit $Ax \leq b$ zu finden bzw. zu entscheiden (mit mathematischer Sicherheit), dass es ein solches x nicht gibt.

Idee: Eliminiere die Variable x_1 ; Finde $\tilde{A} \in \mathbb{R}^{\tilde{m} \times (n-1)}$, $\tilde{b} \in \mathbb{R}^{\tilde{m}}$, so dass

$$\exists x \in \mathbb{R}^n : Ax \leq b \Leftrightarrow \exists \tilde{x} \in \mathbb{R}^{n-1} : \tilde{A}\tilde{x} \leq \tilde{b}.$$

Dazu: Multipliziere Zeilen von A und entsprechende Einträge von b mit positiven Konstanten. Dann hat das System $Ax \leq b$ nach Ummummerierung der Zeilen folgende Gestalt:

$$(*) \quad \begin{bmatrix} 1 & (a'_1)^T \\ \vdots & \vdots \\ 1 & (a'_r)^T \\ -1 & (a'_{r+1})^T \\ \vdots & \vdots \\ -1 & (a'_{r+s})^T \\ 0 & (a'_{r+s+1})^T \\ \vdots & \vdots \\ 0 & (a'_m)^T \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \leq \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

wobei $(a'_i)^T \in \mathbb{R}^{1 \times (n-1)}$ die i -te Zeile von A ist, in der das erste Element gelöscht wurde (es kann passieren, dass $r = 0$ oder $s = 0$ ist). Betrachte die ersten r Bedingungen:

$$x_1 + (a'_i)^T \underbrace{\begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix}}_{=\tilde{x}} \leq b_i \Rightarrow x_1 \leq b_i - (a'_i)^T \tilde{x}, \quad i = 1, \dots, r.$$

Genauso die nächsten s Bedingungen:

$$x_1 + (a'_{r+j})^T \tilde{x} \leq b_{r+j} \Rightarrow x_1 \geq (a'_{r+j})^T \tilde{x} - b_{r+j}, \quad j = 1, \dots, s.$$

Zusammen gilt also

$$(**) \quad \max_{j=1, \dots, s} (a'_{r+j})^T \tilde{x} - b_{r+j} \leq x_1 \leq \min_{i=1, \dots, r} b_i - (a'_i)^T \tilde{x}.$$

(Falls $s = 0$, dann ist $\max_{j=1, \dots, s} (a'_{r+j})^T \tilde{x} - b_{r+j} = -\infty$ und falls $r = 0$, ist $\min_{i=1, \dots, r} b_i - (a'_i)^T \tilde{x} = +\infty$. In diesen Fällen ist also P in Richtung x_1 unbeschränkt.)

Also kann man x_1 eliminieren und das System (*) ist lösbar genau dann, wenn das System

$$\begin{aligned} (a'_{r+j})^T \tilde{x} - b_{r+j} &\leq b_i - (a'_i)^T \tilde{x}, \quad i = 1, \dots, r, \quad j = 1, \dots, s \\ (a'_i)^T \tilde{x} &\leq b_i, \quad i = r + s + 1, \dots, m \end{aligned}$$

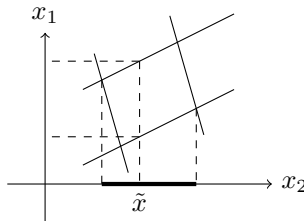
bzw. das System

$$(***) \quad \begin{aligned} ((a'_{r+j})^T + (a'_i)^T) \tilde{x} &\leq b_i + b_{r+j}, \quad i = 1, \dots, r, \quad j = 1, \dots, s \\ (a'_i)^T \tilde{x} &\leq b_i, \quad i = r + s + 1, \dots, m. \end{aligned}$$

lösbar ist. Das neue System hat $r \cdot s + m - (r + s)$ viele Ungleichungen und $n - 1$ Variablen.

Bemerkung 1.1.

a) (***) entspricht der Projektion des Polyeders $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ entlang der x_1 -Achse.



- b) Eine Lösung \tilde{x} kann zu einer Lösung (x_1, \tilde{x}) von (*) erweitert werden. Dazu muss x_1 die Ungleichungen (**) erfüllen.
- c) Das Verfahren wird fortgesetzt, indem nun sukzessive die Variablen x_2, x_3, \dots eliminiert werden, bis man bei x_n angekommen ist.
- d) Für x_n ist es offensichtlich, ob das finale System eine Lösung besitzt. Das finale System hat eine Lösung \Leftrightarrow das Ursprungssystem (*) hat eine Lösung.

2. LÖSEN VON LPS MIT FOURIER-MOTZKIN

Wir wollen das LP

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

mit dem Eliminationsverfahren von Fourier und Motzkin lösen. Dazu führen wir eine zusätzliche Variable λ ein und betrachten das System

$$Ax \leq b, \quad \lambda \leq c^T x.$$

Die Idee ist, dass λ dem größtmöglichen Wert der Zielfunktion $c^T x$, also dem Maximum, so dass alle Ungleichungen erfüllt sind, entsprechen soll. Wegen $\lambda \leq c^T x \Leftrightarrow \lambda - c^T x \leq 0$, ist das System äquivalent zu

$$\begin{bmatrix} A & 0 \\ -c^T & 1 \end{bmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} \leq \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

Man kann nun das LP lösen, indem man eine Lösung $\begin{pmatrix} x \\ \lambda \end{pmatrix}$ von diesem System findet, so dass λ so groß wie möglich ist. Dazu eliminiert man x_1, \dots, x_n bis λ die letzte Variable ist. Dann wählt man λ so groß wie möglich.

3. FARKAS LEMMA MIT FOURIER-MOTZKIN

Lemma 3.1. (Farkas)

$Ax \leq b$ hat keine Lösung $\Leftrightarrow \exists y \geq 0 : y^T A = 0, y^T b < 0$.

Beweis.

„ \Leftarrow “: (hatten wir schon) Angenommen $Ax \leq b$ hat eine Lösung. Dann

$$0 = y^T Ax \leq y^T b < 0, \text{ Widerspruch.}$$

„ \Rightarrow “: (algorithmisch mit Fourier-Motzkin)

Behauptung: $\underbrace{(0, \dots, 0)}_{y^T A}, \underbrace{-1)}_{y^T b} \in \mathbb{R}^{n+1}$ ist eine nicht-negative Linearkombination der Zeilen der Matrix $[A|b]$, gegeben durch y .

Beweis: (per Induktion nach n)

$n = 1$: Angenommen $Ax_1 \leq b$ hat keine Lösung, das heißt

$$\begin{bmatrix} 1 \\ \vdots \\ 1 \\ -1 \\ \vdots \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} x_1 \leq \begin{pmatrix} b_1 \\ \vdots \\ b_r \\ b_{r+1} \\ \vdots \\ b_{r+s} \\ b_{r+s+1} \\ \vdots \\ b_m \end{pmatrix}$$

hat keine Lösung. Es gibt zwei Fälle, wie dies zustande kommen kann:

1.Fall: $0 \cdot x_1 \leq b_{r+s+k}$ und $b_{r+s+k} < 0$.

Wähle $y = (0, \dots, 0, \underbrace{-\frac{1}{b_{r+s+k}}}_{\text{Position } r+s+k}, 0, \dots, 0)^T$.

2.Fall: $b_i < -b_{r+j}$ für ein $i \in \{1, \dots, r\}$ und ein $j \in \{1, \dots, s\}$.

Dann hat das System $Ax_1 \leq b$ keine Lösung, denn $-b_{r+j} \leq x_1 \leq b_i$, aber

$b_i < -b_{r+j}$. Wähle

$$y = (0, \dots, 0, \underbrace{-\frac{1}{b_i + b_{r+j}}}_{\text{Position } i}, 0, \dots, 0, \underbrace{-\frac{1}{b_i + b_{r+j}}}_{\text{Position } r+j}, 0, \dots, 0)^T.$$

$n > 1$: Betrachte das System $A'x' \leq b'$, in dem die Variable x_1 eliminiert wurde. Das System besitzt keine Lösung, also ist der Vektor $(0, \dots, 0, -1) \in \mathbb{R}^n$ eine nicht-negative Linearkombination der Zeilen von $[A'|b']$. Nach Konstruktion sind die Zeilen der Matrix $[0 \ A|b]$ nicht-negative Linearkombinationen der Zeilen von $[A|b]$. Also ist der Vektor $(0, \dots, 0, -1) \in \mathbb{R}^{n+1}$ eine nicht-negative Linearkombination der Zeilen von $[A|b]$.

□

PROF. DR. F. VALLENTIN, DR. A. GUNDELT, MATHEMATISCHES INSTITUT, UNIVERSITÄT ZU KÖLN,
WEYERTAL 86–90, 50931 KÖLN, DEUTSCHLAND

E-mail address: frank.vallentin@uni-koeln.de, anna.gundert@uni-koeln.de

KAPITEL 8 — DAS SIMPLEXVERFAHREN

F. VALLENTIN, A. GUNDERT

In diesem Kapitel lernen wir eine weitere Methode kennen, lineare Programme zu lösen: den Simplexalgorithmus.

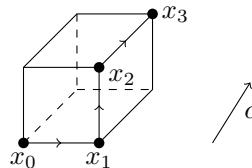
Ziel: Löse (LP)

$$(LP) \quad \begin{aligned} p^* &= \max c^T x \\ x &\in \mathbb{R}^n \\ Ax &\leq b. \end{aligned}$$

Wissen: Angenommen $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ ist ein Polytop. Dann wird das Maximum an einer Ecke von P angenommen (\rightarrow Kapitel 5.5).

Geometrische Idee: Finde eine Sequenz von Ecken $x_0, x_1, \dots, x_N \in P$, so dass

$$c^T x_0 \leq c^T x_1 \leq \dots \leq c^T x_N = p^*.$$



1. SIMPLEXALGORITHMUS MIT BEKANNTER STARTECKE

Zunächst treffen wir die spezielle Annahme, dass das Polyeder P eine Ecke x_0 besitzt, die wir kennen. Später beseitigen wir diese spezielle Annahme.

Simplexalgorithmus

- Wähle ein Teilsystem $A_0 x \leq b_0$ von $Ax \leq b$ mit einer regulären quadratischen Matrix A_0 , wobei $A_0 x_0 = b_0$.
- Bestimme $u \in \mathbb{R}^m$ mit $c^T = u^T A$ und $u_i = 0$, falls Zeile i von A nicht zu A_0 gehört. Dazu berechne $c^T A_0^{-1}$ und füge Nullen an den entsprechenden Stellen hinzu.

1. Fall: $u \geq 0$.

Dann ist x_0 optimal, weil u eine optimale duale Lösung ist. Denn:

$$\begin{aligned} c^T x_0 &= u^T A x_0 = u^T b \geq \min\{y^T b : y \geq 0, y^T A = c^T\} \\ &= \max\{c^T x : Ax \leq b\}. \end{aligned}$$

2.Fall: $u \not\geq 0$.

Sei i der kleinste Index mit $u_i < 0$.

Wähle $y \in \mathbb{R}^n$ mit $a^T y = 0$ für alle Zeilen a^T von A_0 mit $a^T \neq a_i^T$
und $a_i^T y = -1$.

(y ist die entsprechende Spalte von $-A_0^{-1}$).

2.a) $a^T y \leq 0$ für alle Zeilen a^T von A .

Dann ist $x_0 + \lambda y \in P \forall \lambda \geq 0$. Desweiteren ist

$$\begin{aligned} c^T(x_0 + \lambda y) &= c^T x_0 + \lambda c^T y \\ &= c^T x_0 + \lambda \underbrace{u^T A y}_{=-u_i} \\ &= c^T x_0 - \lambda u_i \rightarrow +\infty \text{ für } \lambda \rightarrow \infty. \end{aligned}$$

Das heißt das LP ist unbeschränkt.

2.b) $a^T y > 0$ für eine Zeile a^T von A .

Setze

$$\begin{aligned} \lambda_0 &= \max\{\lambda : x_0 + \lambda y \in P\} \\ &= \min\left\{\frac{b_j - a_j^T x_0}{a_j^T y} : j = 1, \dots, m, a_j^T y > 0\right\}, \end{aligned}$$

wobei die Gleichheit von Maximum und Minimum im Beweis des Theorems von Minkowski gezeigt wurde. Sei j der kleinste Index, an dem das Minimum angenommen wird. Definiere

A_1 = Matrix, die man aus A_0 erhält,

indem man Zeile a_i^T durch Zeile a_j^T austauscht.

$$x_1 = x_0 + \lambda_0 y.$$

Dann gilt $A_1 x_1 = b_1$.

- Gehe zum Anfang mit A_1, x_1 , anstelle von A_0, x_0 .
- Wiederhole diese Schritte, bis $u \geq 0$ oder bis klar ist, dass das LP unbeschränkt ist.

Satz 1.1. *Der Simplexalgorithmus terminiert nach endlich vielen Schritten.*

Beweis. Wir bezeichnen die Variablen im k -ten Schritt mit $A_k, x_k, u_k, y_k, \lambda_{0,k}$. Es gilt

$$\begin{aligned} c^T x_0 &\leq c^T x_1 \leq \dots, \text{ und} \\ c^T x_k &= c^T x_{k+1} \Leftrightarrow x_k = x_{k+1}, \end{aligned}$$

weil

$$c^T x_{k+1} = c^T(x_k + \lambda_{0,k} y_k) \quad \text{mit } \lambda_{0,k} \geq 0$$

und

$$c^T y_k = (-u_k)_i > 0.$$

Angenommen der Algorithmus landet in einer Endlosschleife. Dann gibt es k, l mit $k < l$ und $A_k = A_l$, weil es nur endlich viele verschiedene Teilmatrizen von A gibt.

Dann

$$c^T x_k = c^T x_l, \text{ also } x_k = x_{k+1} = \dots = x_l.$$

Sei r der größte Index, so dass die Zeile a_r^T in einer Iteration aus A_t genommen wird, wobei $t = k, k + 1, \dots, l$. Dies passiere in Schritt p . Weil $A_k = A_l$, gibt es ein q , so dass a_r^T wieder in A_q aufgenommen wird. Dann

$$k \leq p < q < l.$$

Dann gilt für $j > r$

$$a_j^T \text{ kommt in } A_p \text{ vor} \iff a_j^T \text{ kommt in } A_q \text{ vor.}$$

Es gilt

$$u_p^T A y_q = c^T y_q > 0.$$

Also gibt es ein j mit $(u_p)_j (a_j^T y_q) > 0$. Aber:

1. Fall: a_j^T gehört nicht zu A_p . Dann $(u_p)_j = 0$, Widerspruch.
2. Fall: a_j^T gehört zu A_p .
 - a) $j > r$: Dann $a_j^T y_q = 0$, Widerspruch.
 - b) $j = r$: Dann $(u_p)_j < 0$ und $a_j^T y_q > 0$, Widerspruch.
 - c) $j < r$: Dann $(u_p)_j \geq 0$ und $a_j^T y_q \leq 0$, Widerspruch.

□

2. SIMPLEXALGORITHMUS OHNE STARTECKE

Jetzt beschäftigen wir uns mit der Frage, wie man den Simplexalgorithmus startet, wenn man keine Ecke x_0 kennt.

OBdA: Das LP ist von der Form

$$\max\{c^T x : x \geq 0, Ax \leq b\}.$$

(\rightarrow Aufgabe 9.1: Dann besitzt $P = \{x \in \mathbb{R}^n : x \geq 0, Ax \leq b\}$ eine Ecke.)

Idee: Um eine Ecke von P zu finden, füge eine Extravariablen hinzu und stelle ein neues LP auf, das eine offensichtliche Ecke besitzt und dessen optimale Lösung eine Ecke von P liefert.

Extravariablen: $y \in \mathbb{R}^m, y \geq 0$.

Neues LP:

$$\min e^T y$$

$$\begin{bmatrix} A & -I_m \\ -I_n & 0 \\ 0 & -I_m \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix},$$

mit $e = (1, \dots, 1)^T$.

Offensichtliche Ecke:

$$x = 0, \quad y_j = \begin{cases} 0, & \text{falls } b_j \geq 0 \\ -b_j, & \text{falls } b_j < 0 \end{cases}, \quad j = 1, \dots, m.$$

Dann ist $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+m}$ eine Ecke von

$$P' = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : Ax - y \leq b, x \geq 0, y \geq 0 \right\},$$

weil $\begin{pmatrix} x \\ y \end{pmatrix} \in P'$ und $\text{rang} \begin{bmatrix} A & -I_m \\ -I_n & 0 \\ 0 & -I_m \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = n + m$.

Jetzt kann man den Simplexalgorithmus mit der Startecke $\begin{pmatrix} x \\ y \end{pmatrix}$ verwenden, um das neue LP zu lösen. Sei $\begin{pmatrix} x^* \\ y^* \end{pmatrix}$ die Ecke von P' , die der Algorithmus liefert.

1.Fall: $e^T y^* > 0$.

Dann ist das Original-LP ungültig, denn

$$\nexists x \geq 0 : Ax \leq b, \text{ da } \exists j : y_j^* > 0 \text{ und } (Ax - y^*)_j \leq b_j.$$

2.Fall: $e^T y^* = 0$.

Dann $y^* = 0$ und x^* ist eine Ecke von P , weil $x^* \in P$ und $\text{rang} \begin{bmatrix} A \\ -I \end{bmatrix}_{x^*} = n$.

3. ZUR PRAKTISCHEN UND THEORETISCHEN EFFIZIENZ DES SIMPLEXALGORITHMUS

- + sehr schnell bei vielen praxisrelevanten Eingaben.
- + sehr gute Implementationen erhältlich (CPLEX, gurobi).
- Klee-Minty-Würfel (1972): Beispiel, dass der Algorithmus exponentiell viele Schritte im worst case benötigt.
- + Spielman-Teng (2004): „smoothed analysis“: Algorithmus ist polynomiell, falls die Eingabe leicht, zufällig „gestört“ ist.
- offenes Problem („polynomielle Hirsch-Vermutung“): Ist der maximale Abstand zwischen zwei Ecken polynomiell in m, n ?

PROF. DR. F. VALLENTIN, DR. A. GUNDELT, MATHEMATISCHES INSTITUT, UNIVERSITÄT ZU KÖLN,
WEYERTAL 86–90, 50931 KÖLN, DEUTSCHLAND

E-mail address: frank.vallentin@uni-koeln.de, anna.gundert@uni-koeln.de

KAPITEL 9 — DIE ELLIPSOIDMETHODE

F. VALLENTIN, A. GUNDERT

1. GRUNDLAGEN ZU ELLIPSOIDEN

Definition 1.1. Eine positiv definite Matrix $A \in \mathbb{R}^{n \times n}$ und ein Vektor $x \in \mathbb{R}^n$ definieren das Ellipsoid

$$\mathcal{E}(A, x) = \{y \in \mathbb{R}^n : (y - x)^T A^{-1} (y - x) \leq 1\}.$$

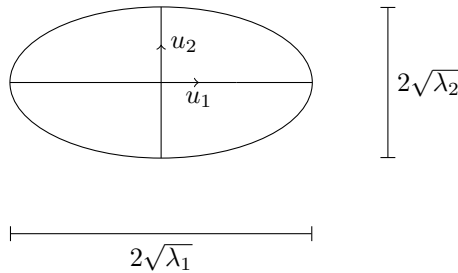
Beispiel 1.2. $\mathcal{E}(r^2 I_n, 0) = r B_n$, wobei $B_n = \{y \in \mathbb{R}^n : \|y\| \leq 1\}$ die n -dimensionale Einheitskugel ist.

Eigenschaften:

- Hauptachsentransformation/Spektralzerlegung:
Eine positiv definite Matrix $A \in \mathbb{R}^{n \times n}$ lässt sich folgendermaßen darstellen:

$$A = \sum_{i=1}^n \lambda_i u_i u_i^T,$$

wobei $\lambda_1, \dots, \lambda_n > 0$ die Eigenwerte von A sind, und $u_1, \dots, u_n \in \mathbb{R}^n$ eine Orthonormalbasis bilden, bestehend aus den zugehörigen Eigenvektoren. Diese Darstellung von A heißt *Spektralzerlegung*. Die Matrix A definiert ein Ellipsoid wie folgt: Die u_i bestimmen die Richtungen der Hauptachsen und die Länge der jeweiligen Hauptachse ist durch $2\sqrt{\lambda_i}$ gegeben.



- Volumen von $\mathcal{E}(A, x)$:

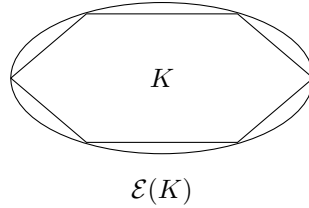
$$\text{vol } \mathcal{E}(A, x) = \sqrt{\det A} \cdot \text{vol } B_n,$$

wobei $\text{vol } B_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}$, mit der Gammafunktion Γ , die für unsere Zwecke ausreichend definiert ist durch

$$\Gamma(1/2) = \sqrt{\pi}, \quad \Gamma(1) = 1, \quad \Gamma(x+1) = x \cdot \Gamma(x).$$

Der Beweis der Volumenformel ist die Lösung von Aufgabe 12.3.

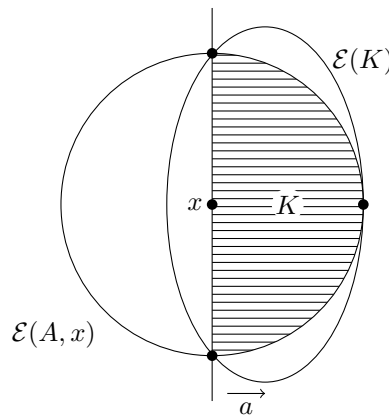
Satz 1.3. Sei $K \subseteq \mathbb{R}^n$ eine konvexe und kompakte Menge. Dann existiert ein eindeutiges Ellipsoid $\mathcal{E}(K)$ mit $K \subseteq \mathcal{E}(K)$ und mit minimalem Volumen (die beste äußere Approximation von K). Dieses Ellipsoid $\mathcal{E}(K)$ heißt das Loewner-John-Ellipsoid von K .



Beweis. → Vorlesung „Konvexe Optimierung“ im Wintersemester 2015/16. □

Hier: Spezialfall

$$K = \mathcal{E}(A, x) \cap \{y \in \mathbb{R}^n : a^T y \geq a^T x\} \text{ mit } a \neq 0.$$



Lemma 1.4. Sei $A \in \mathbb{R}^{n \times n}$ positiv definit, $x, a \in \mathbb{R}^n$, $a \neq 0$. Dann

$$\mathcal{E}(\mathcal{E}(A, x) \cap \{y \in \mathbb{R}^n : a^T y \geq a^T x\}) = \mathcal{E}(A', x')$$

mit

$$A' = \frac{n^2}{n^2 - 1} \left(A - \frac{2}{n+1} b b^T \right), \quad x' = x + \frac{1}{n} b, \quad b = \frac{1}{\sqrt{a^T A a}} A a.$$

Desweiteren gilt

$$\frac{\text{vol } \mathcal{E}(A', x')}{\text{vol } \mathcal{E}(A, x)} < e^{-\frac{1}{2(n+1)}} < 1.$$

Beweis. Anleitung:

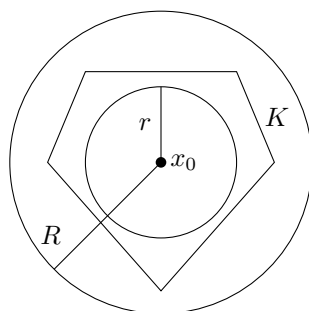
- $\mathcal{E}(A, x) \cap \{y \in \mathbb{R}^n : a^T y \geq a^T x\} \subseteq \mathcal{E}(A', x')$ folgt aus Aufgabe 12.4.
- Dass $\mathcal{E}(A', x')$ tatsächlich das Loewner-John-Ellipsoid ist, erhält man durch elementares Nachrechnen (etwas fummelig).
- Die Aussage über den Quotient der Volumina erhält man ebenfalls durch elementares Nachrechnen (etwas leichter).

□

2. TRENNEN UND OPTIMIEREN

Voraussetzung: Sei \mathcal{K} eine Klasse von konvexen und kompakten Mengen. Angenommen für jedes $K \in \mathcal{K}$ kennen wir $x_0 \in \mathbb{R}^n$, $r, R > 0$ mit

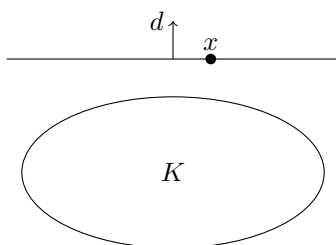
$$x_0 + rB_n \subseteq K \subseteq x_0 + RB_n.$$



Trennungsproblem:

Eingabe: $x \in \mathbb{R}^n$, $K \in \mathcal{K}$.

Ausgabe: „ $x \in K$ “ oder $d \in \mathbb{R}^n$ mit $d^T x > \max_{y \in K} d^T y$.



Die Hyperebene

$$\{y \in \mathbb{R}^n : d^T y = d^T x\}$$

trennt x und K .

Falls $K = \{x \in \mathbb{R}^n : Ax \leq b\}$ ein (beschränktes) Polyeder ist, dann kann man das Trennungsproblem wie folgt lösen: Überprüfe für gegebenes x alle linearen Ungleichungen

$$a_i^T x \leq b_i \quad i = 1, \dots, m.$$

Dann ist $x \in K$ genau dann, wenn alle Ungleichungen erfüllt sind. Falls $a_i^T x > b_i$, wähle $d = a_i$.

Optimierungsproblem

Eingabe: $c \in \mathbb{R}^n$, $\|c\| = 1$, $\varepsilon > 0$, $K \in \mathcal{K}$.

Ausgabe: $x \in K$ mit $c^T x \geq \max_{y \in K} c^T y - \varepsilon$.

Theorem 2.1. (Grötschel, Lovász, Schrijver, 1981; basierend auf Ellipsoidmethode) Für eine Klasse \mathcal{K} lässt sich das Optimierungsproblem durch N -faches Lösen des Trennungsproblems lösen, wobei N so gewählt ist, dass

$$2 \frac{R^2}{r} e^{-\frac{N}{2(n+1)^n}} \leq \varepsilon$$

gilt.

Beweis. Der Beweis folgt aus Lemma 3.2 später in diesem Kapitel. □

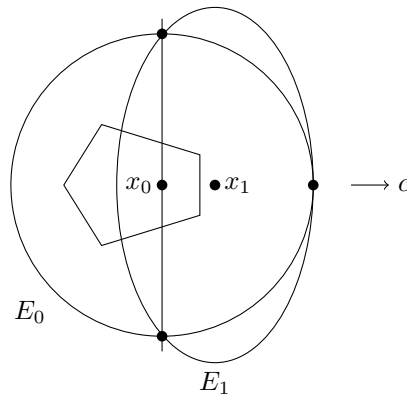
```

 $E_0 = \mathcal{E}(R^2 I_n, x_0)$ 
for  $k = 0$  to  $N - 1$  do
  Löse Trennungsproblem für  $x_k$ , den Mittelpunkt von  $E_k$ 
  if  $x_k \in K$  then
    |  $a = c$ 
  end
  else
    |  $a = -d$ 
  end
   $E_{k+1} = \mathcal{E}(E_k \cap \{y \in \mathbb{R}^n : a^T y \geq a^T x_k\})$ .
end

```

Algorithmus 1 : Ellipsoidmethode

3. ELLIPSOIDMETHODE



Lemma 3.1. *Seien*

$$\xi_k = \max\{c^T x_j : 0 \leq j \leq k, x_j \in K\}$$

und

$$K_k = K \cap \{x \in \mathbb{R}^n : c^T x \geq \xi_k\}.$$

Dann gilt $K_k \subseteq E_k$.

Beweis. \rightarrow Aufgabe 13.1. □

Lemma 3.2. *Sei $k \in \mathbb{N}$ und sei j ein Index mit $0 \leq j < k$ und $\xi_k = c^T x_j$. Dann gilt*

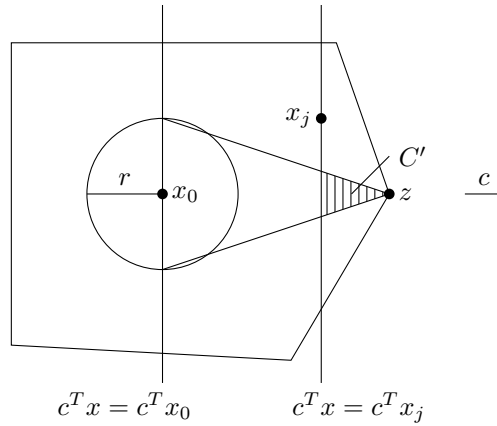
$$c^T x_j \geq \max_{y \in K} c^T y - 2 \frac{R^2}{r} e^{-\frac{k}{2(n+1)n}}.$$

Beweis. Sei $z \in K$ eine optimale Lösung, also $c^T z = \max_{y \in K} c^T y$. Betrachte den runden Kegelstumpf C mit Spitze z und Basis

$$x_0 + rB_n \cap \{y \in \mathbb{R}^n : c^T y = c^T x_0\}.$$

Sei

$$C' = C \cap \{x \in \mathbb{R}^n : c^T x \geq c^T x_j\}.$$



Das Volumen von C' ist

$$\text{vol } C' = \frac{r^{n-1} \cdot \text{vol } B_{n-1}}{n} (c^T z - c^T x_0) \left(\frac{c^T z - c^T x_j}{c^T z - c^T x_0} \right)^n$$

Nach dem vorhergehenden Lemma ist $C' \subseteq K_k \subseteq E_k$. Also

$$\begin{aligned} \text{vol } C' &\leq \text{vol } E_k \\ &\leq \text{vol } E_0 \cdot \left(e^{-\frac{1}{2n+1}} \right)^k \\ &< R^n \text{vol } B_n \cdot e^{-\frac{k}{2n+1}}. \end{aligned}$$

Die Ungleichung von Cauchy-Schwarz ergibt:

$$|c^T z - c^T x_0| \leq \underbrace{\|c\|}_{=1} \cdot \underbrace{\|z - x_0\|}_{\leq R} \leq R.$$

Zusammen:

$$(c^T z - c^T x_j)^n \leq R^n \text{vol } B_n e^{-\frac{k}{2n+1}} \cdot \frac{n \overbrace{(c^T z - c^T x_0)^{n-1}}^{\leq R}}{r^{n-1} \text{vol } B_{n-1}}.$$

Also, da $1 \leq \frac{R}{r}$:

$$c^T z - c^T x_j \leq \left(\frac{n \text{vol } B_n}{\text{vol } B_{n-1}} \right)^{\frac{1}{n}} \frac{R^2}{r} e^{-\frac{k}{2(n+1)n}}.$$

□

Aus dem letzten Lemma folgt das Theorem von Grötschel, Lovász und Schrijver unmittelbar.

Der hier vorgestellte Algorithmus ist an einer Stelle sehr idealisiert: Wir haben angenommen, dass wir mit unendlicher Genauigkeit rechnen können; aber in der Definition von b , das zur Bestimmung von E_k verwendet wird, wird eine Wurzel gezogen.

Der Algorithmus kann leicht verändert werden, so dass er mit vorgegebener Bitkomplexität läuft. Die Analyse wird aber deutlich technischer.

PROF. DR. F. VALLENTIN, DR. A. GUNDERT, MATHEMATISCHES INSTITUT, UNIVERSITÄT ZU KÖLN,
WEYERTAL 86-90, 50931 KÖLN, DEUTSCHLAND
E-mail address: `frank.vallentin@uni-koeln.de`, `anna.gundert@uni-koeln.de`

KAPITEL 10 — DIE INNERE-PUNKTE-METHODE

F. VALLENTIN, A. GUNDERT

Vorteile:

- + Löst effizient lineare Programme (in Theorie und Praxis)
- + erweiterbar (zu einer größeren Klasse von Optimierungsproblemen)
- + einfach zu implementieren

1. DER ZENTRALE PFAD

LP-Dualität

$$(LP) \quad \begin{array}{ll} \min & c^T x \\ & Ax = b \\ & x \geq 0 \end{array} = \begin{array}{ll} \max & b^T y \\ & A^T y \leq c \end{array}$$

Damit die Nebenbedingungen der rechten Seite die gleiche Form haben, wie auf der linken Seite, führen wir sogenannte Schlupfvariablen $s_1, \dots, s_n \in \mathbb{R}$, $s_1, \dots, s_n \geq 0$ ein, bzw. den Vektor $s = (s_1, \dots, s_n)^T \in \mathbb{R}^n$, sodass $A^T y + s = c$. Dann haben die beiden linearen Programme von oben die Form

$$\begin{array}{ll} \min & c^T x \\ & Ax = b \\ & x \geq 0 \end{array} = \begin{array}{ll} \max & b^T y \\ & A^T y + s = c \\ & s \geq 0 \end{array}$$

Wir interessieren uns für optimale Lösungen der beiden Programme und erinnern uns an Kapitel 5:

Optimalitätskriterium

$$\begin{aligned} (x^*, y^*, s^*) \text{ ist optimal} &\iff A^T y^* + s^* = c \\ &Ax^* = b \\ &x_i^* s_i^* = 0, \quad i = 1, \dots, n \\ &x^* \geq 0 \\ &s^* \geq 0 \end{aligned}$$

Idee: Betrachte die (nichtlineare) Abbildung

$$F: \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$$
$$F(x, y, s) = \begin{bmatrix} A^T y + s - c \\ Ax - b \\ X S e \end{bmatrix},$$

wobei $X = \text{diag}(x_1, \dots, x_n)$, $S = \text{diag}(s_1, \dots, s_n)$, $e = (1, \dots, 1)^T$, und finde (x^*, y^*, s^*) mit $F(x^*, y^*, s^*) = 0$, z.B. mit Hilfe des Newton-Verfahrens.

Problem: Es gibt viele (x^*, y^*, s^*) mit $F(x^*, y^*, s^*) = 0$, die nicht zur Menge der zulässigen Lösungen

$$\mathcal{F} = \{(x, y, s) : A^T y + s = c, Ax = b, x \geq 0, s \geq 0\}$$

gehören. Sei

$$\mathcal{F}^0 = \{(x, y, s) : A^T y + s = c, Ax = b, x > 0, s > 0\}$$

die Menge der *strikt zulässigen Lösungen*, das heißt x und s liegen im Inneren des nicht-negativen Orthanten

$$\mathbb{R}_{\geq 0}^n = \{z \in \mathbb{R}^n : z \geq 0\}.$$

Sei $\tau > 0$. Suche (x_τ, y_τ, s_τ) , $x_\tau > 0$, $s_\tau > 0$, mit

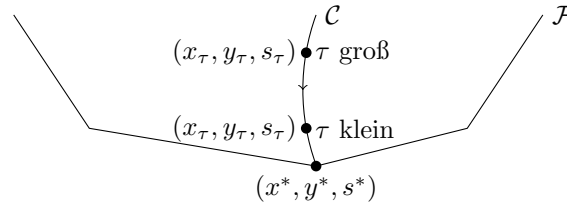
$$F(x_\tau, y_\tau, z_\tau) = \begin{bmatrix} 0 \\ 0 \\ \tau e \end{bmatrix}, \quad \text{d.h. } (x_\tau)_i (s_\tau)_i = \tau \quad \forall i = 1, \dots, n.$$

Wir zeigen gleich, dass $(x_\tau, y_\tau, s_\tau) \in \mathcal{F}^0$ existieren und eindeutig bestimmt sind.

Definition 1.1. Die Menge

$$\mathcal{C} = \{(x_\tau, y_\tau, s_\tau) : \tau > 0\}$$

heißt der zentrale Pfad von (LP).



Lemma 1.2. Es gilt

$$\lim_{\tau \rightarrow 0} c^T x_\tau - b^T y_\tau = 0.$$

Beweis.

$$\begin{aligned} c^T x_\tau - b^T y_\tau &= (A^T y_\tau + s_\tau)^T x_\tau - b^T y_\tau \\ &= y_\tau^T A x_\tau + s_\tau^T x_\tau - b^T y_\tau \\ &= y_\tau^T b + n\tau - b^T y_\tau \\ &= n\tau \rightarrow 0 \quad \text{für } \tau \rightarrow 0. \end{aligned}$$

□

Theorem 1.3. Der zentrale Pfad existiert und ist eindeutig. Mit anderen Worten: Falls $\mathcal{F}^0 \neq \emptyset$ und $\tau > 0$, dann $\exists!(x_\tau, y_\tau, s_\tau)$:

$$F(x_\tau, y_\tau, s_\tau) = \begin{bmatrix} 0 \\ 0 \\ \tau e \end{bmatrix}, \quad x_\tau > 0, s_\tau > 0.$$

Der Beweis benötigt noch etwas Vorarbeit.

Definition 1.4. *Definiere*

$$H^0 = \{(x, s) : \exists y : (x, y, s) \in \mathcal{F}^0\},$$

und die Barrierefunktion

$$f_\tau : H^0 \rightarrow \mathbb{R}$$

$$f_\tau(x, s) = \frac{1}{\tau} x^T s - \sum_{i=1}^n \ln(x_i s_i).$$

[Falls (x, s) gegen den Rand von H^0 läuft, dann $f_\tau(x, s) \rightarrow +\infty$].

Ziel: An (x_τ, s_τ) nimmt f_τ sein Minimum an.

Lemma 1.5. f_τ ist strikt konvex.

Beweis.

i) Der Term $x^T s$ ist linear in H^0 :

Sei \bar{x} mit $A\bar{x} = b$ gegeben. Dann gilt für $(x, s) \in H^0$:

$$\begin{aligned} x^T s &= x^T (c - A^T y) \\ &= x^T c - b^T y \\ &= x^T c - (A\bar{x})^T y \\ &= x^T c - \bar{x}^T A^T y \\ &= x^T c - \bar{x}^T (c - s). \end{aligned}$$

ii) Die Funktion $t \mapsto -\ln t$ ist strikt konvex, weil die zweite Ableitung $t \mapsto \frac{1}{t^2} > 0$ ist.

Aus i) und ii) folgt die Behauptung. □

Lemma 1.6. f_τ ist nach unten beschränkt, d.h.

$$\exists C : f_\tau(x, s) \geq C \quad \forall (x, s) \in H^0.$$

Beweis. \rightarrow Aufgabe 13.3. □

Lemma 1.7. Für $K \geq 0$ ist die Menge $\{(x, s) \in H^0 : x^T s \leq K\}$ beschränkt.

Beweis. Folgt unmittelbar aus Aufgabe 13.2. □

Lemma 1.8. Sei $\kappa \in \mathbb{R}$. Dann ist die Menge

$$\{(x, s) \in H^0 : f_\tau(x, s) \leq \kappa\}$$

in einer kompakten Menge enthalten.

Beweis. Für $g(t) = t - \ln t - 1$ ist

$$f_\tau(x, s) = \sum_{i=1}^n g\left(\frac{x_i s_i}{\tau}\right) + n - n \ln \tau.$$

Dann

$$f_\tau(x, s) \leq \kappa \Leftrightarrow \sum_{i=1}^n g\left(\frac{x_i s_i}{\tau}\right) \leq \bar{\kappa} = \kappa - n + n \ln \tau.$$

Für Index $i = 1, \dots, n$ gilt

$$\sum_{i=1}^n g\left(\frac{x_i s_i}{\tau}\right) \leq \bar{\kappa} - \sum_{\substack{j=1 \\ j \neq i}}^n g\left(\frac{x_j s_j}{\tau}\right) \leq \bar{\kappa},$$

weil $g(t) \geq 0$ für $t \geq 0$.

Weil $g(t) \rightarrow \infty$ für $t \rightarrow 0$ oder $t \rightarrow \infty$, existiert ein M , sodass

$$\frac{1}{M} \leq x_i s_i \leq M, \quad i = 1, \dots, n.$$

Also

$$x^T s = \sum_{i=1}^n x_i s_i \leq nM.$$

Nach Lemma 1.8 existiert ein M_u mit

$$x_i \leq M_u, \quad s_i \leq M_u.$$

Also

$$\frac{1}{M} \leq x_i s_i \leq M_u s_i \quad \Rightarrow \quad s_i \geq \frac{1}{MM_u}, \quad \text{genauso } x_i \geq \frac{1}{MM_u}.$$

Also sind die x_i und s_i nach oben und unten beschränkt und somit ist die Menge $\{(x, s) \in H^0 : f_\tau(x, s) \leq \kappa\}$ in einer kompakten Menge enthalten. \square

Nun können wir beweisen, dass der zentrale Pfad existiert und eindeutig ist.

Beweis. (Theorem 1.3)

Weil f_τ strikt konvex und die Menge $\{(x, s) \in H^0 : f_\tau(x, s) \leq \kappa\}$ in einer kompakten Menge enthalten ist, besitzt f_τ ein eindeutiges Minimum (x^*, s^*) .

Behauptung: $x^* = x_\tau$, $s^* = s_\tau$.

Beweis: (x^*, s^*) löst das Minimierungsproblem

$$\begin{aligned} \min f_\tau(x, s) \\ Ax = b \\ A^T y + s = c \\ x > 0, s > 0 \end{aligned}$$

(in den Variablen x, y, s). Insbesondere ist (x^*, s^*) ein lokales Minimum für

$$\begin{aligned} \min f_\tau(x, s) \\ Ax = b \\ A^T y + s = c. \end{aligned}$$

Nun wenden wir ein Theorem aus der Analysis an:

Theorem 1.9. (*Extrema mit Nebenbedingungen/Lagrange-Multiplikatoren*)

Sei $U \subseteq \mathbb{R}^n$ offen, seien $f: U \rightarrow \mathbb{R}$, $g_1, \dots, g_N: U \rightarrow \mathbb{R}$ stetig differenzierbar. Falls $x^* \in \mathbb{R}^n$ ein lokales Extremum von f unter den Nebenbedingungen $g_1(x^*) = \dots = g_N(x^*) = 0$ ist, dann gibt es $\lambda_1, \dots, \lambda_N \in \mathbb{R}$ (Multiplikatoren) mit

$$\nabla f(x^*) = \sum_{i=1}^N \lambda_i \nabla g_i(x^*).$$

Hier haben wir die Situation:

$$\begin{aligned}\nabla f_\tau(x, s) &= \begin{bmatrix} \frac{s}{\tau} - X^{-1}e \\ 0 \\ \frac{x}{\tau} - S^{-1}e \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} f_\tau(x, s) \\ \frac{\partial}{\partial y} f_\tau(x, s) \\ \frac{\partial}{\partial s} f_\tau(x, s) \end{bmatrix}, \\ g_1(x, y, s) &= [Ax - b]_1, \dots, g_m(x, y, s) = [Ax - b]_m, \\ g_{m+1}(x, y, s) &= [A^T y + s - c]_1, \dots, g_{m+n}(x, y, s) = [A^T y + s - c]_n.\end{aligned}$$

Also $\exists \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^n$:

$$\begin{aligned}\text{(a)} \quad & \frac{s}{\tau} - X^{-1}e = A^T \lambda \\ \text{(b)} \quad & 0 = A\mu \\ \text{(c)} \quad & \frac{x}{\tau} - S^{-1}e = \mu.\end{aligned}$$

Aus (b) und (c) folgt $A(\frac{x}{\tau} - S^{-1}e) = 0$.

(a) multipliziert mit $(\frac{x}{\tau} - S^{-1}e)^T$ liefert:

$$\underbrace{\left(\frac{x}{\tau} - S^{-1}e\right)^T}_{=0} A^T \lambda = \underbrace{\left(\frac{x}{\tau} - S^{-1}e\right)^T}_{\left(\frac{1}{\tau}Xe - S^{-1}e\right)^T} \underbrace{\left(\frac{s}{\tau} - X^{-1}e\right)}_{\left(\frac{1}{\tau}Se - X^{-1}e\right)}.$$

Definiere $X^{\frac{1}{2}} = \text{diag}(\sqrt{x_1}, \dots, \sqrt{x_n})$ und $S^{\frac{1}{2}} = \text{diag}(\sqrt{s_1}, \dots, \sqrt{s_n})$. Dann

$$\begin{aligned}0 &= \left(\frac{1}{\tau}Xe - S^{-1}e\right)^T \overbrace{\left(X^{-\frac{1}{2}}S^{\frac{1}{2}}\right)\left(X^{\frac{1}{2}}S^{-\frac{1}{2}}\right)}^{=I} \left(\frac{1}{\tau}Se - X^{-1}e\right) \\ &= \left\| \frac{1}{\tau}(XS)^{\frac{1}{2}}e - (XS)^{-\frac{1}{2}}e \right\|^2.\end{aligned}$$

$\Rightarrow XSe = \tau e$. □

Definition 1.10. Sei $(x, y, s) \in \mathcal{F}^0$. Das Dualitätsmaß μ von (x, y, s) ist

$$\mu = \frac{x^T s}{n} = c^T x - b^T y.$$

Definition 1.11. Sei $\theta \in [0, 1)$. Die Nachbarschaft des zentralen Pfads ist

$$N(\theta) = \{(x, y, s) \in \mathcal{F}^0 : \|XSe - \mu e\| \leq \theta \mu\}.$$

Lemma 1.12.

- a) $(x, y, s) \in \mathcal{C} \Leftrightarrow (x, y, s) \in N(0)$.
b) Falls $\|XSe - \mu e\| < \mu$, d.h. $\theta < 1$, dann ist $x > 0, s > 0$.

Beweis.

- a) klar.
b) Falls $x_i = 0$ oder $s_i = 0$, dann

$$\|XSe - \mu e\| \geq |x_i s_i - \mu| = \mu, \text{ aber } \theta < 1.$$

□

2. DER ALGORITHMUS UND SEINE ANALYSE

Betrachte

$$F(x, y, s) = \begin{bmatrix} A^T y + s - c \\ Ax - b \\ XSe \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \tau e \end{bmatrix}.$$

Sei $(x, y, s) \in \mathcal{F}^0$. Linearisiere F an (x, y, s) :

$$F(x, y, s) + \nabla F(\Delta x, \Delta y, \Delta s),$$

wobei

$$\nabla F(\Delta x, \Delta y, \Delta s) = \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix}$$

mit

$$(\Delta x, \Delta y, \Delta s) \in \mathbb{R}^{2n+m} \text{ und } F(x, y, s) = \begin{bmatrix} 0 \\ 0 \\ XSe \end{bmatrix}$$

Definition 2.1. Für den Zentrierungsparameter $\sigma \in [0, 1]$ definiere den modifizierten Newton-Schritt $(\Delta x, \Delta y, \Delta s) \in \mathbb{R}^{2n+m}$ durch das lineare Gleichungssystem

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -XSe + \sigma \mu e \end{bmatrix}.$$

Nun können wir den Algorithmus angeben:

Innere-Punkte-Methode:

Input : $\theta = 0.4$, $\sigma = 1 - \frac{0.4}{\sqrt{n}}$, Startpunkt $(x_0, y_0, s_0) \in N(\theta)$, $N =$ Anzahl Iterationen

for $k = 0$ **to** $N - 1$ **do**

Bestimme $(\Delta x_k, \Delta y_k, \Delta s_k)$ durch Löse des LGS

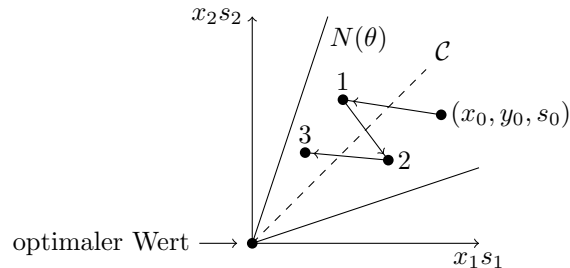
$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S_k & 0 & X_k \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta y_k \\ \Delta s_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -X_k S_k e + \sigma \mu_k e \end{bmatrix}$$

$(x_{k+1}, y_{k+1}, s_{k+1}) = (x_k, y_k, s_k) + (\Delta x_k, \Delta y_k, \Delta s_k)$.

end

Algorithmus 1 : Innere-Punkte-Methode

Bild dazu:



Lemma 2.2. Falls $(x_k, y_k, s_k) \in \mathcal{F}^0$, dann gilt

- a) $(\Delta x_k)^T (\Delta s_k) = 0$.
 b) $\mu_k = \underbrace{\left(1 - \frac{0.4}{\sqrt{n}}\right)}_{\sigma} \mu_{k-1}$.

Beweis.

- a) \rightarrow Aufgabe 13.4.
 b)

$$\begin{aligned} x_{k+1}^T s_{k+1} &= (x_k + \Delta x_k)^T (s_k + \Delta s_k) \\ &= x_k^T s_k + x_k^T \Delta s_k + \Delta x_k^T s_k + \underbrace{\Delta x_k^T \Delta s_k}_{=0} \\ &= \sigma x_k^T s_k, \end{aligned}$$

denn aus der dritten Zeile des LGS im Algorithmus ergibt sich

$$(1) \quad S_k \Delta x_k + X_k \Delta s_k = -X_k S_k e + \sigma \mu_k e,$$

und durch Aufsummieren der Einträge in (1) erhält man

$$s_k^T \Delta x_k + x_k^T \Delta s_k = -x_k^T s_k + n \sigma \mu_k,$$

also

$$x_k^T s_k + s_k^T \Delta x_k + x_k^T \Delta s_k = n \sigma \mu_k = n \sigma \frac{x_k^T s_k}{n} = \sigma x_k^T s_k.$$

□

Beispiel 2.3. $\mu_0 = 10^6$, $n = 10000$, $N = 5000$. Dann

$$\mu_N = \left(1 - \frac{0.4}{\sqrt{10000}}\right)^N \mu_0 = (0.996)^N \mu_0 \leq 0.002$$

Jetzt fehlt nur noch zu zeigen, dass $(x_k, y_k, s_k) \in \mathcal{F}^0$. Wir zeigen, dass (x_k, y_k, s_k) sogar in $N(\theta)$ liegt.

Lemma 2.4. Seien $u, v \in \mathbb{R}^n$, $u^T v \geq 0$. Dann gilt

$$\|UVe\| \leq 2^{-\frac{3}{2}} \|u + v\|^2.$$

Beweis. Es ist

$$\begin{aligned} 0 \leq u^T v &= \sum_{u_i v_i > 0} u_i v_i + \sum_{u_i v_i < 0} u_i v_i \\ &= \sum_{i \in P} |u_i v_i| - \sum_{i \in M} |u_i v_i|, \end{aligned}$$

wobei

$$P = \{i : u_i v_i > 0\}, \quad M = \{i : u_i v_i < 0\},$$

also

$$(*) \quad \sum_{i \in M} |u_i v_i| \leq \sum_{i \in P} |u_i v_i|.$$

Somit

$$\begin{aligned}
\|UVe\| &= (\|(u_i v_i)_{i \in P}\|^2 + \|(u_i v_i)_{i \in M}\|^2)^{\frac{1}{2}} \\
&\stackrel{(1)}{\leq} (\|(u_i v_i)_{i \in P}\|_1^2 + \|(u_i v_i)_{i \in M}\|_1^2)^{\frac{1}{2}} \\
&\stackrel{(*)}{\leq} (2\|(u_i v_i)_{i \in P}\|_1^2)^{\frac{1}{2}} \\
&= \sqrt{2}\|(u_i v_i)_{i \in P}\|_1 \\
&\stackrel{(2)}{\leq} \sqrt{2}\left\|\left(\frac{1}{4}(u_i + v_i)^2\right)_{i \in P}\right\|_1 \\
&= 2^{-\frac{3}{2}} \sum_{i \in P} (u_i + v_i)^2 \\
&\leq 2^{-\frac{3}{2}} \|u + v\|^2.
\end{aligned}$$

Dabei gilt (1), weil $\|\cdot\|_2 \leq \|\cdot\|_1$, und (2) gilt, da für $\alpha, \beta \in \mathbb{R}$ stets $\alpha\beta \leq \frac{1}{4}(\alpha + \beta)^2$ erfüllt ist. \square

Lemma 2.5. Sei $(x_k, y_k, z_k) \in N(0.4)$. Dann gilt

$$\|X_{k+1}S_{k+1}e - \mu_{k+1}e\| \leq 0.2\mu_k.$$

D.h. insbesondere $(x_{k+1}, y_{k+1}, s_{k+1}) \in N(0.4)$.

Beweis.

1.Behauptung: $\|X_{k+1}S_{k+1}e - \mu_{k+1}e\| = \|\Delta X_k \Delta S_k e\|$.

Beweis: Betrachte den i -ten Eintrag der Vektoren:

$$\begin{aligned}
&(x_{k+1})_i (s_{k+1})_i - \mu_{k+1} \\
&= ((x_k)_i + (\Delta x_k)_i) ((s_k)_i + (\Delta s_k)_i) - \underbrace{\sigma \mu_k}_{\text{Lemma 2.2.b)}} \\
&= (x_k)_i (s_k)_i + (x_k)_i (\Delta s_k)_i + (\Delta x_k)_i (s_k)_i + (\Delta x_k)_i (\Delta s_k)_i - \sigma \mu_k \\
&= (x_k)_i (s_k)_i - (x_k)_i (s_k)_i + \sigma \mu_k + (\Delta x_k)_i (\Delta s_k)_i - \sigma \mu_k \\
&= (\Delta x_k)_i (\Delta s_k)_i
\end{aligned}$$

2.Behauptung: $\|\Delta X_k \Delta S_k e\| \leq 0.2\mu_k$.

Beweis: Definiere $D_k = X_k^{\frac{1}{2}} S_k^{-\frac{1}{2}}$. Dann gilt

$$\begin{aligned}
\|\Delta X_k \Delta S_k e\| &= \|D_k^{-1} \Delta X_k D_k \Delta S_k e\| \\
&\leq 2^{-\frac{3}{2}} \|D_k^{-1} \Delta x_k + D_k \Delta s_k\|^2.
\end{aligned}$$

nach Lemma 2.4. Nun ist

$$D_k^{-1} \Delta x_k + D_k \Delta s_k = (X_k S_k)^{-\frac{1}{2}} (-X_k S_k e + \sigma \mu_k e),$$

weil

$$S_k \Delta x_k + X_k \Delta s_k = -X_k S_k e + \sigma \mu_k e$$

gilt, und dies multipliziert mit $(X_k S_k)^{-\frac{1}{2}}$ liefert die obige Gleichung. Also

$$\begin{aligned} \|\Delta X_k \Delta S_k e\| &\leq 2^{-\frac{3}{2}} \sum_{i=1}^n \frac{-(x_k)_i (s_k)_i + \sigma \mu_k}{(x_k)_i (s_k)_i} \\ &\leq 2^{-\frac{3}{2}} \frac{\|X_k S_k e - \sigma \mu_k e\|^2}{\min_{i=1, \dots, n} (x_k)_i (s_k)_i}. \end{aligned}$$

Da $(x_k, y_k, s_k) \in N(\theta)$, gilt $\min_{i=1, \dots, n} (x_k)_i (s_k)_i \geq (1 - \theta) \mu_k$, denn $\|X_k S_k e - \mu_k e\| \leq \theta \mu_k \Rightarrow |(x_k)_i (s_k)_i - \mu_k| \leq \theta \mu_k \Rightarrow (x_k)_i (s_k)_i - \mu_k \geq -\theta \mu_k$.

Desweiteren

$$e^T (X_k S_k e - \mu_k e) = x_k^T s_k - n \mu_k = 0.$$

Also

$$\begin{aligned} &\|X_k S_k e - \sigma \mu_k e\|^2 \\ &= \|(X_k S_k e - \mu_k e) + (1 - \sigma) \mu_k e\|^2 \\ &= \|X_k S_k e - \mu_k e\|^2 + 2(1 - \sigma) \mu_k e^T (X_k S_k e - \mu_k e) + (1 - \sigma)^2 \mu_k^2 e^T e \\ &\leq \theta^2 \mu_k^2 + (1 - \sigma)^2 \mu_k^2 n. \end{aligned}$$

Zusammen

$$\begin{aligned} \|\Delta X_k \Delta S_k e\| &\leq 2^{-\frac{3}{2}} \frac{\theta^2 \mu_k^2 + (1 - \sigma)^2 \mu_k^2 n}{(1 - \theta) \mu_k} \\ &\leq 0.2 \mu_k. \end{aligned}$$

□

LITERATUR

- [1] S.J. Wright, *Primal-dual interior-point-methods*, SIAM, 1997.

PROF. DR. F. VALLENTIN, DR. A. GUNDELT, MATHEMATISCHES INSTITUT, UNIVERSITÄT ZU KÖLN,
WEYERTAL 86–90, 50931 KÖLN, DEUTSCHLAND
E-mail address: frank.vallentin@uni-koeln.de, anna.gundert@uni-koeln.de