

Operations Research

Jason Uhing

15. Januar 2025

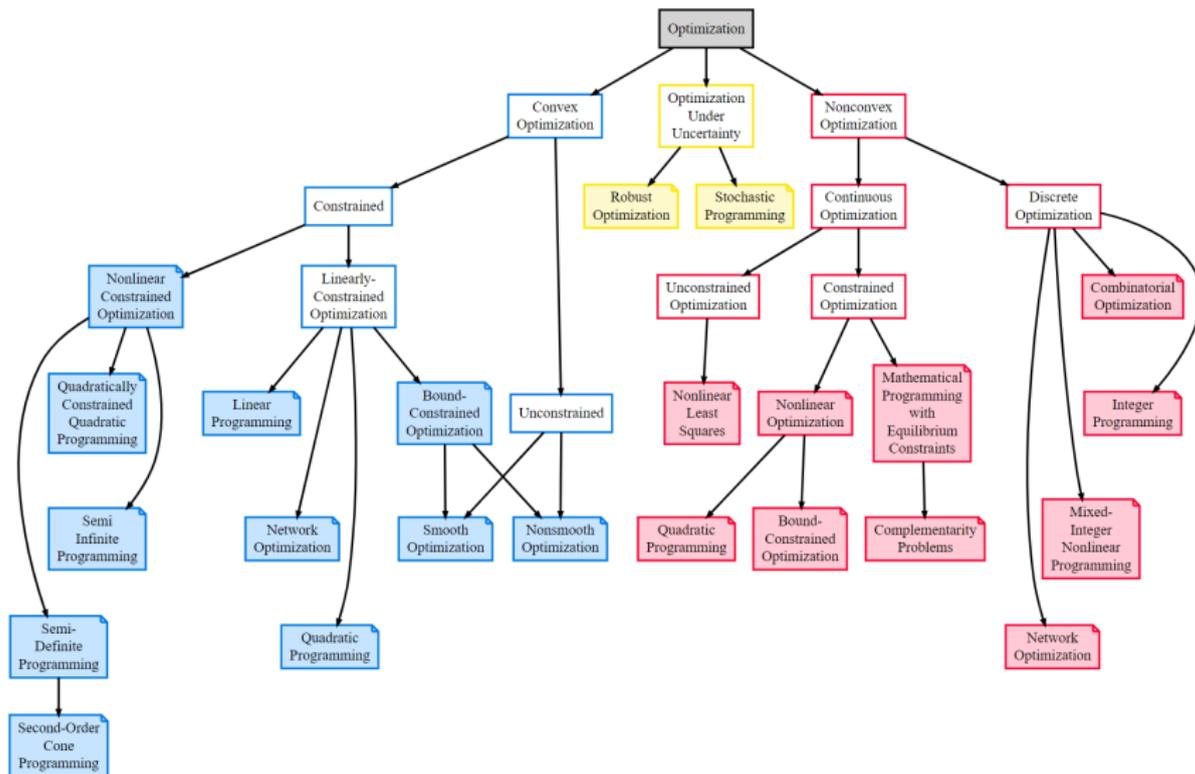
Was ist OR?

- 'Operations Research' (OR) als Disziplin unter diesem Namen war ab Mitte der 1930er Jahre als Methodologie militärischer Strategien entstanden (Einsatz der eigenen Luftstreitkräfte sowie den Schiffsverkehr optimal planen). Später hat sich die OR in vor allem in der Ökonomie verbreitet, z.B. unter dem Titel '*Unternehmensforschung*'.
- Aus der Sicht der Mathematik ist OR eine Mischung aus verschiedenen lange bekannten Techniken (lineare Algebra, Konvexitätsprobleme, Kombinatorik, Transportprobleme) abgestimmt auf die jeweilige Anwendung. Die Disziplin definiert sich also eher *nicht* durch ganz bestimmte Methoden, sondern durch ihr *Ziel*.
- Das *Ziel der OR* ist die Vorbereitung von möglichst guten Entscheidungen durch Anwendung mathematischer Methoden (stark praxisorientiert).

Was ist OR?

- Die *Hauptaufgabe* der OR ist die Übersetzung eines realen Entscheidungsproblems in ein Optimierungs- bzw. Simulationsmodell.
- *OR im engeren Sinne* bezieht sich auf mathematische Modellierung von Entscheidungsproblemen und Entwicklung von Algorithmen zur Anwendung und Lösung mathematischer Modelle. OR in diesem engeren Sinne hat viele Teilgebiete, z.B.
 - Lineare Optimierung (Maximierung unter linearen Nebenbedingungen)
 - Nichtlineare Optimierung (z.B. quadratisch)
 - Optimaler Transport
 - Spieltheorie
 - Graphentheorie und Netzplantechnik
 - Diskrete und kombinatorische Optimierung
 - Dynamische Optimierung
 - Warteschlangentheorie

Was ist OR?



Was lesen?

- K. H. Borgwardt, *Optimierung, Operations Research, Spieltheorie*, Springer Basel AG, 2001
- W. Domschke, A. Drexl, R. Klein, A. Scholl, *Einführung in Operations Research*, Springer, 2015.
- D. Jungnickel, *Graphs, Networks and Algorithms*, Springer, 2013. (Engl. Version von *Graphen, Netzwerke und Algorithmen*, Spektrum, 1994.)
- D. Jungnickel, *Optimierungsmethoden - Eine Einführung*, Springer, 2015.
- W. Hochstättler, *Lineare Optimierung*, Springer, 2012.
- H. Peters, *Game Theory, A Multi-Levelled Approach*, Springer 2015
- I. Wegener, *Operations Research*, Universität Dortmund, Skript zur Vorlesung WS 1998/99

Problemstellung

Beispiel 2.1.1

- (i) Ein Landwirt hat 100 ha Land. Davon kann er x_1 ha für den Anbau von Kartoffeln nutzen und x_2 ha für den Anbau von Getreide. Das *Ziel* ist es, x_1 und x_2 so zu wählen, dass der Gewinn maximal ausfällt.

	Kartoffeln/ha	Getreide/ha	insg. verfügbar
Anbaukosten	100	200	11000
Arbeitstage	1	4	160
Gewinn	400	1200	

Man erhält folgende Restriktionen:

Problemstellung

Beispiel 2.1.1 (Fortsetzung)

$$\begin{aligned}x_1, x_2 &\geq 0 && \text{(Fläche immer nichtnegativ)} \\100 x_1 + 200 x_2 &\leq 11000 && \text{(wegen Anbaukosten)} \\x_1 + 4 x_2 &\leq 160 && \text{(wegen Arbeitstagen)} \\x_1 + x_2 &\leq 100 && \text{(wegen max. verfügbarer Fläche).}\end{aligned}$$

Der entsprechende Gewinn ist

$$f(x_1, x_2) = 400 x_1 + 1200 x_2.$$

Wir formalisieren diesen Typ von Aufgabenstellung abstrakter. Zur Erinnerung: Eine lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}$ nennt man eine *Linearform* auf \mathbb{R}^n . Zu jeder Linearform f auf \mathbb{R}^n gibt es $a_1, \dots, a_n \in \mathbb{R}$, sodass

$$f(x) = a_1 x_1 + \dots + a_n x_n, \quad \text{für alle } x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Die Menge aller Linearformen auf \mathbb{R}^n bildet einen n -dimensionalen Vektorraum (den *Dualraum* zu \mathbb{R}^n).

Definition 2.1.2

Bei einem *allgemeinen linearen Optimierungsproblem* sind Zahlen $m, n \in \mathbb{N}$ sowie Linearformen f_1, \dots, f_m und f auf \mathbb{R}^n und Zahlen $c_1, \dots, c_m \in \mathbb{R}$ gegeben. Die Menge

$$P := \{x \in \mathbb{R}^n : f_i(x) \leq c_i \quad \text{für alle } i = 1, \dots, m\}$$

heißt *zulässiger Bereich*, die Bedingungen $f_i \leq c_i$, $i = 1, \dots, m$ nennt man die *Restriktionen*. Die Linearform f nennt man die *Zielfunktion*. Ein Punkt $y \in P$ heißt *Lösung* des Problems, falls er eine Maximalstelle der Zielfunktion f auf P ist, also falls

$$f(y) \geq f(x) \quad \text{für alle } x \in P.$$

Bemerkung 2.1.3

- (i) Möchte man statt einer Maximalstelle für f eine Minimalstelle finden, so kann man einfach $-f$ betrachten.

- (ii) Restriktionen der Form $f_i = c_i$ kann man ganz einfach umschreiben als $f_i \leq c_i$ und $-f_i \leq -c_i$.
- (iii) Umgekehrt kann man Restriktionen in Ungleichungsform durch Einführung einer zusätzlichen reellen Variablen in Restriktionen mit Gleichheit umwandeln: Man hat

$$f_i(x) \leq c_i \quad \text{genau dann, wenn} \quad f_i(x) + \tilde{x} = c_i \quad \text{und} \quad \tilde{x} \geq 0.$$

Das ist trivial, aber in der Praxis oft sinnvoll, z.B. kann man anstelle der Restriktion

$$\text{'Arbeitszeit} \leq \text{verfügbare Zeit'}$$

die Restriktion

$$\text{'Arbeitszeit} + \text{Freizeit} = \text{verfügbare Zeit'}$$

in das Modell aufnehmen.

Polyeder

Definition 2.2.1

Eine Teilmenge $A \subset \mathbb{R}^n$ heisst *konvex*, falls für alle $x, y \in A$ die Verbindungsstrecke

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

in A enthalten ist.

Wir nutzen auch die Schreibweisen

$$]x, y[:= [x, y] \setminus \{x, y\}, \quad]x, y] := [x, y] \setminus \{x\}, \quad \text{und} \quad [x, y[:= [x, y] \setminus \{y\}.$$

Beispiel 2.2.2

Sei f eine Linearform auf \mathbb{R}^n , $f \neq 0$ und sei $c \in \mathbb{R}$. Dann ist die Menge

$$\{f \leq c\} := \{x \in \mathbb{R}^n : f(x) \leq c\}$$

konvex.

Geometrisch ist die Situation wie folgt: Die Menge

$$\{f = c\} := \{x \in \mathbb{R}^n : f(x) = c\} = \{f \leq c\} \cup \{f \geq c\}$$

ist eine Hyperebene im Raum \mathbb{R}^n , welche letzteren in zwei Halbräume 'teilt', nämlich in $\{f \leq c\}$ und $\{f \geq c\}$.

Definition 2.2.3

Ein *abgeschlossener Halbraum* in \mathbb{R}^n ist eine Teilmenge, die sich in der Form $\{f \leq c\}$ darstellen lässt mit einer Linearform f auf \mathbb{R}^n und einer Konstanten $c \in \mathbb{R}$.

Definition 2.2.4

Ein endlicher Durchschnitt abgeschlossener Halbräume in \mathbb{R}^n heisst ein *Polyeder*.

Lemma 2.2.5

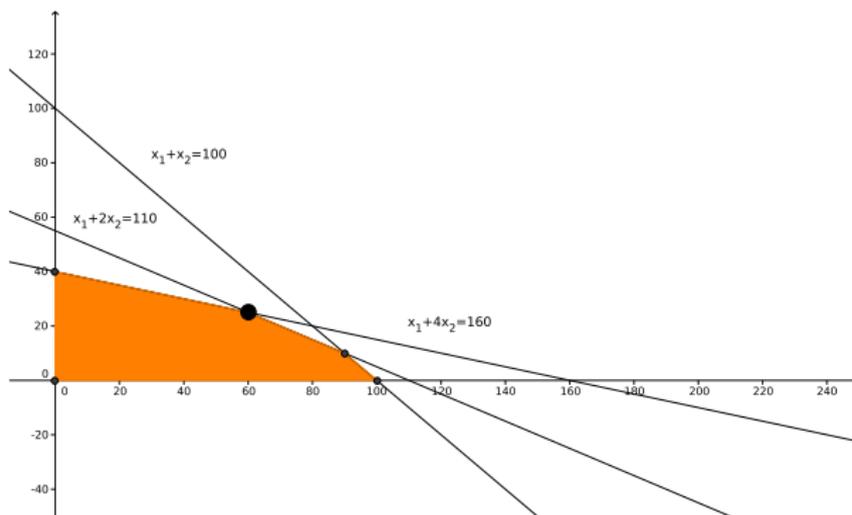
Sei $I \neq \emptyset$ und $\{A_i\}_{i \in I}$ eine Familie konvexer Mengen $A_i \subset \mathbb{R}^n$. Dann ist auch $\bigcap_{i \in I} A_i$ konvex.

Korollar 2.2.6

Jedes Polyeder ist konvex und abgeschlossen.

Beispiel 2.2.7

In der Situation von Beispiel 2.1.1 (i) sind die Hyperebenen Geraden, und gemäss den Restriktionen muss der zulässige Bereich das farbige Polyeder sein, das 'unterhalb' aller drei Geraden liegt und zudem von unten und von links durch die Achsen begrenzt wird:



Ist $A \subset \mathbb{R}^n$ kompakt, so nimmt eine stetige Funktion $f : A \rightarrow \mathbb{R}$ auf A ihr Maximum an. Das bedeutet, es existiert ein Punkt $y \in A$ sodass $f(y) \geq f(x)$ für alle $x \in A$. Ein solches y nennt man dann eine Maximalstelle für f auf A . Da ein beschränktes Polyeder $P \neq \emptyset$ im \mathbb{R}^n kompakt und eine Linearform f auf \mathbb{R}^n stetig ist, nimmt $f|_P$ auf P ihr Maximum an, also:

es gibt einen Punkt $y \in P$ sodass $f(y) \geq f(x)$ für alle $x \in P$.

Proposition 2.2.8

Ist P ein beschränktes Polyeder im \mathbb{R}^n und nimmt eine Linearform $f \neq 0$ ihr Maximum auf P in einem Punkt $y \in P$ an, so kann y kein innerer Punkt von P sein.

Definition 2.2.9

Sei $A \subset \mathbb{R}^n$ konvex.

- (i) Eine Teilmenge S von A ist eine *Seite von A* , falls S konvex ist und für zwei beliebige Punkte $x, y \in A$ folgendes gilt: Gibt es ein $0 < \lambda < 1$, sodass $\lambda x + (1 - \lambda)y \in S$, dann hat man $x, y \in S$.
- (ii) Ein Punkt $x \in A$ heisst *extremal*, falls $\{x\}$ eine Seite von A ist.

Wir schreiben A_e für die Menge der Extrempunkte von A .

Wie üblich schreiben wir $B(x, r) = \{y \in \mathbb{R}^n : |x - y| < r\}$ für die offene Kugel mit Radius $r > 0$ und Mittelpunkt x . Wir schreiben $\overline{B(x, r)}$ oder $\text{cl}(B(x, r))$ für ihren Abschluss, d.h. die abgeschlossene Kugel mit Radius $r > 0$ und Mittelpunkt x . Die Sphäre $\partial B(x, r)$ mit Radius $r > 0$ und Mittelpunkt x bezeichnen wir auch mit $S(x, r)$.

Beispiel 2.2.10

- (i) Die Verbindungsgerade zwischen zwei Eckpunkten eines abgeschlossenen Dreiecks ist eine Seite.
- (ii) Jeder Eckpunkt eines Polyeders ist ein Extrempunkt.
- (iii) Die Sphäre $S(x, r)$ ist die Menge der Extrempunkte der abgeschlossenen Kugel $\overline{B(x, r)}$.

Bemerkung 2.2.11

- (i) Für jede konvexe Menge A sind \emptyset und A (triviale) Seiten.
- (ii) Ist S eine Seite von A , so ist auch $A \setminus S$ konvex, denn:

Lemma 2.2.12

Sei $A \subset \mathbb{R}^n$ konvex. Die folgenden Aussagen sind äquivalent:

- (i) x ist ein Extrempunkt von A .
- (ii) x ist nur trivial konvex kombinierbar aus Punkten von A , d.h. falls $x = \lambda y + (1 - \lambda)z$ ist mit $y, z \in A$ und $0 < \lambda < 1$, so muss $y = z = x$ sein.
- (iii) $A \setminus \{x\}$ konvex.

Für Linearformen beobachten wir nun folgenden Effekt.

Satz 2.2.13

Sei $A \subset \mathbb{R}^n$ konvex, f eine Linearform auf \mathbb{R}^n und $\alpha \in \mathbb{R}$ mit $f \leq \alpha$ auf A . Dann ist

$$S := A \cap \{f = \alpha\}$$

eine Seite von A . Ist A ein Polyeder, so ist auch S ein Polyeder.

Korollar 2.2.14

Ist P der zulässige Bereich eines allgemeinen linearen Optimierungsproblems und f seine Zielfunktion, dann nimmt $f|_P$ ihr Maximum auf einer Seite von P an.

Wir betrachten als nächstes die Extrempunkte von Seiten.

Satz 2.2.15

Sei S die Seite einer konvexen Menge $A \subset \mathbb{R}^n$. Dann gilt

$$S_e = A_e \cap S.$$

Definition 2.2.16

Für $A \subset \mathbb{R}^n$ heisst

$$k(A) := \bigcap_{A \subset B, B \text{ konvex}} B$$

die *konvexe Hülle* von A .

Bemerkung 2.2.17

- (i) $k(A)$ ist wohldefiniert für alle $A \subset \mathbb{R}^n$: Mit $B = \mathbb{R}^n$ existiert stets eine konvexe Menge mit $A \subset B$.
- (ii) Nach Lemma 2.2.5 ist $k(A)$ konvex, und nach Konstruktion ist $k(A)$ die kleinste konvexe Menge, die A enthält.

- (iii) A ist genau dann konvex, wenn $k(A) = A$, denn:
- (iv) $k(A)$ ist genau dann beschränkt, wenn A beschränkt ist, denn:

Satz 2.2.18

Sei $A \subset \mathbb{R}^n$. Dann gilt

$$k(A) = \left\{ \sum_{i=1}^m \lambda_i a_i : m \in \mathbb{N}, a_1, \dots, a_m \in A, \lambda_1, \dots, \lambda_m \geq 0, \sum_{i=1}^m \lambda_i = 1 \right\}.$$

Für endlich viele Punkte $a_1, \dots, a_m \in \mathbb{R}^n$ und $\lambda_1, \dots, \lambda_m \geq 0$ mit $\sum_{i=1}^m \lambda_i = 1$ nennt man

$$\sum_{i=1}^m \lambda_i a_i$$

eine *Konvexkombination* der Punkte a_1, \dots, a_m . Satz 2.2.18 sagt also, dass die konvexe Hülle $k(A)$ die Menge aller Konvexkombinationen von endlich vielen Punkten aus A ist.

Korollar 2.2.19

Für jede endliche Menge $A \subset \mathbb{R}^n$ ist $k(A)$ kompakt.

Beispiel 2.2.20

Man hat $k(B(0, 1)) = B(0, 1)$, eine offene Menge.

Satz 2.2.21 (Satz von Krein-Milman für Polyeder)

Sei $P \subset \mathbb{R}^n$ ein beschränktes Polyeder. Dann gilt $P = k(P_e)$.

Satz 2.2.22 (Satz von Carathéodory)

Sei $P \subset \mathbb{R}^n$ ein beschränktes Polyeder. Dann ist jeder Punkt $x \in P$ Konvexkombination von höchstens $n + 1$ Punkten aus P_e .

Bemerkung 2.2.23

Satz 2.2.21 und Satz 2.2.22 gelten für beliebige kompakte konvexe Mengen in \mathbb{R}^n .

Korollar 2.2.24

Sei $P \subset \mathbb{R}^n$ ein nichtleeres beschränktes Polyeder und f eine Linearform auf \mathbb{R}^n . Dann gibt es ein $x_0 \in P_e$ mit

$$f(x) \leq f(x_0) \quad \text{für alle } x \in P.$$

Das heisst, mindestens ein Extrempunkt von P muss eine Maximalstelle für $f|_P$ sein. Wir kommen zu folgendem Fazit.

Bemerkung 2.2.25

Die Menge aller Punkte in P , in denen f ihr Maximum auf P annimmt, ist eine Seite P' von P und wegen $P' = k((P')_e) = k(P_e \cap P')$ die konvexe Hülle derjenigen Extrempunkte von P , in denen das Maximum angenommen wird.

Bestimmung von Extrempunkten

Im Folgenden seien f_1, \dots, f_m nichttriviale Linearformen auf \mathbb{R}^n , $c_1, \dots, c_m \in \mathbb{R}$ und

$$P = \bigcap_{i=1}^m \{f_i \leq c_i\}.$$

Die folgende Beobachtung reduziert das Auffinden von Extrempunkten auf die Lösung linearer Gleichungssysteme.

Satz 2.3.1

(i) Jeder Extrempunkt x von P ist Lösung eines linearen Gleichungssystems

$$f_{i_k}(x) = c_{i_k}, \quad 1 \leq k \leq n, \quad (1)$$

mit n Gleichungen und n Unbekannten, wobei $1 \leq i_k \leq m$ und f_{i_1}, \dots, f_{i_n} linear unabhängige Linearformen sind. Insbesondere hat man $P_e = \emptyset$ falls $m < n$ und $\#P_e \leq \binom{m}{n}$ falls $m \geq n$.

(ii) Sind f_{i_k} , $1 \leq k \leq n$ linear unabhängig und ist die Lösung x von (1) in P enthalten, so ist x auch ein Extrempunkt von P , d.h. $x \in P_e$.

Die zweite Aussage in (i) kann man sich gut geometrisch vorstellen:

- Für $n = 1$ braucht man mindestens eine Restriktion, also $m = 1$, um einen Extrempunkt zu haben. Hat man zwei Restriktionen, kann man bestenfalls zwei Extrempunkte bekommen.
- Für $n = 2$ sind $m = 1$ Restriktionen zu wenig, um einen Extrempunkt zu bekommen (denn dann ist das Polyeder ein Halbraum). Für $m = 2$ ist (bei guter Konstellation) bereits einen Extrempunkt möglich.
- Ist $n = 2$ und P ein Dreieck, das als Schnitt dreier abgeschlossener Halbräume entsteht, also $m = 3$, so löst jeder Eckpunkt von P ein System (1) bestehend aus zwei Gleichungen (denn er ist Schnittpunkt von genau zwei der drei Geraden.)

Bemerkung 2.3.2

Satz 2.3.1 zeigt, wie man Extrempunkte eines Polyeders P finden kann: Wir wählen $1 \leq i_1 < i_2 < \dots < i_n \leq m$ und lösen das Gleichungssystem

$$f_{i_k}(x) = c_{i_k}, \quad 1 \leq k \leq n.$$

Ist x die eindeutige Lösung dieses Systems und gehört x zu P , so muss x ein Extrempunkt von P sein, also $x \in P_e$. Auf diese Weise kann man *alle* Extrempunkte von P finden.

Korollar 2.3.3

Sei $x \in P_e$. Dann gibt es eine Linearform f auf \mathbb{R}^n mit

$$f(y) < f(x) \quad \text{für alle } y \in P \setminus \{x\}.$$

Bemerkung 2.3.4

Geometrisch bedeutet Korollar 2.3.3, dass es zu $x \in P_e$ einen offenen Halbraum H gibt (nämlich $H = \{f < c_{i_1} + \dots + c_{i_n}\}$ in der Notation des Beweises), sodass $P \setminus \{x\} \subset H$ und $x \in \partial H$.

Man kann mit linearen Gleichungen auch die Lage von Extrempunkten relativ zueinander beschreiben.

Definition 2.3.5

Zwei verschiedene Extrempunkte x und y eines Polyeders P heißen *benachbart*, falls $[x, y]$ eine Seite von P ist.

Satz 2.3.6

Sei $x \in P_e$ und seien $1 \leq i_1 < i_2 < \dots < i_n \leq m$ so, dass f_{i_1}, \dots, f_{i_n} linear unabhängig sind und

$$f_{i_k}(x) = c_{i_k}, \quad 1 \leq k \leq n.$$

Ist $y \in P_e \setminus \{x\}$ und gilt

$$f_{i_k}(y) = c_{i_k} \quad \text{für } n - 1 \text{ der möglichen Werte von } k,$$

so sind x und y benachbart.

Wenn wir uns wieder den Fall eines Dreiecks P in der Ebene ($n = 2$) vorstellen, wird die Grundidee plausibel:

Zwei verschiedene Eckpunkte sind (im Fall des Dreiecks) benachbart, es gibt also genau eine Gerade (hier in der Rolle einer Hyperebene), auf der beide liegen. Also muss genau eine der jeweils zwei Gleichungen, die diese Punkte beschreiben, von beiden Punkten gelöst werden (nämlich jene, die die Gerade beschreibt). Satz 2.3.6 schliesst nun von den Gleichungen auf die geometrische Lage.

Umgekehrt gilt nun folgender Schluss von der geometrischen Lage auf die Gleichungen.

Satz 2.3.7

Seien x und y benachbarte Extrempunkte eines Polyeders P . Dann gibt es n linear unabhängige Linearformen f_{i_1}, \dots, f_{i_n} mit

$$f_{i_k}(x) = c_{i_k}, \quad 1 \leq k \leq n,$$

und

$$f_{i_k}(y) = c_{i_k} \quad \text{für } n-1 \text{ der möglichen Werte von } k.$$

Bemerkung 2.3.8

Im Allgemeinen kann man nicht hoffen, mithilfe nur einer festen Familie $\{f_{i_k}\}_{1 \leq k \leq n}$ alle zu einem gegebenen Punkt $x \in P_e$ benachbarten Extrempunkte zu finden.

Im Folgenden wollen wir den Begriff 'benachbart' für Extrempunkte benutzen, um leichter entscheiden zu können, ob ein Extrempunkt tatsächlich eine Maximalstelle für eine gegebene Linearform auf dem gesamten Polyeder ist.

- Eine (konvexe) Menge $C \subset \mathbb{R}^n$ die für alle $\lambda \geq 0$ die Menge

$$\lambda C := \{\lambda x : x \in C\}$$

enthält, nennt man einen (*konvexen*) *Kegel*.

- Für eine gegebene konvexe Menge $K \subset \mathbb{R}^n$ ist

$$\mathbb{R}_+ K := \{\lambda x : \lambda \geq 0, x \in K\}$$

eine konvexer Kegel, tatsächlich der kleinste konvexe Kegel, der K enthält.

Lemma 2.3.9

Sei P ein beschränktes Polyeder mit $0 \in P_e$, aber $P \neq \{0\}$, und bezeichne P_e^0 die Menge aller zu 0 benachbarten Extrempunkte von P . Dann hat man

$$P \subset \mathbb{R}_+ k(P_e^0).$$

Bemerkung 2.3.10

$$\begin{aligned} \mathbb{R}_+ k(P_e^0) &= \{\lambda x : \lambda \geq 0, x \in k(P_e^0)\} \\ &= \left\{ \sum_{i=1}^m \lambda_i y_i : m \in \mathbb{N}, \lambda_i \geq 0, y_i \in P_e^0 \right\}. \end{aligned}$$

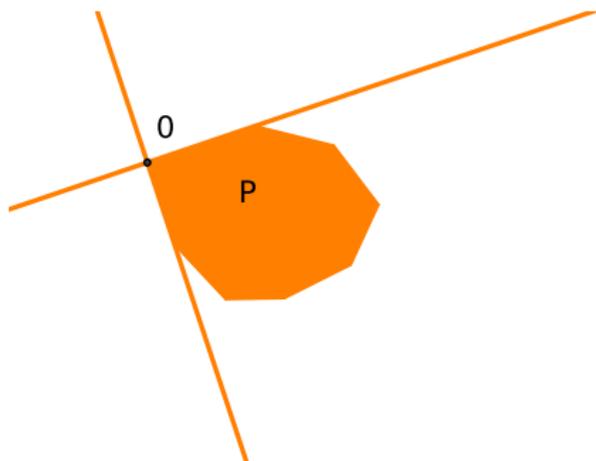


Abbildung: Polyeder als Teilmenge des Kegels

Bevor wir Lemma 2.3.9 beweisen, schauen wir uns folgende praktisch relevante Konsequenz an.

Korollar 2.3.11

Sei P ein beschränktes Polyeder und f eine Linearform auf \mathbb{R}^n . Ist $x \in P_e$ und gilt $f(y) \leq f(x)$ für alle $y \in P_e$, die zu x benachbart sind, so folgt, dass

$$f(x) = \max f(P)$$

- Das bedeutet, um zu entscheiden, ob ein Extrempunkt x eine Maximalstelle von $f|_P$ ist, müssen wir lediglich $f(x)$ mit den Werten $f(y)$ in *benachbarten* Extrempunkten y vergleichen.
- In der Praxis hat man es oft mit sehr vielen Restriktionen zu tun und daher mit sehr vielen Extrempunkten. Das Korollar kommt uns entgegen und reduziert die Komplexität der Frage, ob x Maximalstelle von $f|_P$ ist oder nicht.

Auflösen linearer Gleichungssysteme

Wir betrachten das lineare Gleichungssystem

$$a_{11} x_1 + \cdots + a_{1n} x_n = c_1$$

$$a_{21} x_1 + \cdots + a_{2n} x_n = c_2$$

$$\vdots$$

$$a_{m1} x_1 + \cdots + a_{mn} x_n = c_m$$

mit gegebenen $a_{ij}, c_i \in \mathbb{R}$. Für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ und $1 \leq i \leq m$ setzen wir

$$y_i(x) := -a_{i1} x_1 - \cdots - a_{in} x_n + c_i. \quad (2)$$

Das definiert jeweils eine affine Abbildung $y_i : \mathbb{R}^n \rightarrow \mathbb{R}$, und im Falle $c_i = 0$ eine Linearform. Die Lösungsmenge des Gleichungssystems ist dann

$$\{x \in \mathbb{R}^n : y_1(x) = \cdots = y_m(x) = 0\}.$$

Die Grundidee für einen Lösungsalgorithmus ist nun, die x_1, \dots, x_n durch äquivalente Umformungen als Funktionen der y_1, \dots, y_m auszudrücken und dann y_1, \dots, y_m gleich null zu setzen.

Man verwendet dabei die Darstellung der Relationen (2) durch folgendes Schema:

$$\begin{array}{c|ccc|c}
 & -x_1 & \cdots & -x_n & \\
 \hline
 y_1 & a_{11} & \cdots & a_{1n} & c_1 \\
 \vdots & \vdots & & \vdots & \vdots \\
 y_m & a_{m1} & \cdots & a_{mn} & c_m
 \end{array} \tag{3}$$

Ein *Gauss-Eliminationsschritt* besteht in folgendem:

- 1 Wähle ein (k, l) mit $a_{kl} \neq 0$. Ein solches a_{kl} nennt man ein *Pivotelement*.
- 2 Eliminiere x_l durch Auflösen der k -ten Gleichung nach x_l :

$$y_k = -a_{kl} x_l - \sum_{j \neq l} a_{kj} x_j + c_k$$

genau dann, wenn

$$x_l = -\frac{1}{a_{kl}} y_k - \sum_{j \neq l} \frac{a_{kj}}{a_{kl}} x_j + \frac{c_k}{a_{kl}}.$$

- 3 Eliminiere x_l aus den anderen Gleichungen durch Einsetzen dieses Ausdrucks für x_l . Für $i \neq k$ ergibt sich

$$\begin{aligned}
 y_i &= -a_{il} x_l - \sum_{j \neq l} a_{ij} x_j + c_i \\
 &= \frac{a_{il}}{a_{kl}} y_k - \sum_{j \neq k} \left(a_{ij} - \frac{a_{kj} a_{il}}{a_{kl}} \right) x_j + c_i - \frac{c_k a_{il}}{a_{kl}}.
 \end{aligned}$$

Dies führt zu folgendem, zu (3) äquivalenten Schema:

	$-x_1$	\dots	$-y_k$	\dots	$-x_n$	
y_1						\vdots
\vdots						\vdots
x_l	$\frac{a_{k1}}{a_{kl}} \dots$	\dots	$\frac{1}{a_{kl}}$	\dots	$\frac{a_{kn}}{a_{kl}}$	$\frac{c_k}{a_{kl}}$
\vdots						\vdots
y_i	$a_{i1} - \frac{a_{k1} a_{il}}{a_{kl}} \dots$	\dots	$-\frac{a_{il}}{a_{kl}}$	\dots	$a_{in} - \frac{a_{kn} a_{il}}{a_{kl}} \dots$	$c_i - \frac{c_k a_{il}}{a_{kl}}$
\vdots						\vdots
y_m						\vdots

(4)

Die k -te Zeile (die *Pivotzeile*) ist die Zeile, die mit x_l beginnt. Die l -te Spalte (die *Pivotspalte*) ist jene, die oben $-y_k$ stehen hat.

Das neue Schema (4) wird dabei nach folgenden *Pivotregeln* für den Austausch von k -ter Zeile und l -ter Spalte gebildet:

- (i) Das Pivotelement wird durch seinen Kehrwert ersetzt.
- (ii) Die übrigen Elemente der Pivotzeile werden durch das Pivotelement geteilt.
- (iii) Die übrigen Elemente der Pivotspalte werden durch das Pivotelement dividiert, und das Vorzeichen wird geändert.
- (iv) Alle übrigen Elemente werden nach der folgenden *Rechteckregel* verändert:
Ersetze in

$$\begin{array}{ccc} * & \dots & a \\ \vdots & & \vdots \\ b & \dots & \text{Pivot} \end{array}$$

den Eintrag $*$ durch

$$* - \frac{ab}{\text{Pivot}} = * + \frac{-a}{\text{Pivot}} b.$$

Man kann das nun wie folgt 'optisch vereinfachen':

- Der Quotient $\frac{-a}{\text{Pivot}}$ ist derjenige Eintrag, der im neuen Schema (4) in der l -ten Spalte (der Pivotspalte) auf der Höhe von $*$ steht.
- Es empfiehlt sich daher, als Zwischenschritt zunächst die neue Pivotspalte (ausser dem Kehrwert des Pivotelements) neben das alte Schema zu schreiben:

$$\begin{array}{c|ccc|c|c}
 & -x_1 & \cdots & -x_n & & \\
 \hline
 y_1 & a_{11} & \cdots & a_{1n} & c_1 & -\frac{a_{1l}}{a_{kl}} \\
 \vdots & \vdots & & \vdots & \vdots & \vdots \\
 y_m & a_{m1} & \cdots & a_{mn} & c_m & -\frac{a_{ml}}{a_{kl}}
 \end{array}$$

Dann ändert sich die Rechteckregel (iv) zur Berechnung der Elemente ausserhalb von Pivotzeile und -spalte zur folgenden *Dreiecksregel*: Ersetze in

$$\begin{array}{c}
 * \quad \dots \quad d := \frac{-a}{\text{Pivot}} \\
 \vdots \\
 b
 \end{array}$$

den Eintrag $*$ durch

$$* + db.$$

- Hier ist die letzte Spalte die neue, wie oben hinzugefügte Pivotspalte, und b ist das Element aus der Pivotzeile, das in derselben Spalte wie $*$ steht.
- Die neue Pivotspalte kann natürlich auch gleich an ihren Platz im neuen Schema geschrieben werden.

Beispiel 2.4.1

Das lineare Gleichungssystem

$$2x_1 + 4x_2 + x_3 = -4$$

$$3x_1 + 5x_2 + x_3 = 3$$

$$x_1 + x_2 + x_3 = 0$$

führt auf das (erweiterte) Schema

	$-x_1$	$-x_2$	$-x_3$		
y_1	2	4	1	-4	-4
y_2	3	5	1	3	-5
	1	1	1	0	

Als **Pivot** haben wir hier das Element $a_{32} = 1$ ausgewählt. Die **neue Pivotspalte** (abgesehen vom Kehrwert des Pivots) haben wir rechts hinzugefügt. Zum Beispiel ergibt sich

$$-\frac{a_{12}}{a_{32}} = -\frac{4}{1} = -4.$$

Wir tauschen nun die Bezeichnungen der 3. Zeile und der 2. Spalte (abgesehen vom Vorzeichen) aus und berechnen die **übrigen Elemente** (ausserhalb der **Pivotzeile und -spalte**) mit der Dreiecksregel. Zum Beispiel wird aus

$$* = a_{11} = 2$$

mit $b = a_{31} = 1$ und $d = -4$ der Eintrag

$$* + bd = 2 + 1 \cdot (-4) = -2.$$

Wir erhalten das neue Schema

	$-x_1$	$-y_3$	$-x_3$	
y_1	-2	-4	-3	-4
y_2	-2	-5	-4	3
x_2	1	1	1	0

Die Elemente der Pivotspalte werden nun nicht weiter benutzt, da für die Lösung am Ende $y_1 = y_2 = y_3 = 0$ gesetzt wird. Für den nächsten Gauss-Eliminationsschritt genügt daher das (erweiterte) Schema

	$-x_1$	$-y_3$	$-x_3$		
y_1	-2	#	-3	-4	
y_2	-2	#	-4	3	-1
x_2	1	#	1	0	0.5

Hier haben wir als neues Pivot -2 gewählt. Wir folgen nun wie zuvor den Pivotregeln und erhalten als nächstes Schema

	$-y_1$	$-y_3$	$-x_3$	
x_1	-0.5	#	1.5	2
y_2	-1	#	-1	7
x_2	0.5	#	-0.5	-2

Für den folgenden Eliminationsschritt genügt das (erweiterte) Schema

	$-y_1$	$-y_3$	$-x_3$		
x_1	#	#	1.5	2	1.5
y_2	#	#	-1	7	
x_2	#	#	-0.5	-2	-0.5

Als neues Pivot haben wir -1 gewählt. Mit den Pivotregeln ergibt sich nun das Schema

	$-y_1$	$-y_3$	$-y_2$	
x_1	#	#	1.5	$\frac{25}{2}$
x_3	#	#	-1	-7
x_2	#	#	-0.5	$-\frac{11}{2}$

Jetzt hat man also

	$-y_1$	$-y_3$	$-y_2$	
x_1	#	#	#	$\frac{25}{2}$
x_3	#	#	#	-7
x_2	#	#	#	$-\frac{11}{2}$

und damit die Lösung $x_1 = \frac{25}{2}$, $x_2 = -\frac{11}{2}$ und $x_3 = -7$ des Gleichungssystems.

Einschub: Lösung des LGS mit herkömmlicher Methode zusammen mit den y_i

Das Gauss-Eliminationsverfahren kann man auch zur Rangbestimmung für $(m \times n)$ -Matrizen oder zum Invertieren regulärer $(n \times n)$ -Matrizen verwenden.

Beispiel 2.4.2

Wir bestimmen den Rang der $(n \times m)$ -Matrix mit $n = 4$ und $m = 5$ im Schema

	$-x_1$	$-x_2$	$-x_3$	$-x_4$
y_1	4	3	2	1
y_2	3	-1	0	2
y_3	5	7	4	0
y_4	7	2	2	3
y_5	5	-6	-2	5

(Hier gibt es keine Spalte für Zahlen c_i , denn die sind ja nicht Teil dieser Aufgabenstellung.) Als Pivot ist hier 1 ausgewählt.

Austausch von y_1 und $-x_4$ nach den Pivotregeln ergibt das neue Schema

	$-x_1$	$-x_2$	$-x_3$	$-y_1$
x_4	4	3	2	1
y_2	-5	-7	-4	-2
y_3	5	7	4	0
y_4	-5	-7	-4	-3
y_5	-15	-21	-12	-5

Zum Beispiel ergibt sich für den Eintrag a_{21} der neue Wert $3 + 4(-2) = -5$ und für den Eintrag a_{43} der neue Wert $2 + 2(-3) = -4$. Als neues Pivot wählen wir nun **-5**. Austausch von y_2 und x_1 nach den Pivotregeln ergibt das neue Schema

	$-y_2$	$-x_2$	$-x_3$	$-y_1$
x_4	$\frac{4}{5}$	$\frac{-13}{5}$	$\frac{-6}{5}$	$\frac{-3}{5}$
x_1	$\frac{-1}{5}$	$\frac{7}{5}$	$\frac{4}{5}$	$\frac{2}{5}$
y_3	1	0	0	-2
y_4	-1	0	0	-1
y_5	-3	0	0	1

- Nun bricht das Verfahren aber ab: An allen Positionen, an denen neue Pivotelemente stehen könnten (d.h. für welche ein Tausch der Zeilen- und Spaltenbezeichnung strategisch Sinn machen würden), stehen Nullen.
- Die Linearformen y_3 , y_4 und y_5 sind Linearkombinationen von y_1 und y_2 :

$$y_3 = 2 y_1 - y_2$$

$$y_4 = y_1 + y_2$$

$$y_5 = -y_1 + 3y_2.$$

- Rang der Matrix = $\dim \operatorname{lin}\{y_1, \dots, y_5\} = \dim \operatorname{lin}\{y_1, y_2\} \leq 2$
- Die Linearformen im header der Tabelle (hier im Beispiel also x_2 , x_3 , y_2 und y_2) sind stets linear unabhängig, denn sie ergeben sich durch Basisaustausch im Dualraum. Da es stets n Elemente sind, bilden sie eine Basis des Dualraums. Insbesondere sind daher hier y_1 und y_2 linear unabhängig. Also

$$\text{Rang der Matrix} = 2.$$

- *Einschub*: Rang auf herkömmliche Art und Basistausch im Dualraum

Satz 2.4.3

Für eine reelle $(m \times n)$ -Matrix A sind folgende Aussagen äquivalent:

- (i) Man hat $\text{Rang}(A) = k$.
- (ii) Das Gaussverfahren (Austauschverfahren) bricht nach k Schritten ab.

Bemerkung 2.4.4

Ähnlich kann man mit dem Gaussverfahren die zu einer gegebenen $(n \times n)$ -Matrix inverse finden, falls sie existiert:

- Die Matrix ist genau dann invertierbar, wenn das Verfahren erst nach n Schritten abbricht. In diesem Falle stehen dann am Ende alle x_k oben und alle y_i links.
- Sortiert man nun das Schema nach der natürlichen Ordnung der Indizes, so ergibt sich genau die inverse Matrix (siehe ersten Einschub).

Erster Simplexalgorithmus

Gegeben sind Linearformen f_1, \dots, f_m und f auf \mathbb{R}^n und Zahlen $c_1, \dots, c_m \in \mathbb{R}$.
Gesucht ist das Maximum von f unter den Nebenbedingungen (Restriktionen)

$$x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0$$

und

$$f_1 \leq c_1, f_2 \leq c_2, \dots, f_m \leq c_m.$$

Die Linearformen besitzen Darstellungen

$$f_i(x) = a_{i1} x_1 + \dots + a_{in} x_n, \quad 1 \leq i \leq m,$$

und

$$f(x) = b_1 x_1 + \dots + b_n x_n$$

mit $a_{ij}, b_j \in \mathbb{R}$.

Für $1 \leq i \leq m$ setzen wir nun

$$y_i(x) := -f_i(x) + c_i = - \sum_{j=1}^n a_{ij} x_j + c_i,$$

man nennt die y_i auch *Schlupfvariablen*. Der zulässige Bereich (= Menge der *zulässigen Punkte*, d.h. derjenigen Punkte, die alle Nebenbedingungen erfüllen)

ist das Polyeder

$$P := \bigcap_{j=1}^n \{x_j \geq 0\} \cap \bigcap_{i=1}^m \{y_i \geq 0\}.$$

In diesem Abschnitt nehmen wir stets an, dass $0 \in P$ ist. Äquivalent dazu ist die Forderung, dass

$$c_i \geq 0 \quad \text{für alle } 1 \leq i \leq m \text{ gilt.}$$

Ähnlich wie im vorigen Abschnitt beschreiben wir das lineare Optimierungsproblem durch ein Ausgangsschema der Gestalt

	$-x_1$	\cdots	$-x_n$	
y_1	a_{11}	\cdots	a_{1n}	c_1
\vdots	\vdots		\vdots	\vdots
y_m	a_{m1}	\cdots	a_{mn}	c_m
f	$-b_1$	\cdots	$-b_n$	0

(S₀)

Definition 2.5.1

Ein Schema der Gestalt (S₀) heisst ein zum linearen Optimierungsproblem zugehöriges *Simplex-Tableau*.

Im Folgenden formen wir dieses Tableau durch Zeilen- und Spaltentausch in äquivalente Tableaus um der Gestalt

$$\begin{array}{c|ccc|c}
 & -u_1 & \cdots & -u_n & \\
 \hline
 v_1 & \alpha_{11} & \cdots & \alpha_{1n} & \gamma_1 \\
 \vdots & \vdots & & \vdots & \vdots \\
 v_m & \alpha_{m1} & \cdots & \alpha_{mn} & \gamma_m \\
 \hline
 f & \beta_1 & \cdots & \beta_n & \delta
 \end{array} \tag{S}$$

Hier sind dann $\alpha_{ij}, \beta_j, \gamma_i, \delta \in \mathbb{R}$ geeignete Koeffizienten und ähnlich wie zuvor gilt:

- (i) $u_1, \dots, u_n, v_1, \dots, v_m$ sind die (affin linearen Funktionen) $x_1, \dots, x_n, y_1, \dots, y_m$ in veränderter Reihenfolge,
- (ii) $v_i = -\sum_{j=1}^n \alpha_{ij} u_j + \gamma_i, \quad 1 \leq i \leq m,$
- (iii) $f = -\sum_{j=1}^n \beta_j u_j + \delta.$

Bemerkung 2.5.2

Für ein gegebenes Tableau (S) nennt man v_1, \dots, v_m auch die *Basisvariablen*, u_1, \dots, u_n die *Nichtbasisvariablen* und f die *Zielfunktion*. Das Schema (S) selbst nennt man auch eine *Basisform*.

Bei gegebener Basisform werden also die Basisvariablen und die Zielfunktion durch die Nichtbasisvariablen ausgedrückt. Der zulässige Bereich ist

$$P = \{u_1 \geq 0\} \cap \cdots \cap \{u_n \geq 0\} \cap \{v_1 \geq 0\} \cap \cdots \cap \{v_m \geq 0\}.$$

Definition 2.5.3

Eine Basisform (S) heisst *zulässig*, falls $\gamma_i \geq 0$ für alle $1 \leq i \leq m$.

Die Voraussetzung $0 \in P$ garantiert also, dass (S_0) eine zulässige Basisform ist.

Die folgenden Propositionen liefern einen ersten Simplexalgorithmus. Die Beweise kommen später.

Proposition 2.5.4

Sei (S) eine Basisform. Es gibt genau eine Lösung $x_S = ((x_S)_1, \dots, (x_S)_n) \in \mathbb{R}^n$ des linearen Gleichungssystems $u_j(x_S) = 0, 1 \leq j \leq n$, und diese Lösung x_S ist gegeben durch

$$(x_S)_I = \begin{cases} 0 & \text{falls } x_I \in \{u_1, \dots, u_n\} \text{ und} \\ \gamma_i & \text{falls } x_I = v_i. \end{cases} \quad (I)$$

Ist (S) eine zulässige Basisform, so ist x_S ein Extrempunkt von P . Ferner ist der Wert der Zielfunktion f in x_S gegeben durch

$$f(x_S) = \delta.$$

Bemerkung 2.5.5

Die Gleichungen $u_j(x) = 0$ sind entweder von der Form

$$x_l = 0 \quad (\text{falls } u_j = x_l)$$

oder

$$f_i = c_i \quad (\text{falls } u_j = y_i = -f_i + c_i).$$

In jedem Fall definieren die Gleichungen

$$u_j(x) = 0, \quad j = 1, \dots, n,$$

ein lineares Gleichungssystem im Sinne von Satz 2.3.1 mit linear unabhängigen Gleichungen. Für die eindeutige Lösung x_S gilt:

$$x_S \in P \text{ (und damit } x_S \in P_e) \Leftrightarrow (S) \text{ zulässig.}$$

Speziell für das Ausgangsschema (S_0) gilt

$$u_j = x_j, \quad \text{also } x_S = 0 \text{ und } f(x_S) = 0.$$

Man kann nun verschiedene Fälle unterscheiden.

Proposition 2.5.6

Sei (S) ein zulässiges Schema.

(i) Ist

$$\beta_j \geq 0 \quad \text{für alle } 1 \leq j \leq n,$$

so gilt

$$f(x_S) = \delta = \max f(P),$$

d.h. x_S ist Lösung des linearen Optimierungsproblems.

(ii) Gibt es ein $1 \leq l \leq n$, sodass $\beta_l < 0$ und

$$\alpha_{il} \leq 0 \quad \text{für alle } 1 \leq i \leq m,$$

so ist f auf P nicht nach oben beschränkt, d.h. $\sup f(P) = +\infty$.

(iii) Gibt es ein $1 \leq l \leq n$, sodass $\beta_l < 0$ und

existiert ein $1 \leq i \leq m$ mit $\alpha_{il} > 0$,

so gilt folgendes: Bildet man für alle solche $\alpha_{il} > 0$ die Quotienten $\frac{\gamma_i}{\alpha_{il}}$ und wählt man den Index k so, dass der Quotient minimal wird, also

$$\frac{\gamma_k}{\alpha_{kl}} = \min \left\{ \frac{\gamma_i}{\alpha_{il}} : i \text{ so, dass } \alpha_{il} > 0 \right\}, \quad (5)$$

dann liefert ein Austauschschritt mit Pivotelement α_{kl} ein neues zulässiges Schema (S'). Für den nach Proposition 2.5.4 dadurch bestimmten Extrempunkt $x_{S'}$ gilt

$$f(x_{S'}) \geq f(x_S).$$

Die Quotienten $\frac{\gamma_i}{\alpha_{il}}$ nennt man auch *charakteristische Quotienten*.

Durch Iteration erhält man folgenden *Simplexalgorithmus* für den Spezialfall, dass $x_1, \dots, x_n \geq 0$ und $0 \in P$ (wie vorausgesetzt):

- ① Stelle zulässiges Ausgangstableau auf
- ② Falls $\beta_j \geq 0$ für alle $1 \leq j \leq n$ (Maximum erreicht), dann setze Nichtbasisvariablen $u_j = 0$. Die Werte der x_j liefern eine Maximalstelle, und δ ist der Wert von f an dieser Maximalstelle. STOPP.
- ③ Falls für ein $\beta_l < 0$ gilt, dass $\alpha_{il} \leq 0$ für alle $1 \leq i \leq m$, so ist das Problem unbeschränkt, $\sup f(P) = +\infty$. STOPP.
- ④ Wähle $\beta_l < 0$ und wähle Zeile k , für die $\alpha_{kl} > 0$ ist und der charakteristische Quotient $\frac{\gamma_k}{\alpha_{kl}}$ minimal ist wie in (5).
- ⑤ Führe Austauschschritt aus mit Pivotelement α_{kl} .
- ⑥ Gehe zu 2.

Beispiel 2.5.7

Wir erinnern uns an Beispiel 2.1.1 (i), in dem es um die Frage ging, wieviel x_1 ha Kartoffeln und wieviel x_2 ha Getreide angebaut werden müssten, um die Zielfunktion

$$f(x_1, x_2) = 400 x_1 + 1200 x_2$$

unter den Restriktionen (Nebenbedingungen)

$$\begin{aligned}x_1, x_2 &\geq 0 && \text{(Fläche immer nichtnegativ)} \\100 x_1 + 200 x_2 &\leq 11000 && \text{(wegen Anbaukosten)} \\x_1 + 4 x_2 &\leq 160 && \text{(wegen Arbeitstagen)} \\x_1 + x_2 &\leq 100 && \text{(wegen max. verfügbarer Fläche)}\end{aligned}$$

zu maximieren.

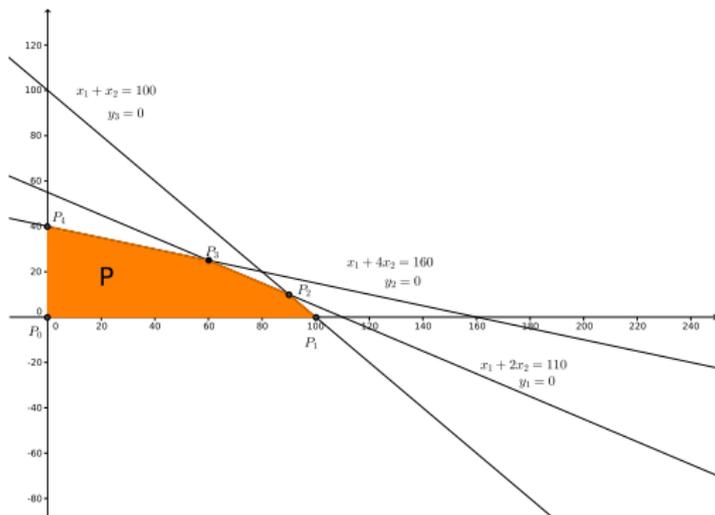


Abbildung: Simplexalgorithmus: Sukzessives Testen der Extrempunkte.

Wir starten mit dem Ausgangsschema (der Ausgangs-Basisform)

	$-x_1$	$-x_2$	
y_1	100	200	11000
y_2	1	4	160
y_3	1	1	100
f	-400	-1200	0

Hier sind also $n = 2$ und $m = 3$. Man hat $u_1(x) = x_1$ und $u_2(x) = x_2$, nach Proposition 2.5.4 ist $x_S = (0, 0) = P_0$ die eindeutige Lösung von $u_j(x_S) = 0$, $j = 1, 2$. Weil $\gamma_1 = 11000$, $\gamma_2 = 160$ und $\gamma_3 = 100$ alle nichtnegativ sind, ist das Schema zulässig, und daher P_0 ein Extrempunkt von P .

Nun hat man z.B. $\beta_2 = -b_2 = -1200 < 0$, also $\beta_l < 0$ für $l = 2$, somit ist nach Proposition 2.5.6 der Punkt P_0 möglicherweise noch keine Maximalstelle von $f|_P$. Da $\alpha_{22} = 4 > 0$, ist man im Fall wie in Proposition 2.5.6 (iii) beschrieben.

Der charakteristische Quotient $\frac{\gamma_2}{\alpha_{22}} = \frac{160}{4} = 40$ ist kleiner als die anderen beiden, nämlich $\frac{\gamma_1}{\alpha_{12}} = \frac{11000}{200} = 55$ und $\frac{\gamma_3}{\alpha_{32}} = \frac{100}{1} = 100$. Daher machen wir einen Austauschschritt mit Pivotelement $\alpha_{22} = 4$.

Wir tauschen also die Bezeichnungen x_2 und y_2 aus, folgen den Pivotregeln und erhalten folgendes neues Schema:

	$-x_1$	$-y_2$	
y_1	50	-50	3000
x_2	$\frac{1}{4}$	$\frac{1}{4}$	40
y_3	$\frac{3}{4}$	$-\frac{1}{4}$	60
f	-100	300	48000

Hier sind $u_1(x) = x_1$ und $u_2(x) = y_2$, also ist die Lösung von $u_j(x_S) = 0$, $j = 1, 2$, diesmal $x_S = (0, 40) = P_4$, denn $x_2 = -\frac{1}{4}x_1 - \frac{1}{4}y_2 + 40 = 40$. Wiederum ist das Schema zulässig, also P_4 extremal. Wir 'testen', ob P_4 Maximalstelle ist: da $\beta_1 = -100 < 0$, ist auch P_4 möglicherweise keine Maximalstelle. Es gibt für $l = 1$ Elemente α_{kl} , die grösser Null sind (sogar alle in der ersten Spalte), und die charakteristischen Quotienten für $l = 1$ sind $\frac{\gamma_1}{\alpha_{11}} = \frac{3000}{50} = 60$, $\frac{\gamma_2}{\alpha_{21}} = \frac{40}{1/4} = 160$ und $\frac{\gamma_3}{\alpha_{31}} = \frac{60}{3/4} = 80$, davon ist der erste minimal. Wir machen wieder einen Austauschschritt, diesmal mit Pivot $\alpha_{11} = 50$.

Tausch der Bezeichnungen x_1 und y_1 liefert nach Pivotregeln das neue Schema

	$-y_1$	$-y_2$	
x_1	#	#	60
x_2	#	#	25
y_3	#	#	15
f	2	200	54000

Jetzt sind $u_j(x) = y_j$, $j = 1, 2$, und die eindeutige Lösung von $u_j(x_S) = 0$ kann man ablesen, nämlich $x_S = (60, 25) = P_3$. Da das Schema zulässig ist, ist P_3 extremal. Wir haben hier nun zuerst die letzte Zeile berechnet, und sehen schon, dass $\beta_1 = 2$ und $\beta_2 = 200$ beide nichtnegativ sind. Nach Proposition 2.5.6 (i) ist dann P_3 garantiert eine Maximalstelle für $f|_P$, und $f(P_3) = f(x_S) = \delta = 54000$. Die Stellen α_{kl} im neuen Schema müssen nun gar nicht mehr berechnet werden.

Bemerkung 2.5.8 (nach Beweis 2.5.4)

Proposition 2.5.4 sagt also, dass in jedem Simplextableau (S) die Nichtbasisvariablen ein System von n linear unabhängigen Linearformen $u_j - u_j(0)$, $1 \leq j \leq n$, definieren, die ein lineares Gleichungssystem $u_j(x) - u_j(0) = u_j(0)$, $1 < j < n$, wie in Satz 2.3.1 (i) bilden.

Bemerkung 2.5.9

Bei der Wahl des Pivotelementes in Schritt 4. des Simplexalgorithmus und dem anschliessenden Austauschschritt sind nun zwei Fälle (a) und (b) zu unterscheiden:

(a) In allen Zeilen mit $\alpha_{ij} > 0$ gilt $\gamma_i > 0$. Dann gilt

$$f(x_{S'}) = \delta - \underbrace{\frac{\beta_l \gamma_k}{\alpha_{kl}}}_{<0} > \delta = f(x_S),$$

d.h. das neue Schema beschreibt einen neuen Extrempunkt $x_{S'} \neq x_S$. Der Wert der Zielfunktion nimmt dabei echt zu. Da die Gleichungssysteme, die x_S und $x_{S'}$ beschreiben, sich nur in einer Gleichung unterscheiden (denn $n - 1$ der Nichtbasisvariablen sind gleich), sind x_S und $x_{S'}$ benachbart. Also entspricht in diesem Falle ein Austauschschritt geometrisch einem Schritt zu einem benachbarten Extrempunkt, für welchen dann der Wert der Zielfunktion echt grösser ist.

- (b) Unter allen Zeilen mit $\alpha_{ij} > 0$ gibt es eine mit $\gamma_i = 0$. Eine solche Zeile mit $\gamma_i = 0$ ist dann als Pivotzeile zu wählen, weil dann der charakteristische Quotient

$$\frac{\gamma_i}{\alpha_{ij}} = 0$$

ist und somit minimal sein muss. Es gilt dann also $\alpha_{kl} > 0$ und $\gamma_k = 0$. Das bedeutet aber, dass die letzte Spalte beim Austausch unverändert bleibt, d.h.

$$\gamma'_i = \gamma_i, \quad 1 \leq i \leq m, \quad \text{und} \quad f(x_{S'}) = \delta' = \delta = f(x_S).$$

Es gilt sogar $x_{S'} = x_S$, denn weil der Austauschschritt dann der Gestalt

$$\begin{array}{c|c|c} & -u_l & \\ \hline v_k & & 0 \end{array} \quad \rightarrow \quad \begin{array}{c|c|c} & -v_k & \\ \hline u_l & & 0 \end{array}$$

ist, führen beide Schemata bei Nullsetzen der Nichtbasisvariablen u_1, \dots, u_n (im ursprünglichen Schema) bzw. $u_1, \dots, u_{k-1}, v_k, u_{k+1}, \dots, u_n$ (im neuen Schema) auf die Gleichungen

$$u_j(x_S) = u_j(x_{S'}) = 0, \quad 1 \leq j \leq n \quad \text{und} \quad v_k(x_S) = v_k(x_{S'}) = 0.$$

Insbesondere lösen x_S und $x_{S'}$ beide das homogene System der linear unabhängigen Gleichungen $u_1, \dots, u_{k-1}, v_k, u_{k+1}, \dots, u_n$. Also folgt $x_{S'} = x_S$.

Definition 2.5.10

Ein Austausch mit Pivotelement α_{kl} , für welches $\gamma_k = 0$ ist, heisst *stationärer Austausch*.

Bemerkung 2.5.11

- (i) Ein stationärer Austausch beschreibt also den Übergang zwischen zwei Gleichungssystemen, die im Sinne von Satz 2.3.1 denselben Extrempunkt beschreiben.
- (ii) Es können dabei zwei Situationen auftreten:
 - (b.1) Der Punkt x_S besitzt einen benachbarten Extrempunkt \tilde{x} mit $f(x_S) < f(\tilde{x})$. Dann gibt es nach Satz 2.3.7 eine endliche Folge stationärer Austauschschritte, nach welchen wieder der Fall (a) eintritt.
 - (b.2) Der Punkt x_S besitzt keinen benachbarten Extrempunkt \tilde{x} mit $f(x_S) < f(\tilde{x})$. In diesem Falle ist x_S bereits ein Extrempunkt, an welchem die Zielfunktion f maximal wird, oder f ist unbeschränkt auf P .
- (iii) Eine Folge stationärer Austauschschritte kann einen Zyklus (loop) bilden. (Dann würde sich der Algorithmus aufhängen.)

Um das zu vermeiden, gibt es geeignete Auswahlregeln, z.B. *Bland's Antizyklusregel*: Ordne die Variablen in einer Reihenfolge, z.B.

$$z_1 = x_1, \dots, z_n = x_n, z_{n+1} = y_1, \dots, z_{n+m} = y_m.$$

Wenn bei einem Austausch mehrere Spalten wählbar sind, so wähle stets die Variable z_i mit dem kleinsten Index i . Sind mehrere Zeilen wählbar, so wähle auch hier den kleinsten Index. (Hochstätter Satz 5.30)

Wir ergänzen Schritt 4. im Simplexalgorithmus durch folgende *Auswahlregel*

- (i) Versuche stationäre Austauschschritte zu vermeiden: Gibt es mehrere $\beta_l < 0$ (mehrere mögliche Pivotspalten), so wähle wenn möglich eine solche, für welche der Fall (a) eintritt. Falls nicht möglich, wende Blands Antizyklusregel an.
- (ii) (*Regel von Dantzig*) Sind mehrere Pivotspalten für einen nicht stationären Austausch möglich, so wähle die Spalte mit dem betragsmässig grössten β_l (Faustregel, Grundidee ist maximale Kostenreduktion; aber nicht unbedingt optimal).

Korollar 2.5.12

Mit der zusätzlichen Auswahlregel terminiert der Simplex-Algorithmus nach endlich vielen Austauschschritten.

Bemerkung 2.5.13

Selbst wenn der zulässige Bereich ein unbeschränktes Polyeder P ist, gilt: Ist die Zielfunktion f auf P beschränkt, so nimmt $f|_P$ sein Maximum in einem Extrempunkt von P an.

Bemerkung 2.5.14

Man kann die Effekte der Auswahlregel auch noch stärker formalisiert beschreiben, allerdings ist das eher notationsintensiv, siehe zum Beispiel Jungnickel Kapitel 4.3.

Variablen nicht nach unten beschränkt

Wir möchten $f(x) = b_1 x_1 + \dots + b_n x_n$ maximieren unter den Nebenbedingungen

$$f_i(x) = \sum_{j=1}^n a_{ij}x_j \leq c_i, \quad 1 \leq i \leq m,$$

wobei wie zuvor $c_i \geq 0$ gelten soll für alle i . (Also soll 0 wieder ein Punkt des zulässigen Bereiches sein.) Wir verlangen aber nun *nicht* mehr, dass $x_j \geq 0$ gelten soll für alle $1 \leq j \leq n$.

Für jedes j , für das $x_j \geq 0$ keine vorgegebene Restriktion mehr ist, kann man nun wie folgt vorgehen: Man ersetzt x_j einfach durch die Differenz zweier neuer Variablen x'_j und x''_j , also

$$x_j = x'_j - x''_j,$$

und stellt die neuen Nebenbedingungen

$$x'_j \geq 0 \quad \text{und} \quad x''_j \geq 0.$$

(Geometrisch erhöht man also künstlich die Dimension.)

Beispiel 2.6.1

Nehmen wir an, $f(x) = x_1 - x_2$, $x = (x_1, x_2) \in \mathbb{R}^2$, soll maximiert werden unter den Nebenbedingungen

$$3x_1 - x_2 \leq 6$$

$$x_1 - 2x_2 \leq 4$$

$$x_2 \leq 1$$

$$x_1 \geq 0.$$

Wir setzen $x_2 := x_2' - x_2''$. Das führt auf das modifizierte Problem,

$$f(x) = x_1 - x_2' + x_2'', \quad x = (x_1, x_2', x_2'') \in \mathbb{R}^3,$$

zu maximieren unter

$$3x_1 - x_2' + x_2'' \leq 6$$

$$x_1 - 2x_2' + 2x_2'' \leq 4$$

$$x_2' - x_2'' \leq 1$$

sowie $x_1 \geq 0$, $x_2' \geq 0$ und $x_2'' \geq 0$. Das führt auf das Ausgangsschema

	$-x_1$	$-x_2'$	$-x_2''$	
y_1	3	-1	1	6
y_2	1	-2	2	4
y_3	0	1	-1	1
f	-1	1	-1	0

Man kann nun wie im letzten Abschnitt mit dem Simplexverfahren lösen, und am Ende erhält man x_2 aus $x_2 = x_2' - x_2''$:

Der erste Austauschschritt mit Pivot 2 ergibt das Schema

	$-x_1$	$-x_2'$	$-y_2$	
y_1	$\frac{5}{2}$	0	$-\frac{1}{2}$	4
x_2''	$\frac{1}{2}$	-1	$\frac{1}{2}$	2
y_3	$\frac{1}{2}$	0	$\frac{1}{2}$	3
f	$-\frac{1}{2}$	0	$\frac{1}{2}$	2

Der zweite, mit Pivot $\frac{5}{2}$, ergibt

	$-y_1$	$-x_2'$	$-y_2$	
x_1	$\frac{2}{5}$	0	$-\frac{1}{5}$	$\frac{8}{5}$
x_2''	$-\frac{1}{5}$	-1	$\frac{6}{10}$	$\frac{6}{5}$
y_3	$-\frac{1}{5}$	0	$\frac{6}{10}$	$\frac{11}{5}$
f	$\frac{1}{5}$	0	$\frac{4}{10}$	$\frac{14}{5}$

Nun ist man aber schon in der Situation von Proposition 2.5.6 (i), denn die Einträge β_j in der letzten Zeile sind alle nichtnegativ. Die eindeutige Lösung des Gleichungssystems, welches aus diesem Schema durch Nullsetzen der Nichtbasisvariablen entsteht, ist

$$(x_1, x_2', x_2'') = \left(\frac{8}{5}, 0, \frac{6}{5}\right),$$

und f hat in diesem Punkt den Wert $\frac{14}{5}$. Rückübersetzen ergibt

$$x_2 = x_2' - x_2'' = -\frac{6}{5},$$

also nimmt die Zielfunktion $f|_P$ ihr Maximum in $(x_1, x_2) = \left(\frac{8}{5}, -\frac{6}{5}\right)$ an,

$$\max f(P) = \frac{14}{5} = f\left(\frac{8}{5}, -\frac{6}{5}\right).$$

Ausgangsschema nicht zulässig

Wir nehmen an, dass es ein $i \in \{1, \dots, m\}$ gibt mit $c_i < 0$. Dann ist also $0 \notin P$. Das Simplexverfahren in der besprochenen Form benötigt aber ein zulässiges Ausgangsschema. Wir ermitteln daher ein solches. Nehmen wir an, die gegebenen Restriktionen lauten

$$\sum_{j=1}^n a_{ij}x_j \leq c_i, \quad 1 \leq i \leq m,$$

und $x_1, \dots, x_n \geq 0$. Der zulässige Bereich ist dann das Polyeder

$$P = \left\{ x = (x_1, \dots, x_n) \in \mathbb{R}_+^n : \sum_{j=1}^n a_{ij}x_j \leq c_i, \quad i = 1, \dots, m \right\}.$$

Wir betrachten nun ein Hilfsproblem mit $n + 1$ Variablen x_1, \dots, x_n, t und Zielfunktion

$$\tilde{f}(x, t) = -t$$

unter den Nebenbedingungen

$$\sum_{j=1}^n a_{ij}x_j - t \leq c_i, \quad 1 \leq i \leq m,$$

und $x_1, \dots, x_n, t \geq 0$. Der zugehörige zulässige Bereich ist das Polyeder

$$\tilde{P} = \left\{ (x, t) \in \mathbb{R}_+^{n+1} : \sum_{j=1}^n a_{ij}x_j - t \leq c_i, \quad i = 1, \dots, m \right\}.$$

Dann gelten:

- (i) Für $c := -\min\{c_i : i = 1, \dots, m\}$ ist $(0, \dots, 0, c) \in \tilde{P}$, d.h. $\tilde{P} \neq \emptyset$ und, da \tilde{f} nach oben beschränkt ist durch 0, ist das Hilfsproblem lösbar.
- (ii) Man hat $x \in P$ genau dann, wenn $(x, 0) \in \tilde{P}$.
- (iii) Man hat $P \neq \emptyset$ genau dann, wenn $\max \tilde{f}(\tilde{P}) = 0$, denn:

In Tableauschreibweise liefert das den folgenden Algorithmus: Das Ausgangsschema ist von der Form

	$-x_1$	\cdots	$-x_n$	
y_1	a_{11}	\cdots	a_{1n}	c_1
\vdots	\vdots		\vdots	\vdots
y_m	a_{m1}	\cdots	a_{mn}	c_m
f	$-b_1$	\cdots	$-b_n$	0

und für mindestens ein i ist $c_i < 0$. Für das Hilfsproblem haben wir das Schema

	$-x_1$	\cdots	$-x_n$	$-t$	
y_1	a_{11}	\cdots	a_{1n}	-1	c_1
\vdots	\vdots			\vdots	\vdots
y_k				-1	
\vdots	\vdots			\vdots	\vdots
y_m	a_{m1}	\cdots	a_{mn}	-1	c_m
\tilde{f}	0	\cdots	0	1	0

und das ist ebenfalls nicht zulässig. Nun wählen wir $k \in \{1, \dots, m\}$ so, dass

$$c_k = \min\{c_i : i = 1, \dots, m\}$$

und tauschen y_k und t mit Pivot -1 , das ergibt ein Schema der Gestalt

	$-x_1$	\cdots	$-x_n$	$-y_k$	
y_1	#	\cdots	#	-1	γ_1
\vdots	\vdots			\vdots	\vdots
t				-1	
\vdots	\vdots			\vdots	\vdots
y_m	#	\cdots	#	-1	γ_m
\tilde{f}	#	\cdots	#	1	#

mit $\gamma_k = -c_k > 0$, und für $j \neq k$ gilt

$$\gamma_j = c_j + c_k \cdot (-1) = c_j - c_k \geq 0.$$

Also ist dieses Schema zulässig.

- ① maximiere $\tilde{f}(x, t) = -t$ auf \tilde{P}
- ② $\max \tilde{f}(\tilde{P}) < 0 \Rightarrow P = \emptyset$
- ③ $\max \tilde{f}(\tilde{P}) = 0 \Rightarrow P \neq \emptyset \Rightarrow$ Maximum in $(x_0, 0) \in \mathbb{R}_+^{n+1} \Rightarrow x_0 \in P$

Nun können wir erreichen, dass t eine Nichtbasisvariable ist, denn ist t Basisvariable, so muss in

	$-u_1$	\cdots	$-u_{n+1}$	
	#	\cdots	#	γ_1
\vdots	\vdots		\vdots	\vdots
t	α_{k1}		α_{km}	γ_k
\vdots	\vdots		\vdots	\vdots
	#	\cdots	#	γ_m
\tilde{f}	#	\cdots	#	0

$\gamma_k = 0$ gelten, weil die Lösung des Systems $(x, t) = (x_0, 0)$ ist.

- 1 t auf \tilde{P} unbeschränkt \Rightarrow es gibt $j \in \{1, \dots, n+1\}$ mit $\alpha_{kj} \neq 0$
- 2 Austausch mit solchem $\alpha_{kj} \neq 0$ als Pivotelement ist stationär, ergibt also zulässiges Schema mit t als Nichtbasisvariable.
- 3 Dieser stationäre Austausch ändert nicht den betrachteten Extrempunkt von \tilde{P} .
- 4 Setze $t = 0$, also streiche die t -Spalte und erhalte ein zulässiges Schema, das die Restriktionen von P beschreibt.
- 5 Ersetze die letzte Zeile durch die korrekte Zeile für die ursprüngliche Zielfunktion f , dargestellt als Funktion der aktuellen Nichtbasisvariablen. Dazu drücke jedes x_k , das eine Basisvariable ist, durch die Nichtbasisvariablen aus:

$$x_k = \sum_{j=1}^n \alpha_{ij} u_j + \gamma_i, \quad \text{falls } x_k \text{ in der } i\text{-ten Zeile steht.}$$

Bemerkung 2.6.2

Man löst also zuerst das Hilfsproblem und nutzt die Lösung, um ein zulässiges Ausgangsschema für das ursprüngliche Problem zu konstruieren. Danach löst man das ursprüngliche Problem wie gewohnt.

Beispiel 2.6.3

Wir wollen $f(x) = 2x_1 - x_2 + 2x_3$ maximieren unter den Restriktionen $x_1, x_2, x_3 \geq 0$ und

$$\begin{aligned}x_1 + x_2 + x_3 &\leq 6 \\ -x_1 + x_2 &\leq -1 \\ -x_2 + x_3 &\leq -1.\end{aligned}$$

Das liefert ein nicht zulässiges Ausgangsschema. Wir maximieren deshalb $\tilde{f}(x, t) = -t$ unter den Restriktionen $x_1, x_2, x_3, t \geq 0$ und

$$\begin{aligned}x_1 + x_2 + x_3 - t &\leq 6 \\ -x_1 + x_2 - t &\leq -1 \\ -x_2 + x_3 - t &\leq -1.\end{aligned}$$

Das zugehörige Schema ist

	$-x_1$	$-x_2$	$-x_3$	$-t$	
y_1	1	1	1	-1	6
y_2	-1	1	0	-1	-1
y_3	0	-1	1	-1	-1
\tilde{f}	0	0	0	1	0

Der minimale Wert in der letzten Spalte ist -1 , er wird in der 2. und der 3. Zeile angenommen. Wir tauschen t mit y_2 , (Pivot ist -1) und erhalten

	$-x_1$	$-x_2$	$-x_3$	$-y_2$	
y_1	2	0	1	-1	7
t	1	-1	0	-1	1
y_3	1	-2	1	-1	0
\tilde{f}	-1	1	0	1	-1

Dieses Schema ist nun zulässig. Die charakteristischen Quotienten für $l = 1$ sind $\frac{7}{2}$, 1 und 0 (stationärer Austausch). Mit Pivot 1 ergibt sich nun

	$-y_3$	$-x_2$	$-x_3$	$-y_2$	
y_1	-2	4	-1	1	7
t	-1	1	-1	0	1
x_1	1	-2	1	-1	0
\tilde{f}	1	-1	1	0	-1

und die charakteristischen Quotienten zu $l = 2$ sind $\frac{7}{4}$ und 1, letzterer minimal.
Mit Pivot 1 erhalten wir

	$-y_3$	$-t$	$-x_3$	$-y_2$	
y_1	2	-4	3	1	3
x_2	-1	1	-1	0	1
x_1	-1	2	-1	-1	2
\tilde{f}	0	1	0	0	0

Dieses Schema ist optimal im dem Sinne, dass \tilde{f} ihren maximalen Wert 0 annimmt, also ist insbesondere $P \neq \emptyset$.

Wir dürfen die t -Spalte 'vergessen' und $t = 0$ setzen. Stellen nun f als Funktion von y_3 , x_3 und y_2 dar: wir hatten $f(x) = 2x_1 - x_2 + 2x_3$. Aus 2. und 3. Zeile folgt, dass (wegen $t = 0$)

$$x_2 = y_3 + x_3 + 1$$

$$x_1 = y_3 + x_3 + y_2 + 2,$$

also

$$f(x) = y_3 + 3x_3 + 2y_2 + 3.$$

Wir haben nun folgendes zulässiges Schema für das ursprüngliche Problem gefunden:

	$-y_3$	$-x_3$	$-y_2$	
y_1	2	3	1	3
x_2	-1	-1	0	1
x_1	-1	-1	-1	2
f	-1	-3	-2	3

Folgen nun wie früher dem Simplexalgorithmus. Für $l = 3$ hat man nur den charakteristischen Quotienten 3, wählen also Pivot **1** und erhalten

	$-y_3$	$-x_3$	$-y_1$	
y_2	2	3	1	3
x_2	-1	-1	0	1
x_1	1	2	1	5
f	3	3	2	9

Da alle β_l nichtnegativ sind, ist das Schema optimal, und Nullsetzen der Nichtbasisvariablen ergibt die Maximalstelle $(5, 1, 0)$ für $f|_P$, und der Wert von f an dieser Stelle ist 9.

Alternativsätze: Notation und Wiederholung

Wir fassen Vektoren als Spaltenvektoren auf,

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

Wir schreiben

$$x \geq 0 \quad \text{falls} \quad x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0$$

und $x \geq y$ oder $y \leq x$, falls $x - y \geq 0$. Für eine Matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad \text{bezeichnet} \quad A^T = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix}$$

ihre transponierte,

insbesondere

$$(x_1, \dots, x_n)^T = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

- Rechenregel: $(A \cdot B)^T = B^T \cdot A^T$
- Euklidisches Skalarprodukt:

$$\langle x, y \rangle := \sum_{i=1}^n x_i y_i = \langle x, y \rangle$$

- $L \subset \mathbb{R}^n$ Untervektorraum von \mathbb{R}^n , dann heisst

$$L^\perp := \{y \in \mathbb{R}^n : \langle x, y \rangle = 0 \text{ für alle } x \in L\}$$

das *orthogonale Komplement* von L

- $L \oplus L^\perp = \mathbb{R}^n$ (orthogonale Summe) und $(L^\perp)^\perp = L$.

Alternativsätze: Sätze und Beweise

Satz 2.7.1 (Lemma von Farkas)

Sei $L \subset \mathbb{R}^n$ ein Untervektorraum. Dann gilt genau eine der beiden Alternativen:

- (i) Es gibt ein $x \in L$ mit $x \geq 0$ und $x_1 > 0$.
- (ii) Es gibt ein $y \in L^\perp$ mit $y \geq 0$ und $y_1 > 0$.

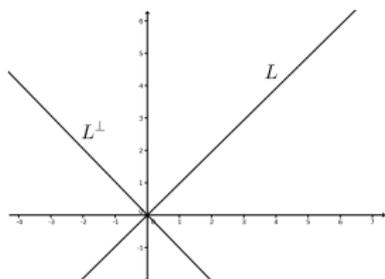


Abbildung: Fall (i)

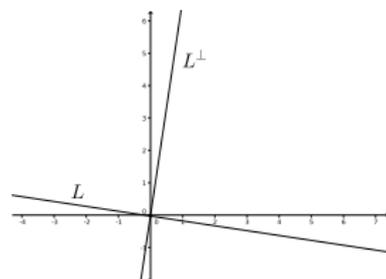


Abbildung: Fall (ii)

Zum Beweis des Lemmas von Farkas verwenden wir folgende Aussage.

Lemma 2.7.2

Sei $n \geq 2$, $L \subset \mathbb{R}^n$ ein Untervektorraum und L^\perp sein orthogonales Komplement.
Seien

$$\tilde{L} := \{\tilde{x} \in \mathbb{R}^{n-1} : (\tilde{x}, 0) \in L\} = L \cap \{x_n = 0\}$$

('Schnitt von L mit der Hyperebene $H := \{x_n = 0\}$ ') und

$$\hat{L}^\perp := \{\hat{y} \in \mathbb{R}^{n-1} : \text{es gibt ein } y_n \in \mathbb{R} \text{ mit } (\hat{y}, y_n) \in L^\perp\}$$

('Bild der Projektion von L^\perp auf H '). Dann sind \tilde{L} und \hat{L}^\perp zueinander orthogonal komplementäre Unterräume des \mathbb{R}^{n-1} , das bedeutet:

$$(\tilde{L})^\perp = \hat{L}^\perp.$$

Bemerkung 2.7.3

Sei A eine reelle $(m \times n)$ -Matrix. Dann gilt ja

$$\langle y, Ax \rangle = y^T (Ax) = (y^T A)x = (A^T y)^T x = \langle A^T y, x \rangle.$$

Daher gilt für Kern

$$\ker(A) = \{x \in \mathbb{R}^n : Ax = 0\}$$

und Bild

$$\text{Bild}(A^T) = \{A^T y : y \in \mathbb{R}^m\}$$

die Gleichung

$$\ker(A) = (\text{Bild}(A^T))^\perp,$$

denn für beliebiges $x \in \mathbb{R}^n$ gilt:

$$\begin{aligned} x \in \ker(A) &\Leftrightarrow Ax = 0 \Leftrightarrow \langle y, Ax \rangle = 0 \text{ für alle } y \in \mathbb{R}^m \\ &\Leftrightarrow \langle A^T y, x \rangle = 0 \text{ für alle } y \in \mathbb{R}^m \Leftrightarrow x \in (\text{Bild}(A^T))^\perp. \end{aligned}$$

Bemerkung 2.7.4

Bemerkung 2.7.3 besagt: A ist nicht injektiv, genau dann wenn A^T nicht surjektiv ist. Also ist A^T surjektiv die Alternative zu A nicht injektiv.

Satz 2.7.5 (Minkowski-Farkas für lineare Gleichungen)

Sei A eine reelle $(m \times n)$ -Matrix und $b \in \mathbb{R}^m$. Dann sind äquivalent:

- (i) Die Gleichung $Ax = b$ besitzt eine Lösung $x \geq 0$.
- (ii) Für alle $u \in \mathbb{R}^m$ mit $u^T A \leq 0$ gilt $u^T b \leq 0$.

Satz 2.7.6 (Minkowski-Farkas für lineare Ungleichungen)

Sei A eine reelle $(m \times n)$ -Matrix und $b \in \mathbb{R}^m$. Dann sind äquivalent:

- (i) Die Ungleichung $Ax \geq b$ besitzt eine Lösung $x \geq 0$.
- (ii) Für alle $u \in \mathbb{R}^m$ mit $u \geq 0$ und $u^T A \leq 0$ gilt $u^T b \leq 0$.

Formulierung und Beweis des Dualitätssatzes

Gegeben sei ein lineares Optimierungsproblem wie zuvor, in dem eine lineare Zielfunktion

$$f(x) = b_1x_1 + \dots + b_nx_n$$

zu maximieren ist unter den Restriktionen $x_1 \geq 0, \dots, x_n \geq 0$ und

$$f_i(x) = a_{i1}x_1 + \dots + a_{in}x_n \leq c_i, \quad 1 \leq i \leq m,$$

in Kurzschreibweise:

$$\text{Maximiere } b^T x \text{ unter den Bedingungen } x \geq 0 \text{ und } Ax \leq c, \quad (6)$$

wobei

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad c = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$$

gegeben sind und $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$. Im Kontext von Dualitätsdiskussionen nennen wir (6) auch das *primale Problem*.

Definition 2.8.1

Das zu (6) *duale Problem* lautet:

Minimiere $c^T y$ unter den Bedingungen $y \geq 0$ und $A^T y \geq b$, $y \in \mathbb{R}^m$.

Bemerkung 2.8.2

- (i) Durch Übergang zu $-c$, $-A$, $-b$ kann man auch das duale Problem wieder wie gewohnt schreiben, nämlich:

Maximiere $(-c)^T y$ unter den Bedingungen $y \geq 0$ und $(-A)^T y \leq -b$.

- (ii) Das zum dualen Problem duale Problem ist wieder das primale Problem.

Im Folgenden schreiben wir

$$P := \{x \in \mathbb{R}^n : x \geq 0, Ax \leq c\}$$

und

$$\tilde{P} := \{y \in \mathbb{R}^m : y \geq 0, A^T y \geq b\}$$

für die zulässigen Bereiche des primalen und des dualen Problems.

Wir beobachten folgenden *schwachen Dualitätssatz*:

Lemma 2.8.3

Für alle $x \in P$ und $y \in \tilde{P}$ gilt

$$b^T x \leq c^T y. \quad (7)$$

Bemerkung 2.8.4

- Ist $\tilde{P} \neq \emptyset$, so ist die Zielfunktion $b^T x$ des primalen Problems nach oben beschränkt.
- Ist $P \neq \emptyset$, so ist die Zielfunktion $c^T y$ des dualen Problems nach unten beschränkt.

Satz 2.8.5

Gibt es $x_0 \in P$ und $y_0 \in \tilde{P}$ mit $b^T x_0 \geq c^T y_0$, so folgt

$$b^T x_0 = \max_{x \in P} b^T x = \min_{y \in \tilde{P}} c^T y = c^T y_0.$$

Satz 2.8.6

Folgende Aussagen sind äquivalent:

(i) Das primale Problem besitzt eine Lösung $x_0 \in P$,

$$b^T x_0 = \max_{x \in P} b^T x.$$

(ii) Das duale Problem besitzt eine Lösung $y_0 \in \tilde{P}$,

$$c^T y_0 = \min_{y \in \tilde{P}} c^T y.$$

(iii) Es gilt $P \neq \emptyset$ und $\tilde{P} \neq \emptyset$.

Falls eine (und damit alle) dieser Aussagen gelten, so folgt, dass

$$b^T x_0 = c^T y_0. \quad (8)$$

Bemerkung 2.8.7

Nach Satz 2.8.6 können also vier Fälle auftreten:

- 1 Man hat $P = \emptyset$ und $\tilde{P} = \emptyset$.
- 2 Man hat $P \neq \emptyset$ und $\tilde{P} = \emptyset$. Dann ist das primäre Problem unbeschränkt.
- 3 Man hat $P = \emptyset$ und $\tilde{P} \neq \emptyset$. Dann ist das duale Problem unbeschränkt.
- 4 Man hat $P \neq \emptyset$ und $\tilde{P} \neq \emptyset$. Dann sind beide Probleme lösbar (wie im Satz).

Insbesondere gilt also:

Ist $P \neq \emptyset$, so ist das primale Problem genau dann lösbar, wenn $\tilde{P} \neq \emptyset$.

und analog

Ist $\tilde{P} \neq \emptyset$, so ist das duale Problem genau dann lösbar, wenn $P \neq \emptyset$.

Übertragung auf den Simplexalgorithmus

- Primales Problem: maximiere $b^T x$ unter den Bedingungen $x \geq 0$ und $Ax \leq c$.
- Duales Problem: minimiere $c^T y$ unter den Bedingungen $y \geq 0$ und $A^T y \geq b$.
- Modifiziertes Ausgangstableau für beide Probleme:

x_1	\cdots	x_n	g	
a_{11}	\cdots	a_{1n}	c_1	y_1
\vdots		\vdots	\vdots	\vdots
a_{m1}	\cdots	a_{mn}	c_m	y_m
$-b_1$	\cdots	$-b_n$	0	

(\tilde{S}_0)

- *Duale Interpretation*: fasse y_1, \dots, y_m als Linearformen auf \mathbb{R}^m auf, also $y_i : \mathbb{R}^m \rightarrow \mathbb{R}$, $y_i(y) = y_i$, und definiere affin lineare Funktionen

$$x_j(y) := \sum_{i=1}^m a_{ij} y_i - b_j, \quad 1 \leq j \leq n, \quad \text{und} \quad g(y) := \sum_{i=1}^m c_i y_i.$$

- Für den zulässigen Bereich gilt dann

$$\tilde{P} = \underbrace{\{x_1 \geq 0, \dots, x_n \geq 0\}}_{=\{A^T y \geq b\}} \cap \{y_1 \geq 0, \dots, y_m \geq 0\}.$$

- Austauschritte nach Pivotregeln ergeben ein neues Tableau

u_1	\cdots	u_n	g	
α_{11}	\cdots	α_{1n}	γ_1	v_1
\vdots		\vdots	\vdots	\vdots
α_{m1}	\cdots	α_{mn}	γ_m	v_m
β_1	\cdots	β_n	δ	

 (\tilde{S})

Proposition 2.8.8

Die Gleichungen

$$u_j = \sum_{i=1}^m \alpha_{ij} v_i + \beta_j, \quad 1 \leq j \leq n, \quad \text{und} \quad g = \sum_{i=1}^m \gamma_i v_i + \delta$$

des Schemas (\tilde{S}) sind äquivalent zu den Gleichungen des Schemas (\tilde{S}_0) . Insbesondere gilt

$$\tilde{P} = \{u_1 \geq 0, \dots, u_n \geq 0\} \cap \{v_1 \geq 0, \dots, v_m \geq 0\}.$$

Korollar 2.8.9

Startet man mit einem zulässigen Ausgangsschema für das primale Problem und findet man mittels Simplexalgorithmus das dafür optimale Schema (S) , dann gilt:

- (i) Man hat $\gamma_1 \geq 0, \dots, \gamma_m \geq 0$, denn (S) ist ein zulässiges Schema.
- (ii) Man hat $\beta_1 \geq 0, \dots, \beta_n \geq 0$, denn das Schema ist optimal.
- (iii) Auch für (S) gilt die duale Interpretation.

Satz 2.8.10

Findet man mit dem Simplexalgorithmus ein optimales zulässiges Tableau, so bestimmt dies simultan Lösungen des primalen und des dualen Problems. Für das duale Problem gilt: Der minimale Wert der Zielfunktion ist δ , und dieser wird angenommen im Punkt $y_0 \in \mathbb{R}^m$ mit

$$(y_0)_k = \begin{cases} 0 & \text{falls } y_k = v_i \text{ für ein } i \\ \beta_j & \text{falls } y_k = u_j \text{ für ein } j. \end{cases}$$

Beispiel 2.8.11

Minimiere $y_1 + y_2$, $y = (y_1, y_2) \in \mathbb{R}^2$ unter den Bedingungen $y_1 \geq 0$, $y_2 \geq 0$ und

$$4 y_1 + y_2 \geq 2$$

$$3 y_1 + 2 y_2 \geq 4$$

$$y_1 + 8 y_2 \geq 10$$

Umformen in ein primales Problem ergibt: maximiere $-y_1 - y_2$ unter den Bedingungen $y_1 \geq 0$, $y_2 \geq 0$ und

$$-4 y_1 - y_2 \leq -2$$

$$-3 y_1 - 2 y_2 \leq -4$$

$$-y_1 - 8 y_2 \leq -10.$$

Das zugehörige Ausgangsschema ist nicht zulässig.

Aber dieses Problem ist das dual zu: maximiere $2 x_1 + 4 x_2 + 10 x_3$,
 $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ unter den Bedingungen $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$ und

$$4 x_1 + 3 x_2 + x_3 \leq 1$$

$$x_1 + 2 x_2 + 8 x_3 \leq 1.$$

	$-x_1$	$-x_2$	$-x_3$	
y_1	4	3	1	1
y_2	1	2	8	1
	-2	-4	-10	0

Austausch mit Pivot 8 liefert

	$-x_1$	$-x_2$	$-y_2$	
y_1	$\frac{31}{8}$	$\frac{22}{8}$	$-\frac{1}{8}$	$\frac{7}{8}$
x_3	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$
	$-\frac{6}{8}$	$-\frac{12}{8}$	$\frac{10}{8}$	$\frac{10}{8}$

Ein weiterer Austausch mit Pivot $\frac{22}{8}$ ergibt das optimale Schema

	$-x_1$	$-y_1$	$-y_2$	
x_2	#	#	#	$\frac{7}{22}$
x_3	#	#	#	$\frac{1}{22}$
	$\frac{30}{22}$	$\frac{12}{22}$	$\frac{26}{22}$	$\frac{38}{22}$

Hier ist der maximale Wert also $\frac{19}{11} = \frac{38}{22}$, und er ergibt sich in $x_0 = (0, \frac{7}{22}, \frac{1}{22})$. Die duale Interpretation

x_1	y_1	y_2		
#	#	#	$\frac{7}{22}$	x_2
#	#	#	$\frac{1}{22}$	x_3
$\frac{30}{22}$	$\frac{12}{22}$	$\frac{26}{22}$	$\frac{38}{22}$	

dazu liefert nun die Lösung des ursprünglichen Problems, nämlich den minimalen Wert $\frac{19}{11}$ für $y_1 + y_2$ auf dem zulässigen Bereich im Punkt $y_0 = (\frac{12}{22}, \frac{26}{22})$.

Interpretation des dualen Problems

Die besprochene Dualität besitzt eine ökonomische Interpretation.

Beispiel 2.9.1

In Beispiel 2.1.1 gibt es für den Landwirt es zwei mögliche *Aktivitäten*, nämlich den Anbau von Kartoffeln auf $x_1 \geq 0$ ha und den Anbau von Getreide auf $x_2 \geq 0$ ha. Er kann über drei beschränkte *Ressourcen* verfügen, nämlich

$$\text{Kapital:} \quad 100 x_1 + 200 x_2 \leq 11000$$

$$\text{Arbeitszeit:} \quad x_1 + 4 x_2 \leq 160$$

$$\text{Land:} \quad x_1 + x_2 \leq 100.$$

Der Gewinn ergibt sich als

$$f(x_1, x_2) = 400 x_1 + 1200 x_2.$$

Wir hatten gesehen, dass sich der maximale Gewinn 54000 im Punkt (60, 25) ergibt.

- Frage: Inwiefern ist es sinnvoll, die Schranken 11000, 160 und 100 zu vergrössern? Welcher Preis wäre für den Zukauf von Ressourcen gerechtfertigt?
- An der Maximalstelle gilt in unserem Beispiel: die Restriktionen für Kapital und Arbeitszeit sind *straff*, d.h. mit Gleichheit erfüllt,

$$100 \cdot 60 + 200 \cdot 25 = 11000$$

$$60 + 4 \cdot 25 = 160.$$

Die Restriktion für Land ist nicht straff, man hat

$$60 + 25 < 100.$$

Es würde also nichts bringen, Land dazuzukaufen.

- Welcher Preis je Einheit wäre für zusätzliches Kapital und zusätzliche Arbeitszeit gerechtfertigt? Wir suchen einen Preisvektor

$$y = (y_1, y_2, y_3)^T.$$

Hier stehen y_1 , y_2 und y_3 jeweils für die Preise für eine Einheit Kapital, Arbeitszeit und Land.

Die Kosten am Markt für den Zukauf zur Produktion von Kartoffeln pro ha sind

$$100 y_1 + y_2 + y_3,$$

die Kosten für die Produktion von Getreide pro ha sind

$$200 y_1 + 4 y_2 + y_3.$$

Falls diese Kosten den Gewinn, den der Landwirt je ha erwirtschaften kann, übersteigen, wird er nicht kaufen. Das also ist der Fall wenn

$$100 y_1 + y_2 + y_3 \geq 400$$

$$200 y_1 + 4 y_2 + y_3 \geq 1200.$$

gilt. Dies sind die Restriktionen des dualen Problems. Der Landwirt wird unter diesen Umständen aber eventuell seine Ressourcen verkaufen. Käufer*innen werden interessiert sein, den Gesamtpreis der Ressourcen

$$11000 y_1 + 160 y_2 + 100 y_3$$

zu minimieren. Dieser definiert die Zielfunktion des dualen Problems. Der Preisvektor y ist also die Lösung des dualen Problems.

Das Ausgangsschema des primalen Problems und das daraus folgende optimale Schema sehen so aus:

	$-x_1$	$-x_2$	
y_1	100	200	11000
y_2	1	4	160
y_3	1	1	100
	-400	-1200	0

	$(-)y_1$	$(-)y_2$	
x_1	#	#	60
x_2	#	#	25
y_3	#	#	15
	2	200	54000

Daraus kann man nun ablesen, dass sich das Minimum für das duale Problem in $y = (y_1, y_2, y_3)^T$ mit

$$y_1 = 2, \quad y_2 = 200, \quad y_3 = 0$$

ergibt.

Dass $y_3 = 0$ ist korrespondiert zu der nicht straffen Restriktion für Land: braches Land steht zur Verfügung, Käufer*innen sind nicht bereit, Geld für Land auszugeben.

Der Preisvektor y beschreibt die Preise, die zu bezahlen sinnvoll wäre, um die jeweiligen Ressourcen zu vergrößern. Man nennt solche fiktiven Preise auch *Schattenpreise*.

Bemerkung 2.9.2

Ist eine Restriktion für das primale Problem nicht straff an der Maximalstelle, so gilt für die zugehörige Variable y_i des dualen Problems an der Minimalstelle $y_i = 0$ (Satz vom schwachen komplementären Schlupf).

Allgemein kann man folgende *Aktivitätsanalyse* formulieren:

- Gegeben sind n *Aktivitäten*, durch welche sich eine Ware erzeugen lässt und m verfügbare *Ressourcen*
- c_i = Vorrat, der von der i -ten Ressource zu Verfügung steht.
- a_{ij} = Menge der i -ten Ressource, die bei der j -ten Aktivität verbraucht wird
- b_j = Menge der Ware, die mit der j -ten Aktivität produziert wird
- Gesucht sind die *Aktivitätsniveaus* $x_j \geq 0$, $1 \leq j \leq n$, mit

$$\sum_{j=1}^n a_{ij}x_j \leq c_i, \quad 1 \leq i \leq m,$$

also Verbrauch \leq Ressourcen, für welche die Menge der produzierten Ware maximal wird:

$$\sum_{j=1}^n b_jx_j \stackrel{!}{=} \max.$$

- Beim dazu dualen Problem sucht man *Schattenpreise* y_i pro Einheit der i -ten Ressource, sodass der Gesamtpreis aller Materialien minimiert wird:

$$\sum_{i=1}^m c_i y_i \stackrel{!}{=} \min,$$

und die Nebenbedingungen

$$\sum_{i=1}^m a_{ij} y_i \geq b_j, \quad 1 \leq j \leq n,$$

erfüllt sind. Letztere besagen, dass die Gesamtkosten der j -ten Aktivität mindestens b_j sein sollen.

- Primales und duales Problem können Rollen tauschen:

'Gewinnmaximierung ohne Erschöpfung der Ressourcen'
'Aufwandsminimierung unter Einhaltung minimaler Standards'.

Beispiel 2.9.3

Die Lebensmittel A und B enthalten die Nährstoffe I , II , III und IV gemäss folgender Tabelle in Nährstoffeinheiten pro Mengeneinheit des Lebensmittels:

	I	II	III	IV
A	2.5	2.0	1.5	1.0
B	0.5	1.0	1.5	3.0

Lebensmittel A habe den Preis 20 und Lebensmittel B den Preis 10. Man sucht eine preisgünstigste Kombination beider Lebensmittel, welche aber insgesamt mindestens 140 Einheiten von I , 300 Einheiten von II , 270 Einheiten von III und 300 Einheiten von IV enthält.

Bezeichnen x_1 und x_2 die Einheiten von A bzw. B in der Kombination, so ist $20x_1 + 10x_2$ zu minimieren unter den Nebenbedingungen $x_1 \geq 0$, $x_2 \geq 0$ und

$$2.5x_1 + 0.5x_2 \geq 140$$

$$2.0x_1 + 1.0x_2 \geq 300$$

$$1.5x_1 + 1.5x_2 \geq 270$$

$$1.0x_1 + 3.0x_2 \geq 300.$$

Das dazu duale Problem kann man nun wie folgt interpretieren:

Man sucht die Preise, zu welchen es sich lohnt, die Nährstoffe zuzukaufen, oder anders gesagt, die Preise y_I , y_{II} , y_{III} und y_{IV} , welche ein Hersteller für Präparate mit jeweils einer Einheit I , II , III bzw. IV verlangen kann.

Die Präparate dürfen insgesamt nicht teurer sein als A und B (sonst kaufen wir sie nicht). Das ergibt die Restriktionen

$$2.5 y_I + 2.0 y_{II} + 1.5 y_{III} + 1.0 y_{IV} \leq 20$$

$$0.5 y_I + 1.0 y_{II} + 1.5 y_{III} + 3.0 y_{IV} \leq 10.$$

Der Verkaufserlös

$$140 y_I + 300 y_{II} + 270 y_{III} + 300 y_{IV}$$

soll dabei maximiert werden.

Nun kann man den Simplexalgorithmus ansetzen. Das erste Problem führt nach Transformation auf Normalform auf ein nicht zulässiges Ausgangsschema. Wir betrachten deshalb lieber das Ausgangsschema für das zweite Problem:

	$-y_I$	$-y_{II}$	$-y_{III}$	$-y_{IV}$	
x_1	2.5	2.0	1.5	1.0	20
x_2	0.5	1.0	1.5	3.0	10
	-140	-300	-270	-300	0

Ein Austauschschritt mit Pivot 1.0 liefert das Schema

	$-y_I$	$-x_2$	$-y_{III}$	$-y_{IV}$	
x_1	1.5	-2.0	-1.5	-5	0
y_{II}	0.5	1.0	1.5	3.0	10
	10	300	180	600	3000

Die duale Interpretation dieses Schemas erlaubt es nun, die Lösung für das erste Problem abzulesen, nämlich

$$x_1 = 0 \quad \text{und} \quad x_2 = 300.$$

Als Schattenpreise für die Präparate ergeben sich

$$y_I = y_{III} = y_{IV} = 0 \quad \text{und} \quad y_{II} = 10.$$

Interpretation der Schattenpreise

- Gegeben primales Problem $b^T x \stackrel{!}{=} \max$ unter $x \geq 0$ und $Ax \leq c$
- Wie verändert sich der Wert der Zielfunktion an der Maximalstelle, wenn man c durch $c + \delta$ ersetzt?
- Angenommen, das duale Problem besitzt die *eindeutige* Lösung $y_0 \in \tilde{P}_e$, für die dann gilt, dass

$$c^T y_0 < c^T y, \quad \forall y \in \tilde{P}_e \setminus \{y_0\}.$$

- Durch Variation von c verändert sich der zulässige Bereich

$$\tilde{P} = \{y \geq 0, A^T y \geq b\}$$

des dualen Problems nicht. Also gilt für δ hinreichend klein, dass

$$(c + \delta)^T y_0 < (c + \delta)^T y, \quad y \in \tilde{P}_e \setminus \{y_0\}.$$

- Das bedeutet aber, dass y_0 auch noch optimal ist nach dem Übergang von c zu $c + \delta$.
- Also ist die Änderung des optimalen Wertes der dualen Zielfunktion

$$(c + \delta)^T y_0 - c^T y_0 = \delta^T y_0 = \sum_{i=1}^m \delta_i (y_0)_i$$

und das ist nach Dualitätssatz auch die Änderung des optimalen Wertes der primalen Zielfunktion.

- Der Schattenpreis $(y_0)_i$ gibt also an, *welchen Beitrag eine Vergrößerung der i -ten Ressource zum Gesamterlös ergibt.*

Beschreibung allgemeiner Polyeder

- Eine Linearkombination $\sum_{i=1}^m \lambda_i a_i$ von Elementen $a_1, \dots, a_m \in \mathbb{R}^n$ heisst *Affinkombination*, falls $\sum_{i=1}^m \lambda_i = 1$.
- Die Menge aller Affinkombinationen von Elementen einer Menge $S \subset \mathbb{R}^n$ heisst die *affine Hülle* von S , wir schreiben dafür $\text{aff } S$.
- Eine Teilmenge $M \subset \mathbb{R}^n$ heisst *affiner Unterraum* von \mathbb{R}^n , falls sie ihre eigene affine Hülle ist, $M = \text{aff } M$.
- M ist genau dann ein affiner Unterraum von \mathbb{R}^n , wenn $M = a + U$ gilt für ein $a \in \mathbb{R}^n$ und einen linearen Unterraum U von \mathbb{R}^n .

Der lineare Unterraum U ist dabei eindeutig bestimmt als $U = M - a$. Die *Dimension* von M ist definiert als $\dim M := \dim U$.

- Der Schnitt affiner Unterräume ist wieder ein affiner Unterraum.
- Die affine Hülle einer Menge $S \subset \mathbb{R}^n$ ist der kleinste affine Unterraum von \mathbb{R}^n , der S enthält.

Darstellungssatz für Polyeder

Ab nun sei

$$P := \bigcap_{i=1}^m \{f_i \leq c_i\}$$

ein gegebenes Polyeder.

Lemma 2.10.1

Für jede Seite S einer konvexen Menge $A \subset \mathbb{R}^n$ gilt

$$S = A \cap \text{aff } S.$$

Definition 2.10.2

Für jedes i nennen wir

$$H_i := \{f_i = c_i\}$$

eine *Restriktionshyperebene (RHE)* von P .

Man kann eine Beschreibung der affinen Hülle einer Seite von P als Schnitt 'guter' RHE erhalten:

Satz 2.10.3

Sei $S \neq \emptyset$ eine Seite von P und $I := \{i \in \{1, \dots, m\} : S \subset H_i\}$. Dann ist

$$\text{aff } S = \bigcap_{i \in I} H_i.$$

Korollar 2.10.4

Für jede Seite $S \neq \emptyset$ von P gilt

$$S = P \cap \bigcap_{i: S \subset H_i} H_i,$$

wobei die H_i die RHE von P bezeichnen.

Bemerkung 2.10.5

① $S = P$ in Satz 2.10.3 liefert

$$\text{aff } P = \bigcap_{i: P \subset H_i} H_i.$$

② Eine RHE H_i , für die $P \subset H_i$ gilt, heißt *singuläre* RHE.

③ Wir definieren die Dimension von P als

$$\dim P := \dim \text{aff } P.$$

Wir definieren das *relative Innere* von P als

$$\text{int}_r P := \{x \in P : B_\delta(x) \cap \text{aff } P \subset P \text{ für ein } \delta > 0\}$$

und den *relativen Rand* von P als

$$\partial_r P := P \setminus \text{int}_r P.$$

- 4 Da $\text{aff } P \subset \mathbb{R}^n$ abgeschlossen ist, stimmen Abschluss und relativer Abschluss einer Menge in $\text{aff } P$ überein. Insbesondere ist also

$$\partial_r P = \overline{P} \setminus \text{int}_r P = P \setminus \text{int}_r P.$$

- 5 Der Schnitt über alle nichtsingulären RHE liefert das relative Innere:

$$\text{int}_r P = P \cap \bigcap_{i: P \not\subset H_i} \{f_i < c_i\}.$$

- 6 Jeder Punkt $x \in \partial_r P$ ist Element einer *echten Seite* S von P .
7 Umgekehrt gilt für jede echte Seite S von P , dass

$$S \subset \partial_r P.$$

(Übung)

Korollar 2.10.6

Ist S eine echte Seite von P , so gilt

$$\dim S < \dim P.$$

Definition 2.10.7

Ein nichtleeres Polyeder heisst *primitiv*, wenn es nicht die konvexe Hülle seiner echten Seiten ist.

Beispiel 2.10.8

- Affine Unterräume (z.B. Punkte, Geraden, Ebenen) sind primitive Polyeder.
- Affine Halbräume, d.h. nichtleere Schnitte eines affinen Unterraums M mit einem Halbraum $\{f \leq c\}$ mit $M \not\subset \{f \leq c\}$ (z.B. Halbgeraden, Halbebenen), sind primitive Polyeder.

Bemerkung 2.10.9

Jeder affine Halbraum ist die Summe einer Halbgeraden und eines Untervektorraumes, d.h. von der Gestalt

$$x_0 + \mathbb{R}_+ \cdot v + U,$$

hier ist x_0 ein fixierter Punkt, $\mathbb{R}_+ \cdot v$ eine Halbgerade (definiert durch einen gegebenen Vektor v), und U ist ein linearer Unterraum.
(Übung)

Satz 2.10.10

Jedes Polyeder $P \neq \emptyset$ ist die konvexe Hülle seiner primitiven Seiten.

Satz 2.10.11

Ein primitives Polyeder P ist entweder ein affiner Unterraum oder ein affiner Halbraum.

Aus den Sätzen 2.10.10 und 2.10.11 erhalten wir nun einen ersten *Darstellungssatz für Polyeder*:

Korollar 2.10.12

Jedes Polyeder ist die konvexe Hülle endlich vieler affiner Unterräume und affiner Halbräume.

Definition 2.10.13

- (i) Eine Linearkombination $\sum_{i=1}^m \lambda_i a_i$ von Vektoren $a_i \in \mathbb{R}^n$ heisst *konisch*, falls $\lambda_i \geq 0$ für alle i .
- (ii) Eine konvexe Menge $K \subset \mathbb{R}^n$ ist ein *konvexer Kegel*, falls für jedes $x \in K$ und $\lambda \geq 0$ auch $\lambda x \in K$ gilt.
- (iii) Für $S \subset \mathbb{R}^n$ heisst

$$\text{cone}(S) := \left\{ \sum_{i=1}^m \lambda_i a_i : m \in \mathbb{N}, a_i \in S, \lambda_i \geq 0, i = 1, \dots, m \right\}$$

der von S erzeugte konvexe Kegel.

Bemerkung 2.10.14

- (i) $\text{cone}(S)$ ist der kleinste konvexe Kegel, der S enthält.
- (ii) $0 \in K$ für jeden konvexen Kegel K .
- (iii) Für eine reelle $(m \times n)$ -Matrix ist $\{Ax \leq 0\}$ ein *polyedrischer Kegel*, d.h. ein konvexer Kegel, der auch ein Polyeder ist.

Lemma 2.10.15

Jeder affine Unterraum und jeder affine Halbraum ist von der Form

$$u + \text{cone}(Y)$$

mit $u \in \mathbb{R}^n$ und $Y = \{y_1, \dots, y_m\} \subset \mathbb{R}^n$ endlich.

Satz 2.10.16

Für jedes Polyeder P gibt es endliche Mengen $X \subset \mathbb{R}^n$ und $Y \subset \mathbb{R}^n$, sodass

$$P = k(X) + \text{cone}(Y).$$

Lemma 2.10.17

Sei P ein Polyeder. In der Zerlegung

$$P = K + S$$

mit $K = k(\{x_1, \dots, x_k\})$ und $S = \text{cone}(\{y_1, \dots, y_l\})$ ist S eindeutig bestimmt. Für jedes $x_0 \in P$ gilt

$$S = \{y \in \mathbb{R}^n : x_0 + \lambda y \in P \text{ für alle } \lambda \geq 0\}.$$

Ist $P = \bigcap_{i=1}^m \{f_i \leq c_i\}$, so gilt $S = \bigcap_{i=1}^m \{f_i \leq 0\}$.

Korollar 2.10.18

Ist S ein polyedrischer Kegel, so gibt es eine endliche Menge

$Y = \{y_1, \dots, y_m\} \subset \mathbb{R}^n$ mit

$$S = \text{cone}(Y).$$

Lösbarkeit von Optimierungsproblemen

Satz 2.10.19

Sei $P \neq \emptyset$ ein Polyeder und f eine Linearform, die auf P nach oben beschränkt ist. Dann nimmt $f|_P$ auf P ihr Maximum an.

Darstellungssatz von Weyl

Satz 2.10.20

Sei

$$P := \text{cone}(Y)$$

für $Y = \{y_1, \dots, y_k\} \in \mathbb{R}^n$ endlich. Dann gibt es eine reelle $(m \times n)$ -Matrix A mit

$$P = \{x \in \mathbb{R}^n : Ax \leq 0\},$$

d.h. P ist ein polyedrischer Kegel.

Satz 2.10.21 (Darstellungssatz von Weyl)

Seien X und Y endliche Teilmengen des \mathbb{R}^n . Dann ist

$$k(X) + \text{cone}(Y)$$

ein Polyeder.

Korollar 2.10.22

Die folgenden Aussagen sind äquivalent:

- 1 P ist ein beschränktes Polyeder.
- 2 P ist die konvexe Hülle endlich vieler Punkte.

Definition 2.10.23

Sei $P = K + S$ ein Polyeder, wobei K ein beschränktes Polyeder ist und S der gemäss Lemma 2.10.17 eindeutig bestimmte polyedrische Kegel.

- 1 Der Vektorraum $L = S \cap (-S)$ heisst der *Linienraum* von P .
- 2 Das Polyeder P heisst *spitz*, falls $P_e \neq \emptyset$.

Lemma 2.10.24

Sei S ein polyedrischer Kegel. Dann ist

$$L := S \cap (-S)$$

der grösste in S enthaltene Untervektorraum.

Satz 2.10.25

Die folgenden Aussagen sind äquivalent:

- 1 Das Polyeder P ist spitz.
- 2 Der Linienraum von P ist trivial, $L = \{0\}$.
- 3 Ist $P = \{x \in \mathbb{R}^n : Ax \leq c\}$, so ist der Rang von A gleich n .

Bemerkung 2.10.26

Zusätzlich $x \geq 0$ in (3) stellt für ein lineares Optimierungsproblem sicher, dass der zulässige Bereich spitz ist.

Anwendungen auf lineare Optimierungsprobleme

Im folgenden Satz sei wie üblich

$$P = \bigcap_{i=1}^m \{f_i \leq c_i\} = K + S$$

mit Linienraum $L = S \cap (-S)$.

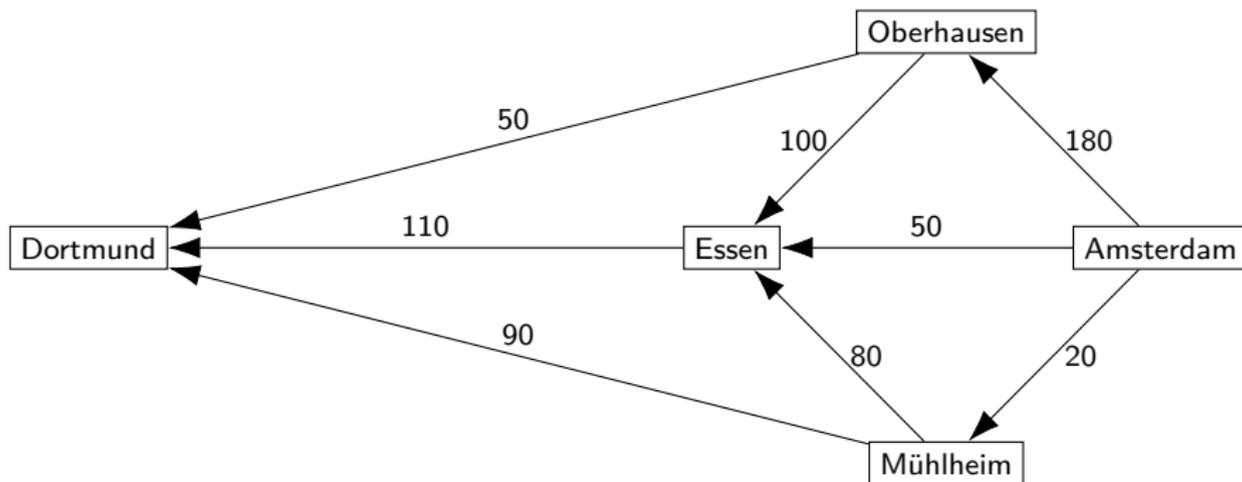
Satz 2.10.27

- 1 Wir haben $\sup_P f < +\infty$ genau dann, wenn $f|_S \leq 0$. In diesem Falle nimmt die Funktion $f|_P$ ihr Maximum an.
- 2 Ist $f|_L \neq 0$, so ist $\sup_P f = +\infty$.
- 3 Ist P spitz und $\sup_P f < +\infty$, so nimmt $f|_P$ ihr Maximum in einem Extrempunkt von P an.

Flussprobleme

Beispiel 3.1.1

Bei der Schokofahrt werden per Rad Schokoladentafeln von Amsterdam nach Dortmund gefahren. Dabei wird zwischendurch je nach Kapazitäten umgeladen.



Wieviele Tafeln können mit diesen Kapazitäten insgesamt befördert werden ?

Definition 3.1.2

Ein *gerichteter Graph* ist ein Paar $G = (X, \Gamma)$, bestehend aus einer Menge X und einer Teilmenge

$$\Gamma \subset (X \times X) \setminus \{(x, x) : x \in X\},$$

sodass $(y, x) \notin \Gamma$ falls $(x, y) \in \Gamma$. Die Elemente $x \in X$ heissen *Knoten* oder *Punkte* des Graphen, die Elemente $\gamma \in \Gamma$ heissen (*gerichtete*) *Kanten*.

Definition 3.1.3

Ein *Netzwerk* ist ein Quadrupel (G, a, b, c) , wobei

- (i) $G = (X, \Gamma)$ ein gerichteter Graph ist mit einer endlichen Menge X ,
- (ii) $a, b \in X$ mit $(b, a) \in \Gamma$; a heisst *Quelle*, b *Senke*, (b, a) *Kontrollkante*,
- (iii) $c : \Gamma \rightarrow \mathbb{N} \cup \{+\infty\}$ ist eine Funktion mit $c(\gamma) \in \mathbb{N}$ für alle $\gamma \in \Gamma \setminus \{(b, a)\}$ und $c((b, a)) = +\infty$; $c(\gamma)$ heisst die *Kapazität* der Kante γ .

Bemerkung 3.1.4

dfnen von Netzwerken können sich in der Literatur unterscheiden.

Definition 3.1.5

- (i) Ein *Fluss* in einem Netzwerk $((X, \Gamma), a, b, c)$ ist eine Abbildung $\varphi : \Gamma \rightarrow \mathbb{N}$, welche die *Kirchhoffsche Regel*

$$\sum_{\gamma \in X \rightarrow} \varphi(\gamma) = \sum_{\gamma \in X \leftarrow} \varphi(\gamma) \quad \text{für alle } x \in X$$

erfüllt. Hier ist

$$X \rightarrow := \{(x, y) \in (X \times X) : (x, y) \in \Gamma\}$$

die Menge aller von x ausgehenden Kanten und

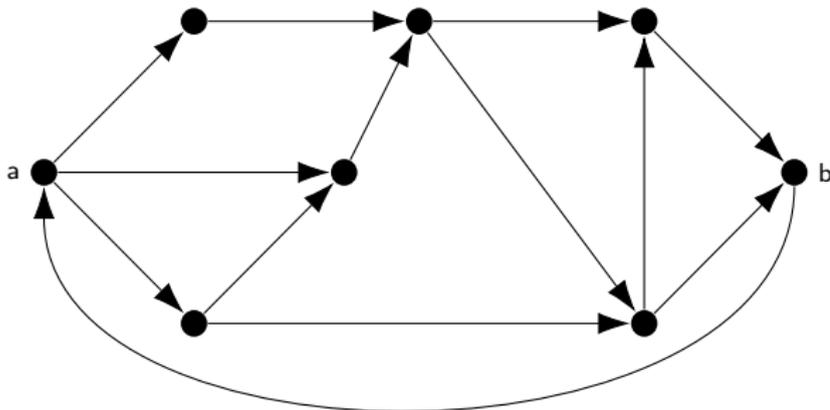
$$X \leftarrow := \{(y, x) \in (X \times X) : (y, x) \in \Gamma\}$$

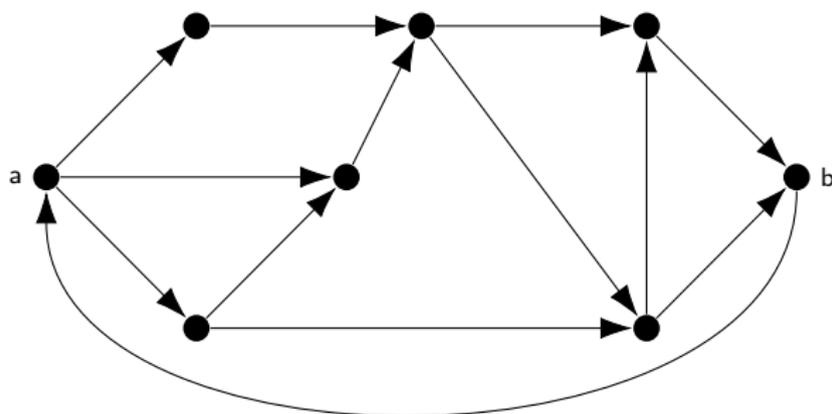
die Menge aller zu x hinführenden Kanten.

- (ii) Ein Fluss φ heisst *zulässig*, falls $\varphi(\gamma) \leq c(\gamma)$ für alle $\gamma \in \Gamma$.
- (iii) Ist φ ein Fluss, so heisst $\varphi(b, a)$ der *Wert* von φ . Ein zulässiger Fluss mit maximalem Wert heisst *Maximalfluss*.

Bemerkung 3.1.6

- (i) Ein Fluss beschreibt einen Transport von Masse von der Quelle a zur Senke b , bei dem an keiner Stelle Masse verloren geht. Letzteres ist codiert in der Kirchhoffsche Regel, die sagt, dass an jedem Knoten die eingehende Gesamtmasse gleich der ausgehenden sein muss.
- (ii) Ein zulässiger Fluss überschreitet auf keiner Kante die Kapazität. Die Kapazität kann man als das Durchleitungsvermögen der Kante interpretieren.
- (iii) Die Kontrollkante ist formal, das heißt sie kommt nicht aus dem praktischen Problem. Sie wird zusätzlich eingeführt, damit die Kirchhoffsche Regel auch in der Quelle a und der Senke b gelten.





- (iv) Der Gesamtfluss von der Quelle a zu der Senke b fließt über die Kontrollkante zurück, ist also gleich dem Wert $\varphi(b, a)$ des Flusses φ . Weil dieser Rückfluss durch die Kontrollkante nicht beschränkt sein soll, setzt man $c(b, a) = +\infty$.

Für eine gegebene Menge $A \subset X$ von Knoten definieren wir nun ganz allgemein die Menge

$$A_{\rightarrow} := \{(x, y) \in \Gamma : x \in A, y \notin A\}$$

der aus A herausführenden Kanten und die Menge

$$A_{\leftarrow} := \{(x, y) \in \Gamma : x \notin A, y \in A\}$$

der nach A hineinführenden Kanten.

Lemma 3.1.7

Sei φ ein Fluss und $A \subset X$. Dann hat man

$$\sum_{\gamma \in A_{\rightarrow}} \varphi(\gamma) = \sum_{\gamma \in A_{\leftarrow}} \varphi(\gamma).$$

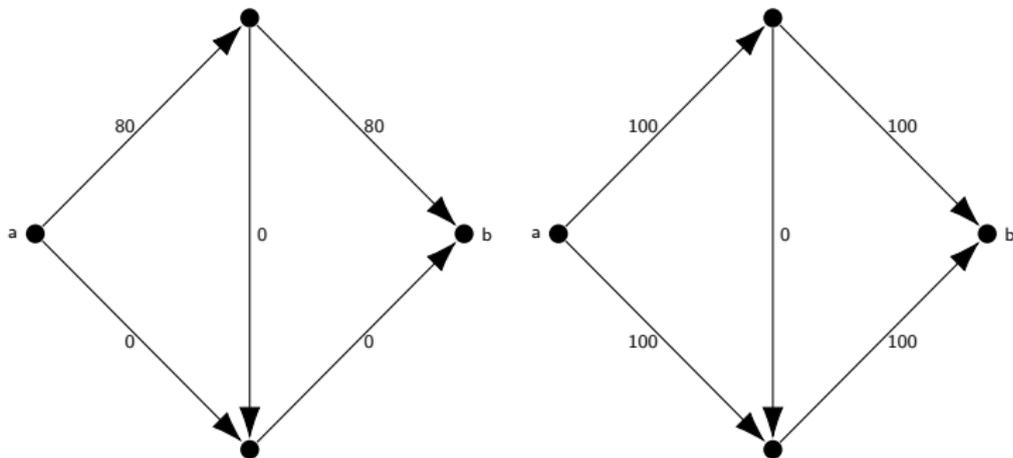
Korollar 3.1.8

Sei φ ein zulässiger Fluss und sei $A \subset X$ so, dass $a \in A$, aber $b \notin A$. Dann ist

$$\varphi(b, a) \leq \sum_{\gamma \in A_{\rightarrow}} c(\gamma).$$

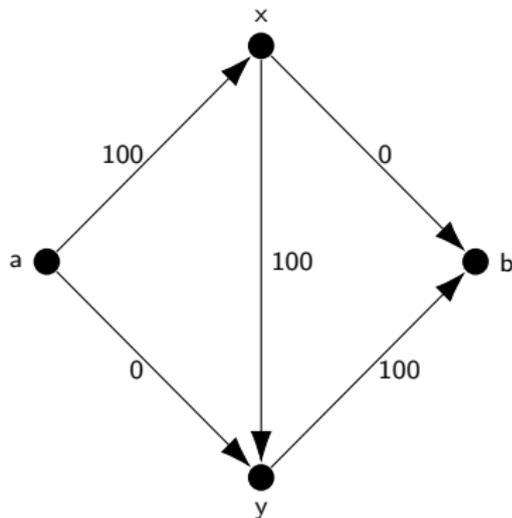
Der Algorithmus von Ford-Fulkerson: Vorbetrachtung

- Wir betrachten ein rekursives Verfahren zur Vergrößerung des Wertes von zulässigen Flüssen. Gegeben sei ein Netzwerk, dessen Kanten alle die Kapazität 100 haben, und ein Fluss mit Werten:



- Durch *Vergrössern* des Flusses durch geeignete Kanten erhält man (rechtes Diagramm) einen zulässigen maximalen Fluss.

- Manchmal ist aber auch ein *Verkleinern* des Flusses durch einzelne Kanten nötig: in dem zulässigen Ausgangsfluss der Form



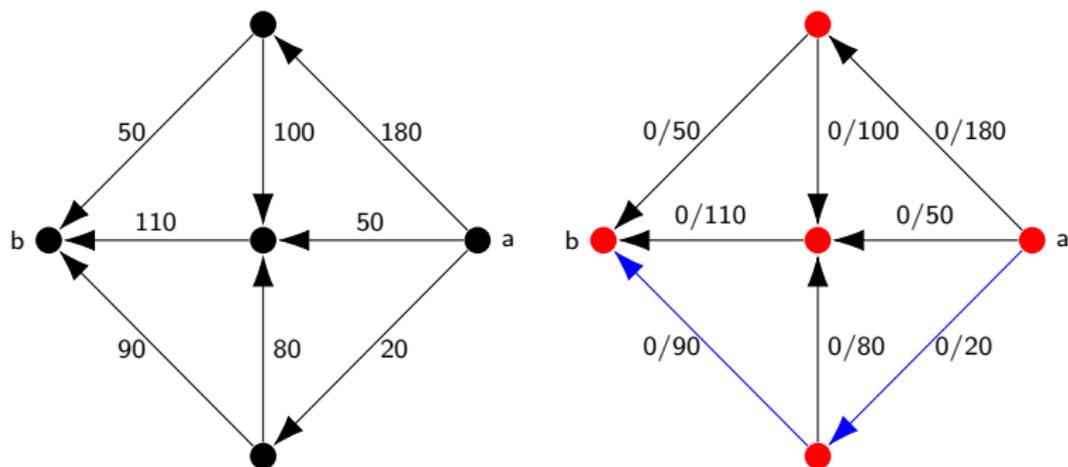
muss der Fluss durch (x, y) auf Null reduziert werden.

- Im Allgemeinen benötigt man beide Mechanismen, Vergrößerung und Verkleinerung entlang Kanten, um zu einem zulässigen Fluss mit maximalem Wert zu kommen.

Algorithmus von Ford-Fulkerson: Beispiele

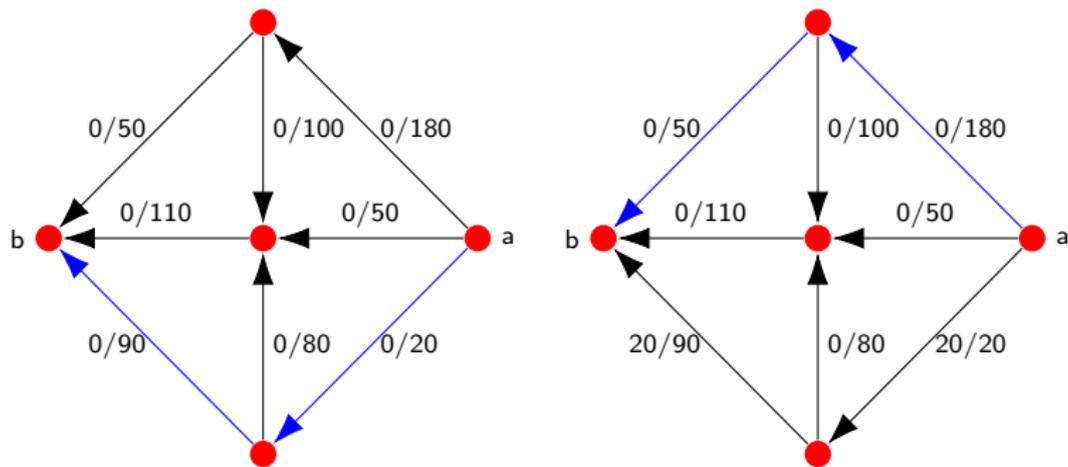
Beispiel 3.2.1

Wir erinnern uns an Beispiel 3.1.1. In abstrakter Form sieht das zugehörige Netzwerk so aus:



Für den Ausgangsfluss $\varphi_0 \equiv 0$ kann man hier alle Knoten rot markieren, und wir können einen blauen Weg von a nach b wählen.

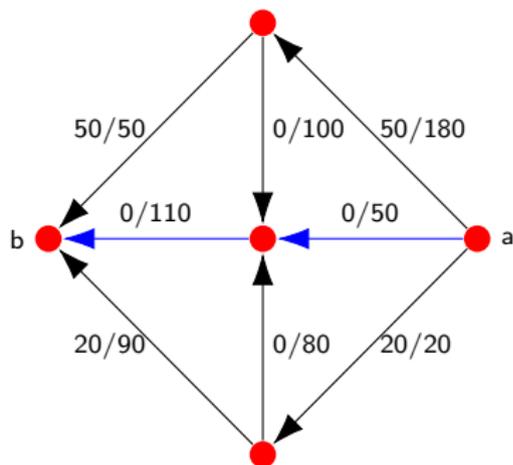
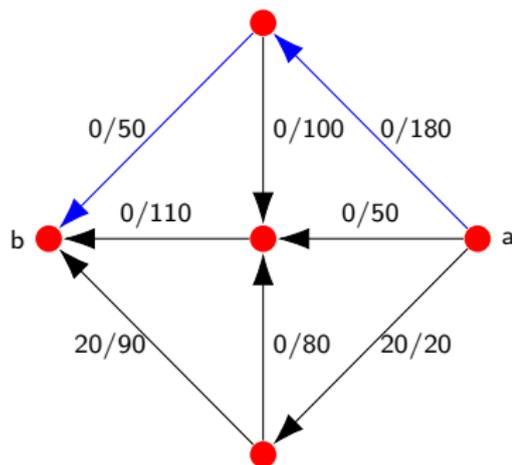
Man hat $\delta = \min(20, 90) = 20$ und erhält (rechtes Diagramm) damit einen neuen Fluss φ_1 :



Hier (rechts) konnten wir wieder alle Knoten rot markieren und einen blauen Weg wählen.

Diesmal ergibt sich $\delta = \min(180, 50) = 50$.

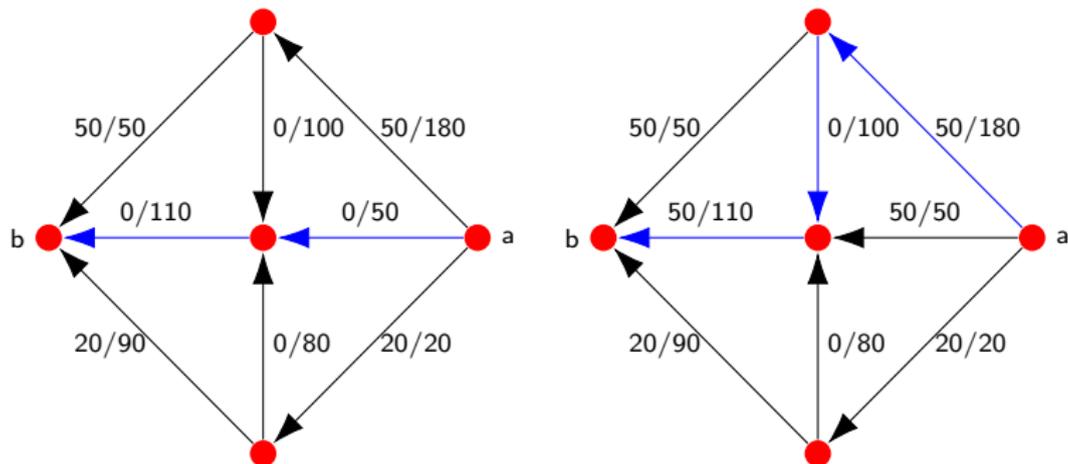
Der neue Fluss φ_2 ergibt sich nun wie im rechten Diagramm



Wieder können alle Knoten rot markiert werden, und wir finden einen blauen Weg von a nach b .

Diesmal folgt $\delta = \min(50, 110) = 50$.

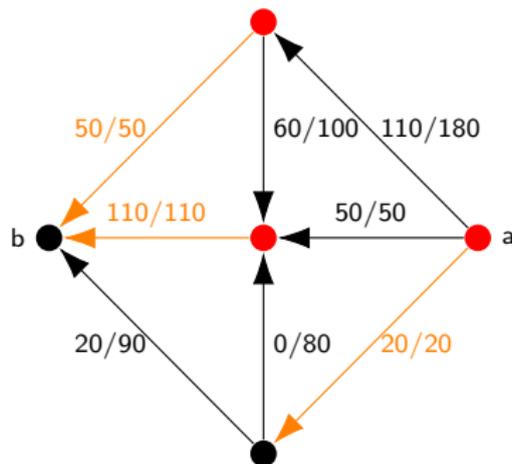
Wir bekommen rechts einen neuen Fluss φ_3



Wieder können alle Knoten markiert werden, und wir können immer noch einen blauen Weg von a nach b finden.

Nun ist $\delta = \min(130, 100, 60) = 60$.

Für den neuen Fluss φ_4 können ausser a nur noch zwei weitere Punkte markiert werden:

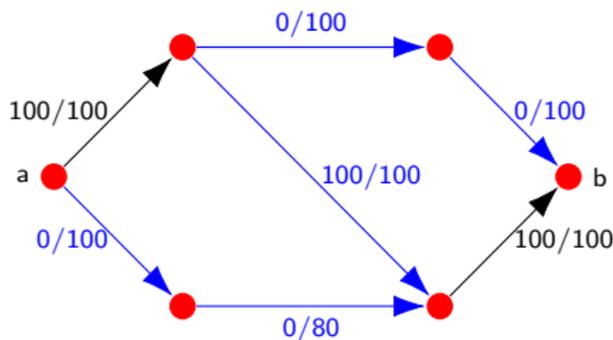
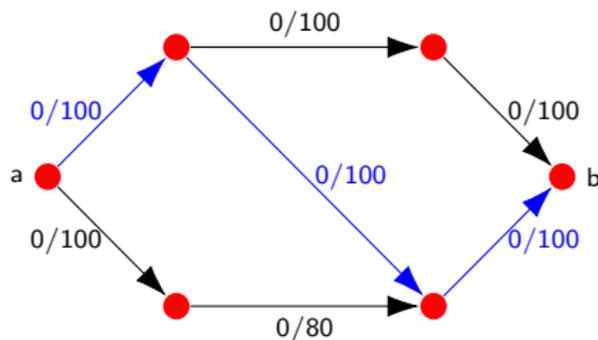


Da b nicht markiert werden kann, ist φ_4 ein Maximalfluss. Sein Wert ist

$$\varphi(b, a) = \sum_{\gamma \in a \rightarrow} \varphi(\gamma) = \sum_{\gamma \in b \leftarrow} \varphi(\gamma) = 110 + 50 + 20 = 180.$$

Beispiel 3.2.2

Alle Knoten können markiert werden, und wir finden einen blauen Weg von a nach b entlang markierter Kanten.

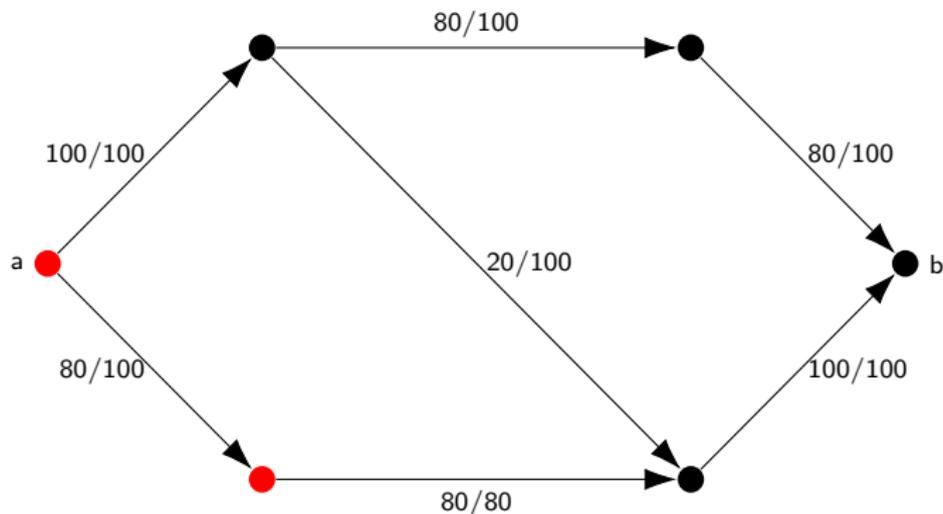


Für diesen Weg ist $\delta = 100$, das ergibt einen neuen Fluss φ_1 im rechten Diagramm.

Wieder können alle Knoten markiert werden, und wir können den eingezeichneten blauen Weg wählen, nun mit einer Rückwärtskante.

Diesmal ist $\delta = 80$.

Wir bekommen einen neuen Fluss φ_2 mit Werten:



Da b nicht mehr markiert werden kann, ist φ_2 ein Maximalfluss mit Wert 180.

In diesem Beispiel haben wir den Fluss nicht nur entlang geeigneter Kanten vergrößert, sondern auch auf der Rückwärtskante verkleinert, um den Flusswert insgesamt zu vergrößern.

Der Algorithmus

- (1) Wähle den zulässigen Ausgangsfluss $\varphi \equiv 0$.
- (2) Zu einem gegebenen Fluss φ markiere Punkte des Graphen durch Anwenden der folgenden Regeln, solange möglich:

Regel 0: Markiere Quelle a .

Regel 1: Ist $\gamma = (x, y) \in \Gamma$, x markiert, y nicht markiert und $\varphi(\gamma) < c(\gamma)$, so markiere y .

Regel 2: Ist $\gamma = (x, y) \in \Gamma$, $\gamma \neq (b, a)$, x nicht markiert, y markiert und $\varphi(\gamma) > 0$, so markiere x .

- (3) Wenn b nicht markiert ist, dann Stopp (haben optimalen Fluss gefunden).
Wenn b markiert, dann gibt es nach Konstruktion einen Weg von a nach b längs gewisser Kanten aus $\Gamma \setminus \{(b, a)\}$ (ohne Beachtung der Richtung) durch markierte Punkte, d.h. durch Punkte

$$a = a_1, a_2, \dots, a_n = b,$$

sodass für $2 \leq i \leq n$ gilt:

Der Algorithmus (Fortsetzung)

(3) (Forts.) Entweder

(α) $(a_{i-1}, a_i) \in \Gamma$ und $\varphi(a_{i-1}, a_i) < c(a_{i-1}, a_i)$ oder

(β) $(a_i, a_{i-1}) \in \Gamma$ und $\varphi(a_i, a_{i-1}) > 0$.

Setze

$$\delta_\alpha := \min\{c(a_{i-1}, a_i) - \varphi(a_{i-1}, a_i) : (a_{i-1}, a_i) \in \Gamma\},$$

$$\delta_\beta := \min\{\varphi(a_i, a_{i-1}) : (a_i, a_{i-1}) \in \Gamma\},$$

wobei das $\min \emptyset := +\infty$ gesetzt wird, und

$$\delta := \min(\delta_\alpha, \delta_\beta) > 0.$$

Definiere einen neuen Fluss $\tilde{\varphi}$ durch

$$\tilde{\varphi}(a_{i-1}, a_i) := \varphi(a_{i-1}, a_i) + \delta \quad \text{falls } (a_{i-1}, a_i) \in \Gamma$$

und

$$\tilde{\varphi}(a_i, a_{i-1}) := \varphi(a_i, a_{i-1}) - \delta \quad \text{falls } (a_i, a_{i-1}) \in \Gamma$$

für $1 \leq i \leq n$, wobei $a_0 := b$, und

$$\tilde{\varphi}(\gamma) := \varphi(\gamma) \quad \text{für alle übrigen } \gamma \in \Gamma.$$

(4) Gehe zu (2).

Proposition 3.2.3

Sei

$$a =: a_1, a_2, \dots, a_n := b$$

ein Weg von a nach b durch markierte Punkte wie in Schritt 3 des Algorithmus und sei $a_0 := b$. Für $\tilde{\varphi}$ gilt:

- (i) $\tilde{\varphi}$ ist ein Fluss (erfüllt Kirchhoffsche Regel).
- (ii) $\tilde{\varphi}$ ist zulässig.
- (iii) $\tilde{\varphi}$ hat einen grösseren Wert als φ , d.h.

$$\tilde{\varphi}(b, a) > \varphi(b, a).$$

Satz 3.2.4

Der Ford-Fulkerson-Algorithmus bricht nach endlich vielen Schritten ab und liefert dabei einen zulässigen Fluss mit maximalem Wert.

Max-flow Min-cut Theorem

Definition 3.2.5

Sei $((X, \Gamma), a, b, c)$ ein Netzwerk und $\Gamma' := \Gamma \setminus \{(b, a)\}$. Wir nennen eine endliche Folge von Knoten

$$a = x_0, x_1, \dots, x_m = b$$

einen *Weg* von a nach b , falls für jedes $i = 1, \dots, m$ entweder

$$(x_{i-1}, x_i) \in \Gamma'$$

oder

$$(x_i, x_{i-1}) \in \Gamma'$$

gilt.

Definition 3.2.6

Sei φ ein zulässiger Fluss auf einem Netzwerk $((X, \Gamma), a, b, c)$ und $\Gamma' := \Gamma \setminus \{(b, a)\}$. Ein Weg

$$a = x_0, x_1, \dots, x_m = b$$

heisst *vergrößernder Weg* bezüglich φ , falls für jedes $i = 1, \dots, m$ gilt: Entweder

(α) $(x_{i-1}, x_i) \in \Gamma'$ und $\varphi(x_{i-1}, x_i) < c(x_{i-1}, x_i)$ oder

(β) $(x_i, x_{i-1}) \in \Gamma'$ und $\varphi(x_i, x_{i-1}) > 0$.

Im Fall (α) nennt man (x_{i-1}, x_i) eine *Vorwärtskante*, im Fall (β) nennt man (x_i, x_{i-1}) eine *Rückwärtskante*.

Korollar 3.2.7

Ein zulässiger Fluss φ auf einem Netzwerk ist genau dann ein Maximalfluss, wenn es keinen vergrößernden Weg bzgl. φ gibt.

Definition 3.2.8

Sei $((X, \Gamma), a, b, c)$ ein Netzwerk und $A \subset X$ eine Menge von Knoten mit $a \in A$ und $b \notin A$. Dann heisst

$$\sum_{\gamma \in A \rightarrow} c(\gamma)$$

die *Schnittkapazität* von A . Als *Schnitt* bezeichnet man die Zerlegung von X in die disjunkten Knotenmengen A und A^c .

Korollar 3.2.9

Der Wert eines Maximalflusses in einem Netzwerk ist gleich der minimalen Schnittkapazität.

Bemerkung 3.2.10

Der Algorithmus von Ford-Fulkerson liefert damit also auch ein konstruktives Verfahren zur Bestimmung des Schnittes mit minimaler Kapazität.

Zyklenzerlegung für Flüsse

Definition 3.2.11

- (i) Ein Graph $G_1 = (X_1, \Gamma_1)$ heisst ein *Teilgraph* eines Graphen $G = (X, \Gamma)$, falls

$$X_1 \subset X \quad \text{und} \quad \Gamma_1 \subset \Gamma.$$

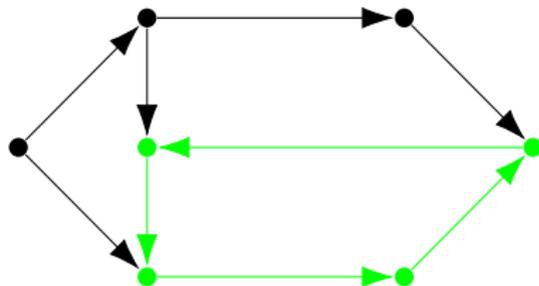
- (ii) Sei $G = (X, \Gamma)$ ein Graph. Ein Teilgraph $K = (X', \Gamma')$ von G heisst *Zyklus* (oder *Kreis*), falls es Punkte $a_1, \dots, a_k \in X$ gibt mit $X' = \{a_1, \dots, a_k\}$ und

$$\Gamma' = \{(a_1, a_2), (a_2, a_3), \dots, (a_{k-1}, a_k), (a_k, a_1)\}.$$

- (iii) Ist $N = ((X, \Gamma), a, b, c)$ ein Netzwerk und φ ein zulässiger Fluss auf N , so heisst φ *Fluss entlang eines Zyklus* $K = (X', \Gamma')$, falls $K = (X', \Gamma')$ ein Zyklus ist und

$$\varphi(\gamma) = 0 \quad \text{für alle } \gamma \in \Gamma \setminus \Gamma'.$$

- Hier ist ein Beispiel für einen Zyklus $K = (X', \Gamma')$:



- Aus der Kirchhoff-Regel folgt, dass ein Fluss φ entlang eines Zyklus $K = (X', \Gamma')$ auf dem Zyklus konstant sein muss, das bedeutet es gibt eine Zahl $c_{\Gamma'} \in \mathbb{N}$ mit

$$\varphi(\gamma) = c_{\Gamma'}, \quad \gamma \in \Gamma'.$$

Satz 3.2.12 (Zyklenerlegung)

Für jeden zulässigen Fluss in einem Netzwerk N existieren zulässige Flüsse $\varphi_1, \dots, \varphi_m$ längs Zyklen, sodass gilt

$$\varphi = \varphi_1 + \dots + \varphi_m.$$

Bemerkung 3.2.13

- (i) Gibt es in einem Netzwerk einen zulässigen Fluss $\varphi \neq 0$, so gibt es einen Zyklus.
- (ii) Der Induktionsschritt liefert ein konstruktives Verfahren zur Zerlegung eines Flusses in Flüsse entlang Zyklen.
- (iii) Ist $\varphi = \varphi_1 + \dots + \varphi_m$, so folgt

$$\varphi(b, a) = \varphi_1(b, a) + \dots + \varphi_m(b, a),$$

d.h. der Wert von φ ist die Summe der Werte der Flüsse φ_i entlang derjenigen Zyklen, die die Kontrollkante (b, a) enthalten.

Korollar 3.2.14

Sei $N = (G, a, b, c)$ ein Netzwerk. Dann existieren Zyklen $K_i = (X_i, \Gamma_i)$ in G mit $(b, a) \in \Gamma_i$ und zulässige Flüsse φ_i entlang K_i , $1 \leq i \leq n$, sodass

$$\varphi = \varphi_1 + \dots + \varphi_n$$

ein Maximalfluss ist.

Definition 3.2.15

Sei $((X, \Gamma), a, b, c)$ ein Netzwerk. Wir nennen einen Weg

$$a = x_0, x_1, \dots, x_m = b$$

von a nach b *gerichtet*, falls für alle $i = 1, \dots, m$ gilt, dass

$$(x_{i-1}, x_i) \in \Gamma' = \Gamma \setminus \{(b, a)\}.$$

(Falls er also mit den Richtungen der darin vorkommenden Kanten kompatibel ist).

Satz 3.2.16

Sei $N = ((X, \Gamma), a, b, c)$ ein Netzwerk und \mathcal{T} die Familie aller Teilmengen $T \subset \Gamma$ derart, dass jeder gerichtete Weg von a nach b mindestens eine Kante von T enthält. Dann gilt für den Wert α eines Maximalflusses die Gleichheit

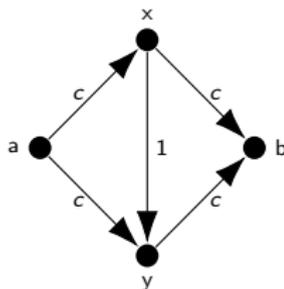
$$\alpha = \min_{T \in \mathcal{T}} \sum_{\gamma \in T} c(\gamma).$$

Algorithmus von Edmond-Karp

- Ziel = Abschätzung für die Anzahl der benötigten Schritte im Algorithmus von Ford-Fulkerson
- $((X, \Gamma), a, b, c)$ Netzwerk, $\Gamma' := \Gamma \setminus \{(b, a)\}$.

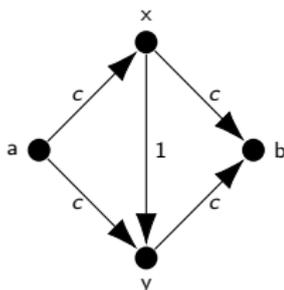
Bemerkung 3.3.1

- (i) Wählt man in FF die vergrößernden Wege ungünstig, so ist die Anzahl der notwendigen Flussvergrößerungen von den Kapazitäten abhängig. Hier



kann der Maximalfluss mit Wert $2c$ mit 2 Flussvergrößerungen erreicht werden.

(i) (Forts.)



Es ist aber auch möglich, als vergrößernde Wege abwechselnd

$$a, x, y, b \quad \text{und} \quad a, y, x, b$$

zu wählen. Dann braucht man $2c$ Schritte.

(ii) Vermeide Kapazitätsabhängigkeit durch Wahl vergrößernder Wege kürzester Länge.

Dazu ändere das Markierungsverfahren im Schritt 2 des Algorithmus:

(ii) (Forts.)

- im ersten Schritt markiere alle Punkte, die von a aus markiert werden können
- im jeweils $(k + 1)$ -ten Teilschritt *alle* Punkte, die ausgehend von *allen* im k -ten Teilschritt markierten erreicht werden können ('Breitensuche').

Nach dem ersten Erreichen von b (falls erreichbar) liefert Rückverfolgung über einzelne Teilschritte einen kürzesten vergrößernden Weg.

(iii) Implementiert man diese Änderungen, erhält man den *Algorithmus von Edmond-Karp*.

- Im Folgenden sei φ_0 ein zulässiger Fluss (z.B. $\varphi_0 = 0$) und

$$\varphi_0, \varphi_1, \dots, \varphi_n$$

eine endliche Folge von zulässigen Flüssen, bei denen φ_{k+1} aus φ_k durch Abänderung längs eines vergrößernden Weges bezüglich φ_k kürzester Länge gewonnen wird.

Definition 3.3.2

Sei $0 \leq k \leq n$, $y, z \in X$, $y \neq z$. Ein *vergrößernder Weg* von y nach z bzgl. φ_k ist eine endliche Folge von Knoten

$$y = x_0, x_1, \dots, x_l = z,$$

sodass für $1 \leq i \leq l$ entweder

$$(x_{i-1}, x_i) \in \Gamma' \quad \text{und} \quad \varphi_k(x_{i-1}, x_i) < c(x_{i-1}, x_i)$$

oder

$$(x_i, x_{i-1}) \in \Gamma' \quad \text{und} \quad \varphi_k(x_i, x_{i-1}) > 0$$

gilt. Die Zahl l nennen wir die *Länge* des Weges. Wir schreiben

- $l_k(y, z)$ für die kürzeste Länge eines vergrößernden Weges von y nach z bezüglich φ_k ,
- $l_k(y, z) := +\infty$, falls kein solcher vergrößernder Weg existiert,
- $l_k(x, x) := 0$ für alle $x \in X$.

Lemma 3.3.3

Sei $0 \leq k \leq n$, $m := l_k(a, b)$ und sei

$$a = x_0, x_1, \dots, x_m = b$$

ein kürzester vergrößernder Weg von a nach b bzgl. φ_k . Dann gilt für jedes $x = x_j$, $0 \leq j \leq m$, dass

$$l_k(a, x) = j \quad \text{und} \quad l_k(x, b) = m - j.$$

Korollar 3.3.4

Sei $0 \leq k \leq n$ und $\gamma = (x, y) \in \Gamma'$.

(i) Ist γ beim Übergang von φ_k zu φ_{k+1} eine Vorwärtskante, so gilt

$$l_k(a, x) + 1 = l_k(a, y).$$

(ii) Ist γ beim Übergang von φ_k zu φ_{k+1} eine Rückwärtskante, so gilt

$$l_k(a, y) + 1 = l_k(a, x).$$

Lemma 3.3.5

Sei $x \in X$ und $0 \leq k < n$. Dann gilt

$$l_{k+1}(a, x) \geq l_k(a, x) \quad \text{und} \quad l_{k+1}(x, b) \geq l_k(x, b).$$

Definition 3.3.6

Eine Kante $\gamma \in \Gamma'$ heisst *Flaschenhals* beim Übergang von φ_k zu φ_{k+1} , falls

$$\varphi_k(\gamma) < c(\gamma) \quad \text{und} \quad \varphi_{k+1}(\gamma) = c(\gamma)$$

oder

$$\varphi_k(\gamma) > 0 \quad \text{und} \quad \varphi_{k+1}(\gamma) = 0.$$

Satz 3.3.7

Sei $0 \leq k < m < n$ und $\gamma \in \Gamma'$ ein *Flaschenhals*, sowohl für den Übergang von φ_k nach φ_{k+1} , als auch für den von φ_m nach φ_{m+1} . Dann gilt

$$l_m(a, b) \geq l_k(a, b) + 2.$$

Satz 3.3.8 (Edmond-Karp)

Führt man den Algorithmus von Ford-Fulkerson in einem Netzwerk $N = ((X, \Gamma), a, b, c)$ mit kürzesten vergrößernden Wegen durch, so wird nach höchstens

$$\frac{1}{2} \cdot \#X \cdot \#\Gamma'$$

Flussvergrößerungen ein Maximalfluss erreicht.

Bemerkung 3.3.9

- (i) In Beispiel 3.1.1 ist $\#X = 5$, $\#\Gamma' = 8$, also ergibt sich die obere Schranke $\frac{1}{2} \cdot 5 \cdot 8 = 20$ für die Laufzeit. Tatsächlich hatten wir nur 4 Schritte benötigt.
- (ii) Sieht man die Vergrößerung des Flusses auf einer Kante als einen Rechenschritt an, so ist der Aufwand für eine einzelne Flussvergrößerung proportional zur Anzahl der Kanten in dem vergrößernden Weg, und man erhält einen Gesamtaufwand der Größenordnung

$$O(\#X \cdot (\#\Gamma')^2).$$

Bemerkung 3.3.10

- Ein weiterer Algorithmus zur Auffindung eines Maximalflusses ist der von *Goldberg-Tarjan*. Er ist der effizienteste bekannte Algorithmus für das Flussproblem.
- Mit $n = \#X$ und $m = \#\Gamma'$ hat man Schranken der Ordnung $O(n \cdot m^2)$ für Ford-Fulkerson (bzw. Edmond-Karp) und $O(n^2 \cdot m)$ für Goldberg-Tarjan.
- In Anwendungen mit $m \sim n^2$ ergibt sich $O(n^5)$ für Ford-Fulkerson und $O(n^4)$ für Goldberg-Tarjan.
- Man kann durch weitere Verbesserungen mit Goldberg-Tarjan $O(n^3)$ erreichen.

Spieltheorie: Einleitung

Was wollen wir unter einem Spiel verstehen?

- Ein Spiel ist ein mathematisches Modell für eine Konfliktsituation zwischen mehreren Beteiligten.
- Die Beteiligten haben Auswahlmöglichkeiten für Handlungen. Diese erbringen einen gewissen Nutzen.
- Wir befassen uns mit strategischen Spielen, bei denen der Ausgang vom Verhalten der Beteiligten abhängt (im Gegensatz zu Glücksspielen).
- Eine Strategie eines Spielers ist eine vollständige Voraus-Festlegung seiner Handlungen für alle denkbaren Spielkonstellationen.
- Ziel der Spieltheorie ist es zum Beispiel Vorhersagen, Erklärungen, Untersuchungen und Anweisungen für Spiele zu liefern

Was für Arten von Spielen gibt es?

- Zahl der Spieler: 2-PS (Zweipersonenspiele) n -PS (n -Personenspiele)
- kooperative und nicht kooperative Spiele
- Spiele mit und ohne Seitenzahlungen (Bestechung)
- Nullsummen- und Nichtnullsummenspiele

Minimax-Strategien

Definition 4.2.1

Ein *Zweipersonen-Nullsummenspiel (2-PNSS)* liegt in Matrixform vor. Die Zeilen stehen für die *Strategien* des Spielers I, die Spalten für die Strategien des Spielers II. Ein Eintrag a_{ij} der Matrix gibt den Gewinn von I und den Verlust von II an für den Fall, dass I die Strategie der Zeile i und II die Strategie der Spalte j spielt.

Die Auswahl von Strategien also Auswahl einer Zeile i und einer Spalte j nennen wir Strategiepaar (i, j) .

Definition 4.2.2

Eine Strategiepaar (i, j) ist im *Gleichgewicht*, falls kein Spieler durch Änderung seiner Strategie einen Auszahlungsvorteil erhalten kann. Die Auszahlung a_{ij} heißt dann Sattelpunkt.

Korollar 4.2.3

Eine Strategiepaar (i, j) einer $m \times n$ -Auszahlungsmatrix $A = (a_{ij})$ ist im Gleichgewicht, genau dann wenn

$$a_{ij} = \min\{a_{ik} \mid 1 \leq k \leq m\}$$

und

$$a_{ij} = \max\{a_{lj} \mid 1 \leq l \leq n\}$$

gilt.

Beispiel 4.2.4

In der folgenden Matrix ist $a_{22} = 2$ ein Sattelpunkt.

$$\begin{pmatrix} 5 & 1 & 3 \\ 3 & 2 & 4 \\ -3 & 0 & 4 \end{pmatrix}$$

Proposition 4.2.5

Sind (i, j) und (i', j') Sattelpunkte in A , dann sind auch (i, j') und (i', j) Sattelpunkte in A und es gilt $a_{ij} = a_{ij'} = a_{i'j} = a_{i'j'}$.

Beispiel 4.2.6

In der folgenden Matrix gibt es keinen Sattelpunkt.

$$\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$$

- Spieler II kann höchstens $3 = \min_j \max_i a_{ij}$ verlieren.
- Spieler I kann mindestens $2 = \max_i \min_j a_{ij}$ gewinnen.

Definition 4.2.7

Bei einem 2-PNSS besitzt Spieler I den *Gewinnsockel*

$$v_I^* := \max_i \min_j \{a_{ij}\}$$

und Spieler II den *Verlustdeckel*

$$v_{II}^* := \min_j \max_i \{a_{ij}\}.$$

Lemma 4.2.8

Es gilt stets

$$v_I^* \leq v_{II}^*.$$

Lemma 4.2.9

Genau dann ist a_{ij} ein Sattelpunkt von A , wenn $v_I^ = v_{II}^* = a_{ij}$ gilt.*

Beispiel 4.2.10

Wir betrachten die Auszahlungsmatrix für das Spiel Stein-Schere-Papier:

$$\begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Es gibt keinen Sattelpunkt, es gilt $v_j^* = -1$ und $v_{ij}^* = 1$.

Gemischte Strategien

Definition 4.3.1

In einem 2-PNSS habe ein Spieler $m \in \mathbb{N}$ Strategien zur Wahl. Eine *gemischte Strategie* ist ein Vektor $x \in \mathbb{R}^m$ mit

$$x \geq 0 \quad \text{und} \quad \sum_{i=1}^m x_i = 1.$$

Lemma 4.3.2

Bei einem 2-PNSS mit Auszahlungsmatrix $A = (a_{ij})$ seien $x \in \mathbb{R}^m$ und $y \in \mathbb{R}^n$ gemischte Strategien für Spieler I bzw. II. Wählen die Spieler die Strategien unabhängig voneinander, so beträgt die erwartete Auszahlung

$$\sum_{i=1}^m \sum_{j=1}^n x_i a_{ij} y_j = x^T A y.$$

Satz 4.3.3

Für 2-PNSS gilt $v_I = v_{II}$. Diese Werte und Strategien für Spieler I und II können mit einem linearen Optimierungsprobleme ermittelt werden.

Bimatrixspiele

- Für ein Spiel mit zwei Spielern habe Spieler i genau $k_i \in \mathbb{N}$ Strategien zur Verfügung.
- Die Auszahlungsfunktion für Spieler i können wir als eine $k_1 \times k_2$ Matrix A_i definieren.
- Spielt Spieler j die Strategie l_j , so erhält Spieler i die Auszahlung $(A_i)_{(l_1, l_2)}$.
- Das Tupel (A_1, A_2) nennen wir *Bimatrixspiel*.

Definition 4.4.1

Ein Paar (x^*, y^*) von gemischten Strategien für ein Bimatrixspiel (A_1, A_2) nennt man im *Gleichgewicht*, falls für alle anderen gemischten Strategien (x, y) gilt:

$$x^T A_1 y^* \leq (x^*)^T A_1 y^* \quad \text{und} \quad (x^*)^T A_2 y \leq (x^*)^T A_2 y^*.$$

Bemerkung 4.4.2

Sind (x^*, y^*) im Gleichgewicht, so kann keiner der Spieler seine erwartete Auszahlung durch eine Änderung seiner Strategie verbessern.

Satz 4.4.3

(Fixpunktsatz von Brouwer) Sei f eine stetige Abbildung von einer nichtleeren, kompakten, konvexen Teilmenge eines endlichdimensionalen Banachraumes in sich selbst. Dann hat f einen Fixpunkt.

Satz 4.4.4

Jedes Bimatrixspiel besitzt ein Strategiepaar im Gleichgewicht.

Beispiel 4.4.5

$$A_1 = \begin{pmatrix} 4 & 0 \\ -1 & 1 \end{pmatrix} \quad A_2 = \begin{pmatrix} 1 & 0 \\ -1 & 4 \end{pmatrix}$$

Gleichgewichtspaare für Spieler 1 und 2 sind:

- $(1, 0)$ und $(1, 0)$ mit Auszahlung $(4, 1)$.
- $(0, 1)$ und $(0, 1)$ mit Auszahlung $(1, 4)$.

Keine Gleichgewichtspaare sind hingegen:

- $(1, 0)$ und $(0, 1)$ mit Auszahlung $(0, 0)$.
- $(0, 1)$ und $(1, 0)$ mit Auszahlung $(-1, -1)$.

Wir betrachten die gemischten Strategien.

- Spieler 1 spiele mit $(p, 1 - p)$ und Spieler 2 mit $(q, 1 - q)$.
- Die erwarteten Auszahlungen für Spieler 1 bzw. 2 sind

$$f_1(p, q) = p(6q - 1) + 1 - 2q, \quad f_2(p, q) = q(6p - 5) + 4(1 - p).$$

- Für $q = \frac{1}{6}$ kann Spieler 1 seine Auszahlung nicht mehr beeinflussen.
Für $p = \frac{5}{6}$ trifft das auf Spieler 2 zu.
- Daher ist

$$(p, 1 - p) = \left(\frac{5}{6}, \frac{1}{6}\right), \quad (q, 1 - q) = \left(\frac{1}{6}, \frac{5}{6}\right)$$

ein Gleichgewichtspaar. Die erwartete Auszahlung beträgt dann $\frac{24}{36}$ für jeden.

- Spielt aber Spieler 2 mit $q > \frac{1}{6}$, so ist $6q - 1 > 0$ und Spieler 2 kann für $p = 1$ seine Auszahlung maximieren. Er wählt also die reine Strategie von oben.
- Spielt Spieler 2 mit $q < \frac{1}{6}$, so ist $6q - 1 < 0$ und Spieler 1 wählt $p = 0$.

Zum Vergleich berechnen wir nun die Minimax-Strategien für beide Spieler.

- Für Spieler 1 verwenden wir Auszahlungsmatrix A_1 und die Strategie $x = (p, 1 - p)$:

$$v_1(x) = \min\{x^T A_1 e_1, x^T A_1 e_2\} = \min\{5p - 1, 1 - p\}$$

also

$$v_1 = \max\{v_1(x) \mid p \in [0, 1]\} = \frac{1}{3} =: p_0.$$

- Analog erhalten wir für Spieler 2 mit Strategie $y = (q, 1 - q)$ und Matrix A_2 :

$$v_2 = \max \{ \min \{ e_1^T A_2 y, e_2^T A_2 y \} \mid q \in [0, 1] \} = \frac{2}{3} =: q_0.$$

- Die Auszahlungen berechnen sich zu

$$f_1(p_0, 1 - p_0) = \frac{2}{3} = f_2(q_0, 1 - q_0).$$

Bemerkung 4.4.6

- Gleichgewichtspunkte sind nicht automatisch Minimax-Strategien und umgekehrt.
- Verschiedene Gleichgewichtspunkte können verschiedene Auszahlungen haben.

Bemerkung 4.4.7

Die Methode im Beweis von Satz 4.4.4 kann man zum Berechnen von Gleichgewichtspunkten verwenden.

Kooperative Spiele

Wir betrachten ein Bimatrixspiel.

- Ein Auszahlungspaar $(x_1, x_2) \in \mathbb{R}^2$ nennen wir *Garantiepunkt*, falls sich Spieler i mindestens x_i sichern kann.
- Mit gemischten Strategien können sich die beiden Spieler mit ihrer Maxmin-Strategie jeweils mindestens v_1 bzw. v_2 an Auszahlung sichern. Daher ist (v_1, v_2) ein Garantiepunkt.

Definition 4.5.1

$$\mathcal{A} := \{(x, A) \mid x \in \mathbb{R}^2, A \subset \mathbb{R}^2 \text{ konvex und kompakt,} \\ \forall a \in A : x \leq a, \exists a' \in A : x < a'\}$$

Bemerkung 4.5.2

- Die Menge \mathcal{A} ist die Menge der Auszahlungskombinationen, die die Spieler für realistisch halten. Ein Element $(x, A) \in \mathcal{A}$ nennen wir *Verhandlungssituation*. Die Menge A selbst *Verhandlungsmenge*.
- Dabei ist $x \in A$ ein Garantiepunkt. Aber es gibt für beide eine echte Verbesserung $a' \in A$.
- Gesucht ist nun eine Verhandlungslösung zu jedem $(x, A) \in \mathcal{A}$. Der Garantiepunkt (v_1, v_2) ist eine solche und kann als *Konfliktlösung* betrachtet werden
- Die Konvexität von A ergibt sich aus der Annahme, dass die Spieler aus einer Verhandlungsmenge durch eine Lotterie neue Verhandlungsmengen erzeugen können bzw. wollen.

Definition 4.5.3

Eine *Verhandlungslösung* auf A ist eine Abbildung

$$\varphi : \mathcal{A} \rightarrow A \subset \mathbb{R}^2,$$

die jedem (x, A) eine Auszahlungspunkt $\varphi(x, A) \in A$ zuordnet.

Man versucht nun vernünftige Anforderung an φ zu stellen, so dass eine Lösung eindeutig festgelegt ist. Die folgenden hat Nash 1950 formuliert.

Axiome 4.5.4 (Nash 1950)

R1 „Individuelle Rationalität“

$$\forall (x, A) \in \mathcal{A} : \varphi(x, A) \geq x.$$

Niemand wird sich mit weniger begnügen, als man ohne Verhandlung erreichen kann.

P1 „Pareto-Optimalität“

$$\varphi(x, A) \in P_W(A) := \{a \in A \mid \nexists y \in A : y > a\}.$$

Es soll keine gleichzeitige echte Verbesserung geben.

S1 „Symmetrie“

Für eine symmetrische Verhandlungssituation (x, A) , das bedeutet

$$x_1 = x_2 \quad \text{und} \quad (a_1, a_2) \in A \Leftrightarrow (a_2, a_1) \in A$$

gilt

$$(\varphi(x, a))_1 = (\varphi(x, a))_2.$$

Symmetrische Verhandlungssituationen ergeben symmetrische Konsequenzen.

T2 „Unabhängigkeit von positiven linearen Transformationen“

Sind $\alpha_1, \alpha_2 > 0$ und $\beta_1, \beta_2 \in \mathbb{R}$ und $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definiert durch

$$T(a_1, a_2) := (\alpha_1 a_1 + \beta_1, \alpha_2 a_2 + \beta_2)$$

dann gilt

$$\forall (x, A) \in \mathcal{A} : \varphi(T(x), T(A)) = T(\varphi(x, A)).$$

Ersetzt ein Spieler seine Verhandlungssituation durch eine Äquivalente, so gewinnt oder verliert er dadurch nichts in der Verhandlungslösung.

U „Unabhängigkeit von irrelevanten Alternativen“ Sind $(x, A), (x, B) \in \mathcal{A}$ mit $\varphi(x, A) \in B \subset A$, dann gilt

$$\varphi(x, B) = \varphi(x, A).$$

Eine anerkannte Lösung soll sich nicht ändern, wenn man einige Verhandlungssituationen, die man vorher nicht betrachtet hat weglässt.

Dazu ein Beispiel. Es sei

$$A := \{(a_1, a_2) \in \mathbb{R}^2 \mid a_1 + a_2 \leq 1\}, \quad x = (0, 0).$$

Nach Axiom S1 und P1 muss die Lösung $(\frac{1}{2}, \frac{1}{2})$ sein. Schränkt man nun A wie folgt ein

$$B := A \cap \{a_1 \leq \frac{1}{2}\}$$

so bleibt $(\frac{1}{2}, \frac{1}{2})$ nach Axiom U die Lösung. Spieler I erhält sein Maximum aber Spieler II könnte mehr bekommen.

Bemerkung 4.5.5

Das sind die Bemerkungen oben in den Axiomen.

Lemma 4.5.6

Sei $(x, A) \in \mathcal{A}$, $A \neq \emptyset$ eine Verhandlungssituation. Dann gibt es genau ein

$$a^* := (a_1^*, a_2^*)^T \in A,$$

das die Funktion

$$f : A \rightarrow \mathbb{R}, \quad f((a_1, a_2)^T) := f(a_1, a_2) := (a_1 - x_1)(a_2 - x_2)$$

maximiert.

Lemma 4.5.7

Vorgelegt sei die Situation aus Lemma (4.5.6) und die Abbildung

$$h(a_1, a_2) := (a_1^* - x_1)a_2 + (a_2^* - x_2)a_1.$$

Dann gilt für alle $(a_1, a_2) \in A$:

$$h(a_1, a_2) \leq h(a_1^*, a_2^*).$$

Satz 4.5.8

Es gibt nur eine Verhandlungslösung, die die Axiome 1 bis 5 erfüllt. Diese Lösung ist der Punkt a^ aus Lemma (4.5.6) und wird Nash-Lösung genannt.*

Beispiel 4.5.9

- Wir betrachten Verhandlungen zwischen einem Unternehmen und Arbeitskräften.
- Das Unternehmen ist durch eine Produktionsfunktion f charakterisiert.
- Die Produktion f hänge nur von der Anzahl a der Arbeitskräfte ab. Sie sei monoton wachsend und konkav.
- Es sei w die Lohnhöhe einer Arbeitskraft. Das Verhandlungsziel ist eine Lohnhöhe und ein Beschäftigungsniveau (w, a) zu bestimmen.
- Bei einer anderen Firma wird mit Lohn w_0 vergütet.
- Der Gewinn der Firma ist $f(a) - wa$. Der Gesamtlohn der Arbeitskräfte beträgt $wa + (A - a)w_0$. Dabei ist A die Anzahl der insgesamt zur Verfügung stehenden Arbeitskräfte für die Firma.
- weiter in der Vorlesung an der Tafel...

Beispiel 4.5.10 (Drohungen)

Wir betrachten ein Bimatrixspiel gegeben durch

$$A = \begin{pmatrix} 4 & 8 \\ 0 & 6 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 8 & 4 \\ -\frac{5}{2} & -2 \end{pmatrix}.$$

- Die Verhandlungsmenge sei

$$k \left(\{ (4, 8)^T, (8, 4)^T, (0, \frac{5}{2})^T, (6, -2)^T \} \right).$$

- Ein Gleichgewichtspunkt reiner Strategien ist $(4, 8)^T$.
- Mit der Minimax-Strategie kann sich Spieler A gemäß seiner Auszahlungsmatrix mindestens 4 sichern, Spieler B hingegen nur -2.

- Vereinbart man $(4, -2)^T$ als Garantiepunkt muss man für die Nash-Lösung auf $y = 12 - x$ und $4 \leq x \leq 8$ die Funktion

$$(x - 4)(y + 2) = (x - 4)(14 - x) =: f(x)$$

maximieren. Das lokale Maximum von f liegt bei $x = 9$. Maximieren von f auf $[4, 8]$ liefert daher die Nashlösung $(8, 4)^T$. Das ist für Spieler A besser als vorher.

- Wenn sich Spieler B auf -2 als Verhandlungsbasis einlässt, dann fährt er aber mit seiner anderen Strategie y_1 , in der er $-\frac{5}{2}$ bekommt, nicht viel schlechter. Das würde aber Spieler A dazu bringen seine erste Strategie x_1 zu spielen, um 4 statt 0 zu erhalten. Daher *droht* Spieler B mit Strategie y_1 .
- Die Drohstrategie für A ist dagegen x_2 , denn dort verringert sich die Auszahlung von B. Spielen beide nun diese gemischte Variante x_2 und y_1 so ergibt sich $(0, -\frac{5}{2})^T$ als Garantiepunkt. Für die Nashlösung muss also die Funktion

$$(x - 0)(y + \frac{5}{2}) = x(14.5 - x) := g(x)$$

maximiert werden. Das ergibt $(7.25, 4.75)^T$ als Nash-Lösung. Für B hat sich das Drohen gelohnt, für A nicht.

Kooperative n-Personenspiele

- Falls bei einem Spiel mit n Spielern Koalitionen erlaubt sind, wird jeder Einzelspieler versuchen, einer für ihn bestmöglichen Koalition beizutreten.
- Die Teilnahme an einer Koalition soll mindestens so viel Auszahlung ergeben, wie man sich alleine sichern könnte.
- Durch Seitenzahlungen können Koalitionäre aus ihrer Koalition herausgelockt werden \implies „Stabilität“

Beispiel 4.6.1

Drei Spieler sollen Koalitionen bilden, bei denen jeweils zwei ein Bündnis eingehen. Dies wird für beide jeweils mit Auszahlung 1 belohnt, der Außenseiter hingegen muss 2 zahlen. Es sind also folgende Auszahlungen möglich:

$$(-2, 1, 1)^T, \quad (1, -2, 1)^T, \quad (1, 1, -2)^T.$$

Falls keine Koalition zustande kommt, erhält jeder nichts also wird $(0, 0, 0)^T$ ausgezahlt.

Es seien nun 2 und 3 verbündet. Jetzt zahlt 1 an 2 einen Betrag von 0.1 gegen das Versprechen mit ihm zu koalieren. Das ergäbe als Auszahlung $(0.9, 1.1, -2)^T$. Beide haben sich verbessert, 3 hat sich verschlechtert. Nun könnte Spieler 3 dem Spieler 2 mehr bieten, zum Beispiel 0.2, das ergäbe $(-2, 1.2, 0.8)^T$ als Auszahlung. ...und so weiter

Die Koalitionen sind „instabil“.

Definition 4.6.2

Sei $N = \{1, \dots, n\}$ die Menge der Spieler. Jede nicht-leere Teilmenge $K \subset N$ heißt eine *Koalition*.

Definition 4.6.3

Die *charakteristische Funktion* $\nu : 2^N \rightarrow \mathbb{R}$ eines n -Personenspiels mit Spielermenge $N = \{1, \dots, n\}$ ist eine Abbildung, die folgende Bedingungen erfüllt: $\nu(\emptyset) = 0$ und

$$\nu(S) + \nu(T) \leq \nu(S \cup T) \quad \text{für alle } S, T \subset N \text{ mit } S \cap T = \emptyset.$$

Bemerkung 4.6.4

Die Subadditivität in Definition 4.6.3 besagt, dass sich eine Koalition mindestens so viel sichern kann, wie sie sich einzeln sichern können.

Gibt es charakteristische Funktionen?

- S_1, \dots, S_n seien Strategiemengen von n Spielern. Die Auszahlungsfunktionen seien $a_k : S_1 \times \dots \times S_n \rightarrow \mathbb{R}$.
- Die Strategiemengen sollen jeweils endlich (nur reine Strategien) oder gemischt (also Simplexes, kompakt reicht auch aus) sein.

Proposition 4.6.5

Für $n \in \mathbb{N}$ sei $N = \{1, \dots, n\}$ die Spielermenge und S_k deren Strategiemengen. Für $K \subset N$ sei S_K die Menge der Strategien der Koalitionäre aus K , also $S_K = \times_{k \in K} S_k$. Wir definieren $\nu : 2^N \rightarrow \mathbb{R}$ durch

$$\nu(K) := \sup_{x \in S_K} \inf_{y \in S_{N-K}} \sum_{k \in K} a_k(x, y).$$

Dann ist ν eine charakteristische Funktion.

Definition 4.6.6

- (i) Ein kooperatives n -PS nennt man *Nullsummenspiel* falls
- a) $\nu(N) = 0$
 - b) $\forall \emptyset \neq K \subset N : \nu(K) + \nu(N - K) = 0.$
- (ii) Ein kooperatives n -PS nennt man *Konstantsummenspiel* falls
- $$\forall \emptyset \neq K \subset N : \nu(K) + \nu(N - K) = \nu(N).$$

Definition 4.6.7

- (i) Ein Spiel (N, ν) heißt *wesentlich*, falls

$$\sum_{i \in N} \nu(i) < \nu(N).$$

- (ii) Ein Spiel (N, ν) heißt *unwesentlich*, falls

$$\sum_{i \in N} \nu(i) = \nu(N).$$

Imputationen

Bemerkung 4.6.8

Unwesentliche Spiele sind aus Sicht von Koalitionen uninteressant, denn schließen sich Spieler zusammen, ist die Subadditivität von ν für sie eine Gleichheit.

Wie sollen die Koalitionäre die Auszahlung aufteilen?

Definition 4.6.9

Ein Vektor $(z_1, \dots, z_n)^T \in \mathbb{R}^n$ heißt *Imputation* oder *Zubilligungsvektor* zu einem n -Personenspiel (N, ν) , falls gilt:

(i) „Individuelle Rationalität“

$$\forall i \in N : z_i \geq \nu(i).$$

(ii) „Kollektive Rationalität“

$$\sum_{i \in N} z_i = \nu(N).$$

Die Menge aller Imputationen heißt *Imputationsraum* $I(\nu)$

Definition 4.6.10

Eine Imputation z *dominiert* eine Imputation w bezüglich einer Koalition K , $K \neq \emptyset$, $K \neq N$, $K \subset N$, $|K| > 1$, falls gilt:

(i) „Überlegenheit“

$$\forall i \in K : z_i > w_i.$$

(ii) „Zulässigkeit“

$$\sum_{i \in K} z_i \leq \nu(K).$$

Wir schreiben $z \xrightarrow{K} w$ und nennen K eine *effektive* Koalition für z .

Bemerkung 4.6.11

In Definition 4.6.10 ergäbe $|K| = 1$ keinen Sinn, denn dann wäre $\nu(1) \geq z_1 > w_1 \geq \nu(1)$. Der Fall $K = N$ kann wegen Definition 4.6.9.(ii) nicht eintreten.

Definition 4.6.12

Eine Imputation z ist *dominationsfähig* gegenüber w , falls es eine effektive Koalition K gibt mit $z \xrightarrow{K} w$. Wir schreiben dann $z \xrightarrow{f} w$.

Definition 4.6.13

Der *Kern* eines kooperativen n -Personenspiels ist

$$\left\{ z \in I(\nu) \mid \nexists w \in I(\nu) : w \xrightarrow{f} z \right\}.$$

Lemma 4.6.14

Es sei (N, ν) ein n -Personenspiel und w eine Imputation. Dann gilt für alle $K \subset N$:

$$\sum_{k \in K} w_k < \nu(K) \iff \exists z \in I(\nu) \text{ mit } z \xrightarrow{K} w. \quad (9)$$

Satz 4.6.15

Für jedes kooperative n -Personenspiel (N, ν) gilt

$$\text{Kern}((N, \nu)) = \left\{ z \in \mathbb{R}^n \mid \forall K \subset N : \sum_{k \in K} z_k \geq \nu(K), \sum_{k=1}^n z_k = \nu(N) \right\}.$$

Der Kern kann leer sein. Wir betrachten folgendes

Beispiel 4.6.16

Drei Spieler wählen eine Zahl aus $\{0, 1\}$. Die Spielermenge sei $N := \{1, 2, 3\}$. Die Auszahlungen $a_i(x, y, z)$ werden in folgender Tabelle festgelegt:

(x, y, z)	a_1	a_2	a_3	$a_1 + a_2$
0,0,0	0	0	0	0
0,0,1	1	1	-2	2
0,1,0	1	-2	1	-1
1,0,0	-2	1	1	-1
1,1,0	1	1	-2	2
0,1,1	-2	1	1	-1
1,0,1	1	-2	1	-1

Es sei $K = \{1, 2\}$. Dann wird diese Koalition K die Strategien 0,0 oder 1,1 spielen. Dies ergibt ein 2-PNSS zwischen der Koalition K und Spieler 3 mit Auszahlungsmatrix

$K \mid 3$	0	1
0,0	0	2
1,1	2	0

Die charakteristische Funktion ν wollen wir als den zu erwartenden Gewinn festlegen. Spielt die Koalition K mit $(p, 1 - p)$ und Spieler 3 mit $(q, 1 - q)$ so ergibt sich als erwartete Auszahlung für K :

$$f(p, q) := (p, 1 - p) \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} q \\ 1 - q \end{pmatrix} = 2q(1 - p) + 2p(1 - q).$$

Der Gradient von f verschwindet bei $p = q = \frac{1}{2}$. Dort ergibt sich die maximale Auszahlung

$$f\left(\frac{1}{2}, q\right) = 1$$

für K unabhängig von der Wahl von q . Wir legen daher $\nu(K) := 1$ fest.

Da die Auszahlungen a_i sich zu Null summieren, definieren wir ν als Nullsummenspiel, also

$$\nu(N) := 0 \quad \text{und} \quad \forall S \subset N : \nu(S) + \nu(N \setminus S) = 0.$$

Wegen der Symmetrie in den Auszahlungen erhalten wir

$$\nu(\{2, 3\}) = \nu(\{1, 3\}) = 1.$$

Aus der Nullsummeneigenschaft ergibt sich

$$\nu(\{1\}) = \nu(\{2\}) = \nu(\{3\}) = -1.$$

Damit ist ν ein wohldefiniertes kooperatives 2 Personennullsummenspiel. Nach Satz 4.6.15 müssen für ein Element (x_1, x_2, x_3) im Kern gelten:

$$x_1 + x_2 + x_3 = 0$$

$$x_1 \geq -1 \quad x_2 \geq -1 \quad x_3 \geq -1$$

$$x_1 + x_2 \geq 1 \quad x_1 + x_3 \geq 1 \quad x_3 + x_2 \geq 1$$

Addition der Ungleichungen ergibt

$$x_1 + x_2 + x_3 \geq \frac{2}{3}.$$

im Widerspruch zu $x_1 + x_2 + x_3 = 0$. Also ist der Kern leer.

Leider ist der Kern für eine ganze Klasse von Spielen leer.

Satz 4.6.17

Für wesentliche Konstantsummenspiele ist der Kern leer.

Der Shapley-Wert

- Es sei (N, ν) ein n -Personenspiel und $\sigma \in S_n$ eine Permutation. Zu einem Spieler $i \in N$ ist die Menge der *Vorgänger* definiert durch

$$P_\sigma(i) := \{r \in N \mid \sigma^{-1}(r) < \sigma^{-1}(i)\}.$$

- Den *Anteil der Auszahlung* von Spieler i bezeichnen wir mit

$$m_i^\sigma := \nu(P_\sigma(i) \cup \{i\}) - \nu(P_\sigma(i)).$$

- Den entsprechenden Auszahlungsvektor schreiben wir als

$$m_\sigma := (m_1^\sigma, \dots, m_n^\sigma).$$

Beispiel 4.7.1

Es sei $N := \{1, 2, 3, 4, 5\}$ und $\sigma \in S_5$ gegeben durch

$$\sigma := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 3 & 5 \end{pmatrix}.$$

Dann hat man $P_\sigma(3) = \{4, 2, 1\}$.

Definition 4.7.2

Der *Shapley-Wert* eines n -Personenspiels (N, ν) ist definiert durch

$$\Phi((N, \nu)) := \frac{1}{n!} \sum_{\sigma \in S_n} m_\sigma.$$

Beispiel 4.7.3

(i) Für $N = \{1, 2\}$ ergibt sich

$$\begin{aligned}\Phi((N, \nu)) &= \frac{1}{2!} \sum_{\sigma \in S_2} m_\sigma = \frac{1}{2}(m_{id} + m_\tau) \\ &= \frac{1}{2}(\nu(N) + \nu(1) - \nu(2), \nu(N) - \nu(1) + \nu(2))\end{aligned}$$

(ii) Für ein additives Spiel gilt $m_\sigma(i) = \nu(i)$ für jedes $i \in N$. Daher folgt

$$\Phi((N, \nu)) = (\nu(1), \dots, \nu(n)).$$

Bemerkung 4.7.4

Für jede Komponente des Shapley-Wertes gilt definitionsgemäß

$$\Phi_i((N, \nu)) = \frac{1}{n!} \sum_{\sigma \in S_n} \nu(P_\sigma(i) \cup \{i\}) - \nu(P_\sigma(i)). \quad (10)$$

Wir erhalten daraus den folgenden

Satz 4.7.5

Für die Komponenten des Shapley-Wertes eines n -Personenspiels (N, ν) gilt

$$\Phi_i((N, \nu)) = \sum_{S \subset N, i \notin S} \frac{|S|!(n-1-|S|)!}{n!} (\nu(S \cup \{i\}) - \nu(S)), \quad 1 \leq i \leq n.$$

Bemerkung 4.7.6

Mit Hilfe von (10) können wir eine stochastische Interpretation des Shapley-Wertes formulieren.

- Betrachte die Permutationen S_n als eine Urne, aus der eine Kugel gezogen wird.
- Die Spieler treten in der Reihenfolge $(\sigma(1), \dots, \sigma(n))$ in einen Raum ein.
- Jedem Spieler wird bei Eintritt sein Anteil an seiner Koalition mit den bereits Eingetretenen ausgezahlt.
- Damit ist Φ_i die erwartete Auszahlung an Spieler i .

Definition 4.7.7

Es seien (N, ν) und (N, μ) jeweils n -Personenspiele.

- (i) Ein Spieler i heißt *Nullspieler*, falls

$$\nu(S \cup \{i\}) = \nu(S)$$

für alle $S \subset N$ gilt.

- (ii) Zwei verschiedene Spieler $i, j \in N$ heißen *symmetrisch*, falls

$$\nu(S \cup \{i\}) = \nu(S \cup \{j\})$$

für alle $S \subset N \setminus \{i, j\}$ gilt.

- (iii) Die *Summe* $(N, \nu + \mu)$ der Spiele (N, ν) und (N, μ) ist definiert durch

$$(\nu + \mu)(S) := \nu(S) + \mu(S).$$

- (iv) Für $\lambda > 0$ ist das n -Personenspiel $(N, \lambda\nu)$ definiert durch

$$(\lambda\nu)(S) := \lambda\nu(S).$$

Zu $N = \{1, \dots, n\}$ sei \mathcal{P}_N die Menge aller n -Personenspiele. Ein Element dieser Menge schreiben wir als $\nu \in \mathcal{P}_N$.

Definition 4.7.8

Eine Abbildung $\Phi : \mathcal{P}_N \rightarrow \mathbb{R}^n$ heißt *Shapley-Funktion*, wenn sie die folgenden Axiome 4.7.9 erfüllt. Der *Shapley-Wert* eines Spiels ist das Bild eines Spiels unter Φ . Der (Shapley-)Wert eines Spielers i ist die i -te Komponente des Shapley-Wertes des Spiels.

Axiome 4.7.9

- 1 „Effizienz“

$$\forall \nu \in \mathcal{P}_N : \sum_{i=1}^n \Phi_i(\nu) = \nu(N).$$

- 2 „Nullspieler-Eigenschaft“: Für jeden Nullspieler $i \in N$ gilt:

$$\forall \nu \in \mathcal{P}_N : \Phi_i(\nu) = 0.$$

- ③ „Symmetrie“: Für je zwei symmetrische Spieler i, j gilt:

$$\forall \nu \in \mathcal{P}_N : \Phi_i(\nu) = \Phi_j(\nu).$$

- ④ „Additivität“

$$\forall \nu, \mu \in \mathcal{P}_N : \Phi(\nu + \mu) = \Phi(\nu) + \Phi(\mu).$$

Bemerkung 4.7.10

Ein Nullspieler trägt zu keiner Koalition etwas bei, daher wird ihm der Wert Null zugeordnet. Symmetrische Spieler tragen zu jeder Koalition dasselbe bei, daher sollen sie denselben Wert erhalten. Werden zwei Spiele hintereinander gespielt, so werden Spieler i die Werte $\Phi_i(\nu)$ und $\Phi_i(\mu)$ zugeordnet. Betrachtet man beide Spiele als nur ein Spiel, ergibt sich das Axiom zur Additivität.

Satz 4.7.11

Der Shapley-Wert aus Definition 4.7.2 erfüllt die Shapley-Axiome 4.7.9.

- Jedes Spiel läßt sich als Linearkombination von Spielen ausdrücken, die eine gewisse Einstimmigkeit beschreiben.
- Dazu definieren wir zu $T \subset N$ das Spiel $u_T \in \mathcal{P}_N$ durch

$$u_T(S) := \begin{cases} 0 & : T \not\subset S \\ 1 & : T \subset S \end{cases} .$$

Lemma 4.7.12

Zu jedem $\nu \in \mathcal{P}_N$ gibt es Zahlen $c_T \in \mathbb{R}$, so dass gilt:

$$\nu = \sum_{T \subset N, T \neq \emptyset} c_T u_T .$$

Lemma 4.7.13

Es seien $\nu \in \mathcal{P}_N$, $c_T \in \mathbb{R}$ die Zahlen aus Lemma 4.7.12 und Ψ eine Shapley-Funktion. Dann gilt

$$\Psi(\nu) = \sum_{c_T > 0} \Psi(c_T u_T) - \sum_{c_T < 0} \Psi(|c_T| u_T). \quad (11)$$

Satz 4.7.14

Der Shapley-Wert aus Definition 4.7.2 ist die einzige Funktion auf \mathcal{P}_N , die die Shapley-Axiome 4.7.9 erfüllt.

Beispiel 4.7.15

Wir betrachten vier Parteien $N = \{1, 2, 3, 4\}$ mit jeweils 5, 20, 25 und 50 Sitzen im Parlament. Es sei $\nu(S) \in \{0, 1\}$ und $\nu(S) = 1$ für alle Koalitionen, die mehr als 50% der Mandate haben. Unter den Gewinnkoalitionen, tragen für die Spieler 1, 2 und 3 jeweils $\{3, 4\}$, $\{2, 4\}$ und $\{1, 4\}$ etwas zum Shapley-Wert bei. Daher gilt

$$\Phi_i(\nu) = \frac{1!(3-1)!}{4!} = \frac{1}{12} \approx 0.083, \quad i \in \{1, 2, 3\}.$$

Die Gewinnkoalitionen für Spieler 4 sind

$$\{1, 4\}, \{2, 3\}, \{3, 4\}, \{1, 2, 4\}, \{2, 3, 4\}, \{1, 3, 4\}, \{1, 2, 3, 4\},$$

also

$$\Phi_4(\nu) = \frac{6 + 6 + 6}{24} = \frac{9}{12} = 0.75$$

Spieler 1 und 4 bzw. 2 und 3 haben mehr bzw. weniger Macht als Mandate.

Shapley-Wert und Betafunktion

- Zu jedem Spieler $i \in N$ eines Spiels (N, ν) sei $x_i \in [0, 1]$ die Bereitschaft an Koalitionen mitzuwirken.
- Für $S \subset N$ ist dann

$$\prod_{i \in S} x_i \prod_{i \in N \setminus S} (1 - x_i)$$

die Wahrscheinlichkeit $\mathbb{P}(S)$, dass die Koalition S zustande kommt.

- Die Zahl

$$\mathbb{E}(\nu) = \sum_{S \subset N} \nu(S) \cdot \mathbb{P}(S) = \sum_{S \subset N} \left(\prod_{i \in S} x_i \prod_{i \in N \setminus S} (1 - x_i) \right) \nu(S)$$

können wir als Erwartungswert des Spiels interpretieren.

- Wir definieren

$$f(x_1, \dots, x_n) := \sum_{S \subset N} \left(\prod_{i \in S} x_i \prod_{i \in N \setminus S} (1 - x_i) \right) \nu(S).$$

Satz 4.7.16

Für jedes $k \in N$ gilt

$$\Phi_k(\nu) = \int_0^1 (\partial_k f)(t, \dots, t) dt.$$

Beispiel 4.7.17

Wir betrachten noch einmal das Beispiel ???. Die Gewinnkoalitionen sind diejenigen in (??), daher berechnen wir

$$\begin{aligned} f(x_1, x_2, x_3, x_4) = & x_1 x_4 (1 - x_2)(1 - x_3) + x_2 x_4 (1 - x_3)(1 - x_1) \\ & + x_3 x_4 (1 - x_2)(1 - x_1) + x_1 x_2 x_4 (1 - x_3) \\ & + x_1 x_3 x_4 (1 - x_2) + x_2 x_3 x_4 (1 - x_1) \end{aligned}$$

Differenzieren nach x_1 und einsetzen von (t, \dots, t) ergibt

$$\partial_1 f(t, \dots, t) = t(1 - t)^2.$$

Daher gilt

$$\Phi_1(\nu) = \int_0^1 t(1 - t)^2 dt = \frac{1}{2}t^2 - \frac{2}{3}t^3 + \frac{1}{4}t^4 \Big|_0^1 = \frac{1}{12}.$$

Nichtlineare Optimierungsprobleme

- Betrachten allgemeine Minimierungsprobleme, bei welchen eine Funktion $f : S \rightarrow \mathbb{R}$ auf $S \subset \mathbb{R}^n$ minimiert werden soll.
- S sei durch *Gleichungs-* und *Ungleichungsrestriktionen* beschrieben.

Definition 5.1.1

Sei $X \subset \mathbb{R}^n$ offen und seien $f, g_1, \dots, g_m, h_1, \dots, h_p : X \rightarrow \mathbb{R}$ gegebene Funktionen. Das Problem, f zu minimieren unter den Restriktionen

$$g_i \leq 0, \quad i = 1, \dots, m, \quad \text{und} \quad h_j = 0, \quad j = 1, \dots, p,$$

nennen wir (allgemeines) *nichtlineares Optimierungsproblem* (MP). Die Menge

$$S := \{x \in X : g(x) \leq 0, h(x) = 0\}$$

heisst *zulässiger Bereich*, hierbei ist

$$g := (g_1, \dots, g_m)^T, \quad \text{und} \quad h := (h_1, \dots, h_p)^T.$$

Soll f nur unter den Restriktionen

$$g_i \leq 0, \quad i = 1, \dots, m,$$

minimiert werden, sagen wir, dass (MP) sei *vom Ungleichungstyp* (MP_{\leq}), in diesem Fall ist

$$S := \{x \in X : g(x) \leq 0\}$$

der *zulässige Bereich*.

Bemerkung 5.1.2

Sind in (MP_{\leq}) alle g_i affin-linear, dann ist S ein Polyeder.

Wir verwenden für (MP) auch die Kurznotation

$$\begin{aligned} f &\stackrel{!}{=} \min, \\ g &\leq 0, \\ h &= 0. \end{aligned}$$

Beispiel 5.1.3

Sei $n = 2$ und $X = \mathbb{R}^2$. Wir suchen das Minimum von

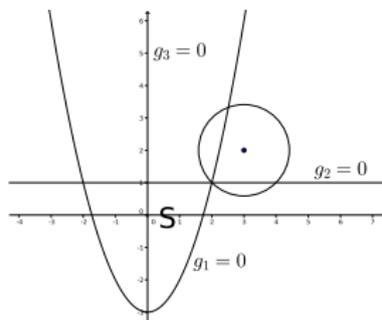
$$f(x, y) = (x - 3)^2 + (y - 2)^2$$

unter den Restriktionen

$$x^2 - y - 3 \leq 0, \quad y - 1 \leq 0, \quad -x \leq 0.$$

Hier sind also $g_1(x, y) = x^2 - y - 3$, $g_2(x, y) = y - 1$ und $g_3(x) = -x$, demnach

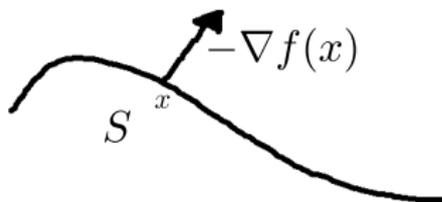
$$S = \{(x, y) \in \mathbb{R}^2 : x^2 - 3 \leq y, y \leq 1, x \geq 0\}.$$



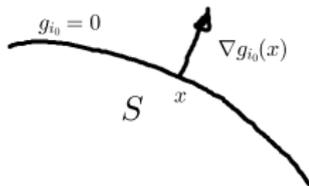
Bemerkung 5.1.4

Wir machen ein paar heuristische Vorüberlegungen, dazu nehmen wir an, die betreffenden Funktionen seien in den betrachteten Punkten differenzierbar.

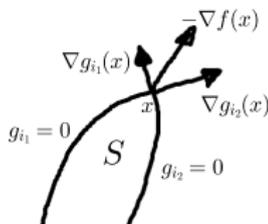
- (i) $\nabla f(x) = 0$ nur für innere Punkte x von S notwendige Bedingung für Minimalstellen $\rightarrow \partial S$?
- (ii) Ist $x \in \partial S$ Minimalstelle für f auf S , so muss $-\nabla f(x)$ 'nach aussen zeigen': denn hätte $-\nabla f(x)$ einen positiven Anteil in eine Richtung 'in S hinein' oder 'entlang ∂S ', so könnte man in diese Richtung gehen und einen kleineren Wert von f finden.



- (iii) Ein Randpunkt x sei gegeben durch *eine* straffe Nebenbedingung $g_{i_0}(x) = 0$. Ist x Minimalstelle, dann zeigen $\nabla g_{i_0}(x)$ und $-\nabla f(x)$ in dieselbe Richtung, denn: Der Gradient $\nabla g_{i_0}(x)$ einer Funktion g_{i_0} steht senkrecht auf ihrer Niveaumenge, welche x enthält. Dass $\nabla g_{i_0}(x)$ dann nur nach aussen zeigen kann, folgt aus der Definition von S . Wegen (ii) muss dann auch $-\nabla f(x)$ in diese Richtung zeigen.



Sei nun x ein Randpunkt, der *zwei* straffe Nebenbedingungen $g_{i_1}(x) = g_{i_2}(x) = 0$ erfüllt. Ist x Minimalstelle, dann muss $-\nabla f(x)$ in dem konvexen Kegel liegen, der von $\nabla g_{i_1}(x)$ und $\nabla g_{i_2}(x)$ aufgespannt wird:



Definition 5.1.5

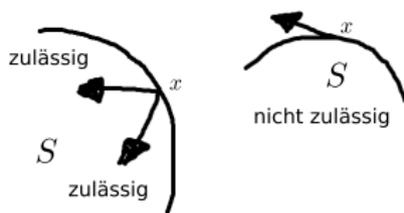
Sei $\emptyset \neq S \subset \mathbb{R}^n$ und $x \in S$. Ein Vektor $d \in \mathbb{R}^n$ heisst *zulässige Richtung* (bzgl. S) in x , falls es ein $\delta > 0$ gibt, sodass

$$[x, x + \delta d] \subset S.$$

Die Menge

$$D_S(x) := \{d : d \text{ zulässige Richtung bzgl. } S \text{ in } x\}$$

heisst der *Kegel der zulässigen Richtungen* in x .



Bemerkung 5.1.6

- (i) Offensichtlich ist $D_S(x)$ ein Kegel: Ist $d \in D_S(x)$ und $\lambda \geq 0$, dann ist $\lambda d \in D_S(x)$.
- (ii) Ist $x \in S$ ein innerer Punkt von S , so gilt $D_S(x) = \mathbb{R}^n$.

Definition 5.1.7

Sei $X \subseteq \mathbb{R}^n$ offen, $x \in X$ und $f : X \rightarrow \mathbb{R}$ differenzierbar in x . Ein Vektor $d \in \mathbb{R}^n$ heisst *Abstiegsrichtung* für f in x , falls es ein $\delta > 0$ gibt mit

$$f(x + \lambda d) < f(x) \quad \text{für alle } \lambda \in (0, \delta).$$

Lemma 5.1.8

Sei $X \subseteq \mathbb{R}^n$ offen, $x \in X$ und $f : X \rightarrow \mathbb{R}$ differenzierbar in x . Erfüllt $d \in \mathbb{R}^n$ die Ungleichung

$$\nabla f(x) \cdot d < 0,$$

so ist d eine *Abstiegsrichtung* für f in x .

Definition 5.1.9

Ist $X \subseteq \mathbb{R}^n$ offen und $f : X \rightarrow \mathbb{R}$ differenzierbar in $x \in X$, so bezeichnen wir mit

$$F_f(x) := \{d \in \mathbb{R}^n : \nabla f(x) \cdot d < 0\}$$

den *Kegel der Abstiegsrichtungen* von f in x .

Bemerkung 5.1.10

Strenggenommen ist $F_f(x)$ selbst noch kein Kegel, sondern $F_f(x) \cup \{0\}$.

Proposition 5.1.11

Sei $X \subset \mathbb{R}^n$ offen, $S \subset X$ und $f : X \rightarrow \mathbb{R}$. Besitzt f in $x \in S$ ein lokales Minimum in S und ist f in x differenzierbar, so gilt

$$D_S(x) \cap F_f(x) = \emptyset.$$

Bemerkung 5.1.12

Proposition 5.1.11 macht insbesondere Bemerkung 5.1.4 (ii) präzise.

Der Ungleichungstyp (MP_{\leq})

Wir betrachten nun (MP_{\leq}),

$$f \stackrel{!}{=} \min, \\ g_i \leq 0, \quad i = 1, \dots, m.$$

Dann ist $D_S(x)$ selbst schlecht handhabbar, man erhält aber Aussagen zu $D_S(x)$ mittels der Restriktionsfunktionen g_i . Für $x \in S = \bigcap_{i=1}^m \{g_i \leq 0\}$ sei

$$I(x) := \{i \in \{1, \dots, m\} : g_i(x) = 0\}$$

die Menge der Indizes der in x *straffen* Restriktionen.

Bemerkung 5.1.13

Im Spezialfall, dass alle g_i affin-linear sind, sammelt man hier die Indizes der Restriktionshyperebenen, in welchen x liegt.

Lemma 5.1.14

Sei (MP_{\leq}) gegeben und $x \in S$. Weiter sei

$$g_i \text{ in } x \begin{cases} \text{differenzierbar, falls } i \in I(x) \\ \text{stetig, falls } i \notin I(x). \end{cases}$$

Dann gilt für

$$G_S(x) := \{d \in \mathbb{R}^n : \nabla g_i(x) \cdot d < 0 \text{ für alle } i \in I(x)\}$$

die Inklusion

$$G_S(x) \subset D_S(x).$$

Zusammen mit Proposition 5.1.11 ergibt sich sofort folgendes Resultat:

Satz 5.1.15

Gegeben sei (MP_{\leq}) und $x \in S$. Ferner sei

$$g_i \text{ in } x \begin{cases} \text{differenzierbar, falls } i \in I(x) \\ \text{stetig, falls } i \notin I(x). \end{cases}$$

Falls f in x ein lokales Minimum in S hat, so ist

$$G_S(x) \cap F_f(x) = \emptyset.$$

Diese Bedingung lässt sich in eine algebraische Aussage umformen, man nennt sie die *Fritz-John-Bedingung*:

Satz 5.1.16

Seien in (MP_{\leq}) mit $x \in S$ die Funktionen g_i wie in Satz 5.1.15. Sei f differenzierbar in x und sei x eine lokale Minimalstelle für f in S . Dann gibt es Konstanten $\mu_0 \geq 0$ und $\mu_i \geq 0$, $i \in I(x)$, die nicht sämtlich null sind, sodass die folgende Fritz-John-Bedingung gilt:

$$\mu_0 \nabla f(x) + \sum_{i \in I(x)} \mu_i \nabla g_i(x) = 0. \quad (\text{FJB})$$

Bemerkung 5.1.17

Ist $\mu_0 > 0$, so sagt die (FJB), dass $-\nabla f(x)$ in dem konvexen Kegel liegt, der von den Gradienten $\nabla g_i(x)$ mit $i \in I(x)$ aufgespannt wird. Das macht insbesondere Bemerkung 5.1.4 (iii) präzise.

Für den Beweis des Satzes 5.1.16 nutzen wir folgenden Alternativsatz von *Gordan*:

Satz 5.1.18

Sei A eine reelle $(m \times n)$ -Matrix. Dann gilt genau eine der folgenden Alternativen:

- (i) Es gibt ein $x \in \mathbb{R}^n$ mit $Ax > 0$.
- (ii) Es gibt ein $y \in \mathbb{R}^m$ mit $A^T y = 0$, $y \geq 0$ und $y \neq 0$.

Korollar 5.1.19

Sei (MP_{\leq}) gegeben und $x \in S$ eine lokale Minimalstelle für f in S . Sind f und alle g_i , $i = 1, \dots, m$, differenzierbar in x , so gibt es $\mu_0, \mu_1, \dots, \mu_m \geq 0$, die nicht alle gleich null sind, mit

$$\mu_0 \nabla f(x) + \sum_{i=1}^m \mu_i \nabla g_i(x) = 0 \quad \text{und} \quad \mu_i g_i(x) = 0, \quad i = 1, \dots, m.$$

Bemerkung 5.1.20

- (i) Die Fritz-John-Bedingung (FJB) ist also eine *notwendige* Bedingung dafür, dass in x ein *lokales* Minimum für f auf S vorliegt, und somit insbesondere dafür, dass ein globales Minimum für f auf S vorliegt. Für 'pseudokonvexe' f gibt es auch hinreichende Bedingungen, dass sich in x ein globales Minimum für f auf S ergibt.
- (ii) Die skalaren Faktoren μ_j heissen auch *Lagrange-Multiplikatoren*, die Gleichungen $\mu_j g_j(x) = 0$ *Bedingungen des komplementären Schlupfes*.
- (iii) Ist in Satz 5.1.16 in der lokalen Minimalstelle x keine der Bedingungen straff gilt also $I(x) = \emptyset$, so liegt x im Inneren von S , und (FJB) reduziert sich wie zu erwarten auf

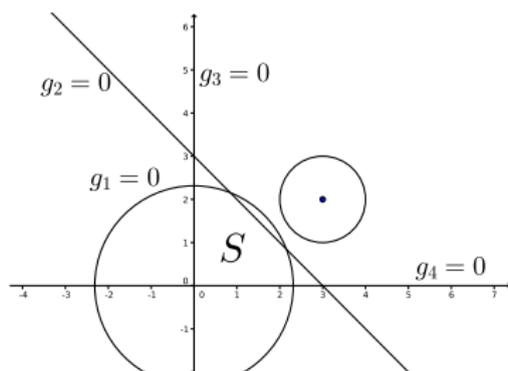
$$\nabla f(x) = 0.$$

In diesem Fall ist x sogar lokale Minimalstelle für f auf einer Umgebung von S .

Beispiel 5.1.21

Sei $n = 2$, $X = \mathbb{R}^2$. Wir wollen $(x, y) \mapsto (x - 3)^2 + (y - 2)^2$ minimieren unter

$$x^2 + y^2 \leq 5, \quad x + y \leq 3, \quad x, y \geq 0.$$



Also haben wir wieder $f(x, y) = (x - 3)^2 + (y - 2)^2$, nun mit

$$g_1(x, y) = x^2 + y^2 - 5, \quad g_2(x, y) = x + y - 3, \quad g_3(x, y) = -x, \quad g_4(x, y) = -y.$$

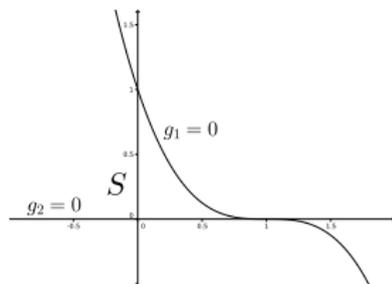
Beispiel weiter an der Tafel...

Beispiel 5.1.22

Sei $n = 2$, $X = \mathbb{R}^2$. Wir wollen $f(x, y) = -x$ minimieren unter den Bedingungen

$$g_1(x, y) := y - (1 - x)^3 \leq 0, \quad g_2(x, y) := -y \leq 0.$$

Hier ist offensichtlich $w = (1, 0)$ eine Minimalstelle von f auf S , und man hat $f(w) = -1$. In w sind sowohl g_1 als auch g_2 straff. Des Weiteren hat man



$$\nabla f(x, y) = (-1, 0), \quad \nabla g_1(x, y) = (3(1 - x)^2, 1), \quad \nabla g_2(x, y) = (0, -1)$$

und insbesondere $\nabla g_1(w) = (0, 1)$ und $\nabla g_2(w) = (0, -1)$. Somit ist die Gleichung (FJB) hier für beliebige $\mu_1 = \mu_2 > 0$ erfüllt, falls $\mu_0 = 0$. Letzteres bedeutet aber, dass keine Information über f in die Gleichung in (FJB) eingeht.

KKT-Bedingung mit LI-Restriktion

Satz 5.1.23

Sei $x \in S$ ein Punkt, in dem g_i für $i \in I(x)$ differenzierbar und für $i \notin I(x)$ stetig ist, und in dem f differenzierbar ist. Weiter gelte

Die Gradienten $\nabla g_i(x)$, $i \in I(x)$, sind linear unabhängig. (LICQ)

Ist x ein lokales Minimum für f in S , so gibt es $\mu_i \geq 0$, $i \in I(x)$, die die folgende KKT-Bedingung erfüllen:

$$\nabla f(x) + \sum_{i \in I(x)} \mu_i \nabla g_i(x) = 0. \quad (\text{KKT})$$

Falls alle g_i , $i = 1, \dots, m$ in x differenzierbar sind, ist dies äquivalent zu

$$\nabla f(x) + \sum_{i=1}^m \mu_i \nabla g_i(x) = 0, \quad \mu_i \geq 0, \quad \mu_i g_i(x) = 0, \quad i = 1, \dots, m.$$

- Mit $\mu = (\mu_1, \dots, \mu_m)^T$ und $g = (g_1, \dots, g_m)^T$ kann man die letzte Identität im Satz auch wie folgt schreiben:

$$\nabla f(x) + \mu^T \cdot \nabla g(x) = 0, \quad \mu \geq 0, \quad \mu^T \cdot g(x) = 0.$$

- Das Kürzel LICQ steht für *linear independence constraint qualification*.

KKT-Bedingung mit Abedi-Restriktion

Definition 5.1.24

Sei $S \subset \mathbb{R}^n$. Ein Vektor $d \in \mathbb{R}^n$ heisst *Tangentialrichtung* für S in $x \in S$, wenn es Folgen $(x_k)_{k=1}^{\infty} \subset S$ und $(\lambda_k)_{k=1}^{\infty} \subset (0, +\infty)$ gibt, sodass

$$\lim_{k \rightarrow \infty} x_k = x, \quad \lim_{k \rightarrow \infty} \lambda_k = 0 \quad \text{und} \quad \lim_{k \rightarrow \infty} \frac{1}{\lambda_k} (x_k - x) = d.$$

Die Menge

$$T_S(x) := \{d \in \mathbb{R}^n : d \text{ Tangentialrichtung für } S \text{ in } x\}$$

aller Tangentialrichtungen für S in x heisst *Tangentialkegel* für S in x .

Bemerkung 5.1.25

- (i) Es gilt $D_S(x) \subset T_S(x)$ für alle $x \in S$:
- (ii) Man kann zeigen, dass für konvexe S die Identität $\overline{D_S(x)} = T_S(x)$ gilt für alle $x \in S$.

Lemma 5.1.26

Sei $X \subset \mathbb{R}^n$ offen, $S \subset X$ und $f : X \rightarrow \mathbb{R}$. Ist x ein lokales Minimum für f auf S und ist f differenzierbar in x , so gilt

$$T_S(x) \cap F_f(x) = \emptyset.$$

Satz 5.1.27

Sei in (MP_{\leq}) $x \in S$, alle g_i , $i \in I(x)$ differenzierbar in x und

$$T_S(x) = G'_S(x), \quad (\text{ACQ})$$

wobei

$$G'_S(x) := \{d \in \mathbb{R}^n : \nabla g_i(x) \cdot d \leq 0 \text{ für alle } i \in I(x)\}.$$

Ist x ein lokales Minimum für f in S , so gibt es $\mu_i \geq 0$, $i \in I(x)$ mit

$$\nabla f(x) + \sum_{i \in I(x)} \mu_i \nabla g_i(x) = 0.$$

Bemerkung 5.1.28

- (i) Man sieht leicht, dass $T_S(x) \subset G'_S(x)$ gilt für alle $x \in S$. Es gilt also

$$G_S(x) \subset D_S(x) \subset T_S(x) \subset G'_S(x)$$

und daher

$$\overline{G_S(x)} \subset \overline{D_S(x)} \subset \underbrace{\overline{T_S(x)}}_{=T_S(x)} \subset \underbrace{\overline{G'_S(x)}}_{=G'_S(x)},$$

Folglich gilt in der Inklusionskette überall Gleichheit, falls

$$\overline{G_S(x)} = G'_S(x). \quad (\text{Cottle-CQ})$$

In diesem Falle gilt auch (ACQ).

- (ii) Man kann zeigen, dass die Bedingung (LICQ) die Bedingung (ACQ) impliziert.

Optima Konvexer Funktionen

Definition 5.2.1

Sei $X \subset \mathbb{R}^n$ konvex. Eine Funktion $f : X \rightarrow \mathbb{R}$ heisst (*strikt*) *konvex*, falls

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad x, y \in X, \quad \lambda \in [0, 1],$$

gilt. Die Funktion f heisst (*strikt*) *konkav*, falls $-f$ (*strikt*) konvex ist.

Beispiel 5.2.2

Affin-lineare Funktionen der Form $f(x) = a^T x + b$ mit $a \in \mathbb{R}^n$ und $b \in \mathbb{R}$ sind gleichzeitig konvex und konkav aber weder strikt konvex, noch strikt konkav, denn

$$\begin{aligned} f(\lambda x + (1 - \lambda)y) &= a^T (\lambda x + (1 - \lambda)y) + b \\ &= \lambda(a^T x + b) + (1 - \lambda)(a^T y + b) \\ &= \lambda f(x) + (1 - \lambda)f(y). \end{aligned}$$

Bemerkung 5.2.3

Ist $f : S \rightarrow \mathbb{R}$ konvex, so ist f stetig auf dem Inneren von S . Unstetigkeitsstellen können aber in ∂S enthalten sein.

Satz 5.2.4

Sei $S \subset \mathbb{R}^n$ konvex und $f : S \rightarrow \mathbb{R}$. Ist S zusätzlich offen und f differenzierbar auf S , so gilt: Die Funktion f ist genau dann konvex, wenn

$$f(x) \geq f(y) + \nabla f(y) \cdot (x - y)$$

für alle $x, y \in S$.

Satz 5.2.5

Sei $X \subset \mathbb{R}^n$ offen und konvex und sei $f : X \rightarrow \mathbb{R}$ differenzierbar.

(i) Die Funktion f ist genau dann konvex, wenn für alle $x, y \in X$ die Ungleichung

$$(\nabla f(y) - \nabla f(x)) \cdot (y - x) \geq 0$$

gilt.

(ii) Ist $f : X \rightarrow \mathbb{R}$ zweimal differenzierbar, dann gilt: Die Funktion f ist genau dann konvex, wenn ihre Hesse-Matrix $H_f(y)$ ist in allen Punkten $y \in X$ positiv semidefinit ist.

pseudokonvex

Definition 5.2.6

Sei $X \subset \mathbb{R}^n$ offen und $f : X \rightarrow \mathbb{R}$. Die Funktion f heisst *pseudokonvex* in $y \in X$, falls f in y differenzierbar ist und für alle $x \in X$ die Implikation

$$\nabla f(y) \cdot (x - y) \geq 0 \quad \Rightarrow \quad f(x) \geq f(y)$$

gilt.

Wir werden in Kürze eine hinreichende Bedingung für das Vorliegen einer (globalen) Minimalstelle für pseudokonvexe Zielfunktionen f beweisen.

Bemerkung 5.2.7

- (i) Sei $S \subset \mathbb{R}^n$ konvex und offen. Dann ist jede konvexe Funktion $f : S \rightarrow \mathbb{R}$, die in $y \in S$ differenzierbar ist, auch pseudokonvex in y : Nach Satz 5.2.4 gilt

$$f(x) \geq f(y) + \underbrace{\nabla f(y) \cdot (x - y)}_{\geq 0} \geq f(y).$$

- (ii) Ein Beispiel für eine Funktion, die pseudokonvex ist auf \mathbb{R} , aber nicht konvex, ist

$$f(x) = -\frac{1}{1+x^2}, \quad x \in \mathbb{R}.$$

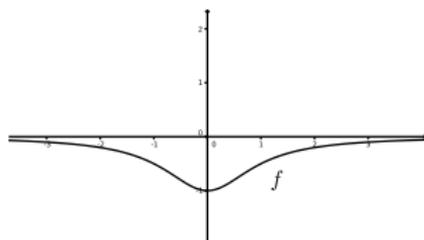
Die Funktion ist auf \mathbb{R} zweimal stetig differenzierbar. Man hat

$$f'(x) = \frac{2x}{(1+x^2)^2}, \quad x \in \mathbb{R}.$$

Ist $y = 0$, so ist $f(x) \geq -1 = f(y)$ für alle $x \in \mathbb{R}$. Ist $y > 0$, so impliziert $f'(y)(x - y) \geq 0$, dass $0 < y \leq x$ gilt und somit $f(x) \geq f(y)$. Analog für $y < 0$. Also ist f pseudokonvex in jedem Punkt $y \in \mathbb{R}$. Andererseits hat man

$$f''(x) = \frac{2(1-3x^2)}{(1+x^2)^3}, \quad x \in \mathbb{R},$$

und das ist nichtnegativ für $|x| \leq 1/\sqrt{3}$, aber negativ für $|x| > 1/\sqrt{3}$. Also ist f nach Satz 5.2.5 (ii) nicht konvex auf \mathbb{R} .



quasikonvex

Definition 5.2.8

Sei $X \subset \mathbb{R}^n$ konvex. Eine Funktion $f : X \rightarrow \mathbb{R}$ heisst *quasikonvex*, falls

$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}, \quad x, y \in X, \quad 0 \leq \lambda \leq 1.$$

Bemerkung 5.2.9

- (i) Es gilt also: Ist f konvex, so ist f auch quasikonvex.
- (ii) Die Funktionen $f(x) = x^3$ und $f(x) = \sqrt{|x|}$ sind quasikonvex auf \mathbb{R} , aber nicht konvex.

Wir betrachten wieder (MP_{\leq}) , d.h.

$$\begin{aligned} f &\stackrel{!}{=} \min, \\ g_i &\leq 0, \quad i = 1, \dots, m, \end{aligned}$$

wobei $f, g_i : X \rightarrow \mathbb{R}$ sind.

Hinreichende KKT-Bedingung

Satz 5.2.10

Es seien $X \subset \mathbb{R}^n$ offen und konvex, $f, g_i : X \rightarrow \mathbb{R}$ und $x \in S := \bigcap_{i=1}^m \{g_i \leq 0\}$, so dass gilt:

- f ist in x pseudokonvex und
- für alle $i \in I(x)$ ist g_i quasikonvex und in x differenzierbar.

Ist in x die KKT-Bedingung

$$\nabla f(x) + \sum_{i \in I(x)} \mu_i \nabla g_i(x) = 0 \quad \text{für geeignete } \mu_i \geq 0$$

erfüllt, so ist x ein globales Minimum für f auf S .