Mathematische Optimierung

Geschrieben von
Jan Pöschko
auf Grundlage der Vorlesung von
Bettina Klinz

TU Graz
Sommersemester 2007

Stand: 27. Oktober 2009

Inhaltsverzeichnis

Einleitung

Extremwertaufgaben

Zielfunktion objective function

Nebenbedingung constraint ($s.t.$ = subject to = unter der Bedingung)

Beispiele

1. *nicht linear:*

$$\min x^2 + 4y^2$$

s.t.

$$x + 3y \leq 5$$
$$x - 2y \geq 1$$

2. *linear:*

$$\min x + 4y$$

s.t.

$$x + 3y \leq 5$$
$$x - 2y \geq 1$$

Wichtige Klassen von Optimierungsproblemen

Lineare Optimierung

- lineare Zielfunktion
- lineare Nebenbedingungen

Allgemeine Form: Gegeben $m \times n$ Matrix A, $a_{ij} \in \mathbb{R}$, Vektor $b \in \mathbb{R}^m$, Vektor $c \in \mathbb{R}^n$.
Lineares Programm:

$$\max c^t x = c_1 x_1 + \cdots + c_n x_n$$

s.t.

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq b_m$$

(lineare Nebenbedingung).

Beispiel (Ernährungsproblem lt. Stigler) *Gegeben: Menge von Nahrungsmitteln, Preis pro Einheit, Einschränkungen eines möglichen Nahrungsplans.*
Ziel: kostengünstiger Ernährungsplan, der alle Bedingungen erfüllt.

- *Entscheidungsvariablen x_i: Anzahl an Portionen pro Nahrungsmittel i*

- *Zielfunktion:* $\min c_1x_1 + \cdots + c_nx_n$ *(c_i: Kosten pro Einheit von i)*
- *Typische Nebenbedingungen:*
  - *Portionsbeschränkung: z.B.* $0 \leq x_1 \leq 4$
  - *Kalorien: z.B.* $110x_1 + 250x_2 + \cdots \geq 1600$

Ganzzahlige Optimierung (integer programming)

Entscheidungsvariablen $\in \mathbb{Z}$.
Spezialfall: ganzzahlige lineare Optimierung ($x_i \in \mathbb{Z}$).

Beispiel

$$\min x_1 + 4x_2 + 5x_3$$

s.t.

$$\begin{aligned} x_1 - 2x_2 + x_3 &\geq 4 \\ x_1 + x_2 + 3x_3 &\leq 5 \\ x_1, x_2, x_3 &\in \mathbb{Z} \end{aligned}$$

Im Allgemeinen schwieriger als lineare Optimierung (oft NP-schwer).
Nichtlineare ganzzahlige Optimierungsprobleme im Allgemeinen hoffnungslos.

Kombinatorische Optimierung

(siehe LV *Kombinatorische Optimierung*)
Typ: endliche (diskrete) Menge von zulässigen Lösungen.

Beispiele *Kürzestes-Weg-Problem, Rundreiseprobleme.*

Nichtlineare Optimierung

Optimalitätsbegriff: *lokales Optimum* (im Gegensatz zum *globalen Optimum*: eigenes Feld, im Allgemeinen schwerer).
Spezielle nichtlineare Optimierung:

- Quadratische Optimierung: quadratische Zielfunktion mit linearen Nebenbedingungen
- Konvexe Optimierung: $\min f(x)$ s.t. $g(x) \leq 0$ (f, g konvex).
  Angenehme Eigenschaft: lokales Optimum = globales Optimum.

Stochastische Optimierung

Zufallsvariablen spielen eine Rolle.

Teil I

Lineare Optimierung

1 Grundlegende Definitionen und Überlegungen

Definition (Lineares Programm in Standardform) *Gegeben sei eine Matrix* $A \in \mathbb{R}^{m \times n}$ *sowie Vektoren* $b \in \mathbb{R}^m, c \in \mathbb{R}^n$*. Das Problem*

$$\max\ c^t x \tag{1.1}$$

s.t.

$$Ax \leq b \tag{1.2}$$
$$x \geq 0 \tag{1.3}$$

heißt lineares Programm in Standardform.
Die affin-lineare Funktion $c^t x$ *in* (1.1) *heißt* Zielfunktion*.*
Die Bedingungen (1.3) *heißen* Vorzeichenbedingungen *oder* Nicht-Negativitätsbedingungen*.*

Komponentenweise Schreibweise:

$$\max c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$$

s.t.

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

m Restriktionen (exkl. Vorzeichenbedingungen), n Entscheidungsvariablen.

Bemerkungen

- c *wird auch* Zielfunktionsvektor *genannt (*c_j*:* j-ter Zielfunktionskoeffizient*).*
- b *wird auch* rechter Seitenvektor *genannt.*
- A *heißt* Restriktionsmatrix*.*

Beispiel

$$\max 5x_1 + 4x_2 + 3x_3$$

s.t.

$$\begin{aligned} 2x_1 + 2x_2 + x_3 &\leq 5 \\ 4x_1 + x_2 + 2x_3 &\leq 11 \\ 3x_1 + 4x_2 + 2x_3 &\leq 8. \end{aligned}$$

Hier wäre

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 4 & 1 & 2 \\ 3 & 4 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 11 \\ 8 \end{pmatrix}, \quad c = \begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix}.$$

Bemerkung *Jedes lineare Programm lässt sich in die Standardform überführen:*

1. $\min c^t x$ *ist äquivalent zu* $\max -c^t x$.
2. *Ungleichungen vom Typ „$\geq$“:*
$$\sum_{j=1}^{n} a_{ij}x_j \geq b_i \quad \longrightarrow \quad \sum_{j=1}^{n} -a_{ij}x_j \leq -b_i.$$
3. *Gleichungsrestriktionen* $\sum_{j=1}^{n} a_{ij}x_j = b_i$ *sind äquivalent zu*
$$\sum_{j=1}^{n} a_{ij}x_j \geq b_i \quad \text{und} \quad \sum_{j=1}^{n} a_{ij}x_j \leq b_i,$$
also
$$\sum_{j=1}^{n} -a_{ij}x_j \leq -b_i \quad \text{und} \quad \sum_{j=1}^{n} a_{ij}x_j \leq b_i.$$
4. *Keine Vorzeichenbedingungen für Variable* x_j*: Ersetze* x_j *durch* $x_j^+ - x_j^-$ *mit* $x_j^+, x_j^- \geq 0$.

Beispiel

$$\max 4x + 3y$$

s.t.

$$\begin{aligned} x + y &\leq 7 \\ x &\geq 0. \end{aligned}$$

$y = y^+ - y^-$ *liefert*

$$\max 4x + 3y^+ - 3y^-$$

s.t.

$$\begin{aligned} x + y^+ - y^- &\leq 7 \\ x, y^+ y^- &\geq 0. \end{aligned}$$

Bemerkung *Andere Umformungen bzw. Standardformen sind möglich. Zum Beispiel wird aus* $\sum_{j=1}^{n} a_{ij}x_j \leq b_i$ *durch Einführen der Variablen* z_i *eine Gleichung und Vorzeichenbedingung:*

$$\begin{aligned} \sum_{j=1}^{n} a_{ij}x_{ij} + z_i &= b_i \\ z_i &\geq 0. \end{aligned}$$

Bemerkung *Unter einem* allgemeinen linearen Programm *versteht man ein lineares Programm der Form*

$$\max c^t x \quad \text{oder} \quad \min c^t x$$

(lineare Zielfunktion) s.t.

$$\sum_{j=1}^{n} a_{ij}x_j \geq b_i$$

oder

$$\sum_{j=1}^{n} a_{ij}x_j \leq b_i$$

oder

$$\sum_{j=1}^{n} a_{ij}x_j = b_i$$

(lineare Nebenbedingungen).

Beispiel

$$\max 5x_1 + 4x_2 + 3x_3$$

s.t.

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &\leq 5 \\ 4x_1 + x_2 + 2x_3 &\leq 11 \\ 3x_1 + 4x_2 + 2x_3 &\leq 8 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

Idee: Umformen in Gleichungen, also

$$\begin{aligned} 2x_1 + 3x_2 + x_3 + x_4 &= 5 \\ 4x_1 + x_2 + 2x_3 + x_5 &= 11 \\ 3x_1 + 4x_2 + 2x_3 + x_6 &= 8 \\ x_1, x_2, x_3, x_4, x_5, x_6 &\geq 0. \end{aligned}$$

Definition (Zulässigkeit, Unbeschränktheit) *Ein Vektor $x \in \mathbb{R}^n$ heißt* zulässig *für ein lineares Programm (P), wenn $Ax \leq b$ und $x \geq 0$. Sonst heißt x* unzulässig.

Das Problem (P) heißt unzulässig, *wenn es für (P) keine zulässige Lösung gibt.*

(P) heißt unbeschränkt, *wenn es für alle $M \in \mathbb{R}$ eine zulässige Lösung x von (P) mit $c^t x > M$ gibt.*

Beispiel *Hier existiert eine zulässige Lösung*

$$x_1 = x_2 = x_3 = 0, x_4 = 5, x_5 = 11, x_6 = 8$$

mit Zielfunktionswert (ZFW) 0.

Frage 1: Ist die vorliegende Lösung bereits optimal?

Definition (Optimallösung) *Sei x zulässig für (P). x heißt* Optimallösung *für (P), wenn keine zulässige Lösunge y existiert mit $c^t y > c^t x$.*

Frage 2: Wenn nein, wie erreicht man eine bessere Lösung?

Beobachtung *Durch Erhöhen von x_1, x_2 oder x_3 kann versucht werden, den ZFW zu erhöhen. (x_1, x_2, x_3 haben positiven Koeffizient in der Zielfunktion.)*

Versuchen wir x_1 zu erhöhen, x_2 und x_3 bleiben 0.

$$\left.\begin{array}{lll} 2x_1 + x_4 = 5 & \Rightarrow x_4 = 5 - 2x_1 \geq 0 & \Rightarrow x_1 \leq \frac{5}{2} \\ 4x_1 + x_5 = 11 & \Rightarrow x_5 = 11 - 4x_1 \geq 0 & \Rightarrow x_1 \leq \frac{11}{4} \\ 3x_1 + x_6 = 8 & \Rightarrow x_6 = 8 - 3x_1 \geq 0 & \Rightarrow x_1 \leq \frac{8}{3} \end{array}\right\} \Rightarrow x_1 \leq \frac{5}{2}.$$

x_1 darf also maximal auf $\frac{5}{2}$ erhöht werden. x_4 wird dann 0.

Neue Lösung $x_1 = \frac{5}{2}, x_2 = x_3 = x_4 = 0, x_5 = 1, x_6 = \frac{1}{2}$ mit ZFW $\frac{25}{2}$.

x_1 muss mit x_4 die Rolle tauschen; eliminiere also x_1 in der 1. Restriktion:

$$\begin{aligned}
2x_1 + 3x_2 + x_3 + x_4 &= 5 \\
\Leftrightarrow x_1 &= \tfrac{5}{2} - \tfrac{3}{2}x_2 - \tfrac{1}{2}x_3 - \tfrac{1}{2}x_4 \\
\Leftrightarrow \tfrac{3}{2}x_2 + \tfrac{1}{2}x_3 + \tfrac{1}{2}x_4 + x_1 &= \tfrac{5}{2}.
\end{aligned}$$

Transformiere Zielfunktion:

$$5 \cdot \left(\tfrac{5}{2} - \tfrac{3}{2}x_2 - \tfrac{1}{2}x_3 - \tfrac{1}{2}x_4\right) + 4x_2 + 3x_3 = \tfrac{25}{2} - \tfrac{7}{2}x_2 + \tfrac{1}{2}x_3 - \tfrac{5}{2}x_4.$$

Nun kann nur mehr x_3 durch Erhöhung zu einer Erhöhung des ZFW beitragen.

Forme nun auch die restlichen zwei Ungleichungen um und erhalte so das neue System

$$\begin{aligned}
\tfrac{3}{2}x_2 + \tfrac{1}{2}x_3 + \tfrac{1}{2}x_4 + x_1 &= \tfrac{5}{2} \\
-5x_2 - 2x_4 + x_5 &= 1 \\
-\tfrac{1}{2}x_2 + \tfrac{1}{2}x_3 - \tfrac{3}{2}x_4 + x_6 &= \tfrac{1}{2}.
\end{aligned}$$

Erhöhe nun x_3:

$$\left.\begin{aligned}
x_1 &= \tfrac{5}{2} - \tfrac{1}{2}x_3 \geq 0 \quad \Rightarrow x_3 \leq 5 \\
x_6 &= \tfrac{1}{2} - \tfrac{1}{2}x_3 \geq 0 \quad \Rightarrow x_3 \leq 1
\end{aligned}\right\} \Rightarrow x_3 \leq 1.$$

x_3 darf also maximal auf 1 erhöht werden. x_6 wird dann 0.

Neue Lösung $x_1 = 2, x_2 = 0, x_3 = 1, x_4 = 0, x_5 = 1, x_6 = 0$ mit ZFW 13.

x_3 und x_6 tauschen nun Rollen; elimiere also x_3:

$$x_3 = 1 + x_2 + 3x_4 - 2x_6.$$

Einsetzen in Zielfunktion liefert

$$\tfrac{25}{2} - \tfrac{7}{2}x_2 + \tfrac{1}{2}x_3 - \tfrac{5}{2}x_4 = \tfrac{25}{2} - \tfrac{7}{2}x_2 + \tfrac{1}{2} + \tfrac{1}{2}x_2 + \tfrac{3}{2}x_4 - x_6 = 13 - 3x_2 - x_4 - x_6.$$

Die vorliegende Lösung ist also optimal!

2 Stoßrichtungen:

1. Korrektheit der Methode und Klärung offener Fragen (Bestimmung einer zulässigen Startbedingung, Erkennung von unzulässigen bzw. unbeschränkten Problemen).
2. Praktische Durchführung, algorithmische Umsetzung.

2 Das Simplexverfahren

Das im Beispiel hergeleitete Verfahren geht zurück auf George Dantzig.

2.1 Herleitung eines Rechenschemas für die Transformationsschritte

Allgemeine Situation: $m \times n$ Matrix A, rechter Seitenvektor $b \in \mathbb{R}^m$, Zielfunktionsvektor $c \in \mathbb{R}^n$, $x \in \mathbb{R}^{m+n}$. Problem in der Form

$$\max c^t x$$

s.t.

$$\begin{aligned} Ax &= b \\ x &\geq 0 \end{aligned}$$

Annahme: $\operatorname{rg} A = m$. Daher enthält A m linear unabhängige Spalten,

$$A = (A_B, A_N)$$

(ggf. nach Vertauschen der Spalten), wobei A_B die m linear unabhängigen Spalten und A_N die restlichen Spalten sind. Ebenso

$$x = (x_B, x_N)$$

(ggf. nach derselben Vertauschung wie bei A).

$$\begin{aligned} Ax = \begin{pmatrix} A_B & A_N \end{pmatrix} \begin{pmatrix} x_B \\ x_N \end{pmatrix} &= b \\ \Leftrightarrow \quad \underbrace{A_B}_{\text{invertierbar}} \quad x_B + A_N x_N &= b, \end{aligned}$$

somit

$$x_B = \underbrace{A_B^{-1} b}_{\tilde{b}} - \underbrace{A_B^{-1} A_N}_{\tilde{A}_N} x_N.$$

Allgemeine Lösung des Gleichungssystems:

$$x = \begin{pmatrix} \tilde{b} - \tilde{A}_N x_N \\ x_N \end{pmatrix}.$$

In der Folge interessieren wir uns für Lösungen mit $x_N = 0$ (auch als *Basislösungen* bekannt).

Beispiel

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Hier ist

$$A_B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad A_N = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

und

$$A_B^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}, \quad A_B^{-1} b = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad A_B^{-1} A_N = \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

Definition (Basis, Nichtbasis) *Sei A eine $m \times n$ Matrix, $\operatorname{rg} A = m$. A_B sei eine Untermatrix von A mit $\operatorname{rg} A_B = m$.*

1. *x heißt* Basislösung *von $Ax = b$ zur Basis B, falls alle Komponenten von x, die nicht Spalten von A_B entsprechen, gleich 0 sind.*
2. *Die Komponenten von x, die zu Spalten von A_B gehören, werden als* Basisvariablen *(BV) bezeichnet, alle anderen als* Nichtbasisvariablen *(NBV).*
3. *Die Menge der Indizes der Basisvariablen wird als* Basis *bezeichnet, die der Nichtbasisvariablen als* Nichtbasis.

Die Matrix bestehend aus den zur Basis (Nichtbasis) gehörenden Spalten von A wird Basismatrix *A_B (*Nichtbasismatrix *A_N) genannt.*

Bemerkung *Durch Vorgabe der Werte der Nichtbasisvariablen sind die Werte der Basisvariablen eindeutig bestimmt.*

Definition (Zulässigkeit, Entartung von Basislösungen bzw. Basen)

1. *Eine Basislösung heißt* entartet *(*degeneriert*), wenn mindestens eine Basisvariable den Wert 0 annimmt.*
2. *Eine Basislösung von $Ax = b$ heißt* zulässig, *wenn $x_B \geq 0$.*
3. *Eine Basis heißt* entartet, *wenn die zugehörige Basislösung entartet ist.*
4. *Eine Basis heißt* zulässig, *wenn die zugehörige Basislösung zulässig ist.*

Bemerkungen

1. *Jeder Basis entspricht genau eine zugehörige Basislösung.*
2. *Jeder nicht-entarteten Basislösung entspricht genau eine Basis (gilt nicht bei Entartung).*

Motivation hinter Basislösungen:

- Sie treten im exemplarisch skizzierten Lösungsverfahren auf.
- Hauptsatz der linearen Optimierung (folgt noch).

2.2 Hauptsatz der linearen Optimierung

Dieser Satz rechtfertigt die im Beispiel verwendete Methode (Einschränkung auf Basislösungen).

Satz 2.1 (Hauptsatz der linearen Optimierung) *Gegeben sei das lineare Programm*

$$\max c^t x$$

s.t.

$$\begin{aligned} Ax &= b \\ x &\geq 0 \end{aligned}$$

mit A eine $m \times n$ Matrix und $\operatorname{rg} A = m$. Dann gilt:

1. *$\{Ax = b, x \geq 0\}$ besitzt genau dann zulässige Lösungen, wenn $\{Ax = b, x \geq 0\}$ zulässige Basislösungen besitzt.*
2. *Es existiert genau dann eine optimale zulässige Lösung (d.h. das lineare Programm ist weder unzulässig noch unbeschränkt), wenn eine optimale zulässige Basislösung existiert.*

Korollar *Auf der Suche nach Optimallösungen für lineare Programme kann man sich auf die Menge der Basislösungen einschränken.*

BEWEIS

1. x zulässige Lösung von $Ax = b$.

$$\sum_{i=1}^{p} y_i a_i = 0 \quad (x_1, \dots, x_p > 0,\ x_{p+1}, \dots, x_n = 0)$$

$$\tilde{x} = x - \varepsilon \cdot \tilde{y}$$

$$A\tilde{x} = b$$

Es muss auch gelten: $\tilde{x} \geq 0$, d.h. $\tilde{x}_j \geq 0\ \forall j$, $\tilde{x}_j = x_j - \varepsilon \tilde{y}_j$.

- Fall a: $\tilde{y}_i \leq 0 \Rightarrow \tilde{x}_j \geq 0$. Da $x_j \geq 0$ (da x zulässig) und $\varepsilon > 0$, hier keine Einschränkung für ε.
- Fall b: $\tilde{y}_j > 0 \Rightarrow x_j - \varepsilon \tilde{y}_j \geq 0 \Rightarrow \varepsilon \leq \frac{x_j}{\tilde{y}_j}$.

$\Rightarrow$ Wähle ε als $\min\left\{\frac{x_j}{\tilde{y}_j} : \tilde{y}_j > 0\right\} =: \varepsilon^*$ (nichtleer $\Rightarrow \varepsilon^*$ wohldefiniert).

Beobachtung: $\tilde{x}$ hat mindestens eine Nullkomponente mehr als x.

Durch iterative Anwendung erhält man nach endlich vielen Schritten eine zulässige Basislösung.

2. Sei x eine optimale Lösung von Problem P. Wir wollen zeigen, dass eine optimale Basislösung existiert. Vorgangsweise: analog zu Teil 1.

Annahme: $x_1, \dots, x_p > 0$, $x_{p+1}, \dots, x_n = 0$.

Fall 1: $\{a_1, \dots, a_p\}$ lin. unabh. $\Rightarrow x$ ist bereits Basislösung.

Fall 2: $\{a_1, \dots, a_p\}$ lin. abh. Dann gibt es $y \neq 0, y \in \mathbb{R}^p$ mit

$$\sum_{i=1}^{p} a_i y_i = 0.$$

Ergänze y auf Vektor im $\mathbb{R}^n$ durch Nullkomponenten $\rightarrow$ Ergebnis $\tilde{y}$.

Betrachte wieder

$$\tilde{x}(\varepsilon) = x - \varepsilon \tilde{y}.$$

Für $0 \leq \varepsilon \leq \varepsilon^*$ ist $\tilde{x}(\varepsilon)$ zulässig.

$$c^t \tilde{x}(\varepsilon) = c^t x - \varepsilon c^t \tilde{y}.$$

x ist optimal $\Rightarrow \varepsilon c^t \tilde{y} \geq 0$, somit $c^t \tilde{y} \geq 0$.

Betrachte $\varepsilon < 0$. Einschränkungen an ε durch Komponenten mit $\tilde{y}_j < 0$. Definiere

$$\varepsilon^{**} = \max\{\frac{x_j}{\tilde{y}_j} : \tilde{y}_j < 0\}.$$

Für $\varepsilon^{**} \leq \varepsilon \leq 0$ ist $\tilde{x}(\varepsilon)$ zulässig und mit analoger Vorgangsweise wie oben erhält man nun $c^t \tilde{y} \leq 0$.

Insgesamt erhält man also $c^t\tilde{y} = 0$.

Nun gilt

$$c^t\tilde{x} = c^t x - \varepsilon c^t y = c^t x,$$

somit ist $\tilde{x}$ ebenfalls Optimallösung.

Durch iterative Wiederholung steigt die Anzahl der Nullkomponenten und wir landen bei einer optimalen Basislösung. ■

2.3 Detailsicht des Simplexverfahrens zur Lösung linearer Programme

Grundidee:

1. Starte mit einer zulässigen Basislösung.
2. (Optimalitätsüberprüfung) Falls die vorliegende Basislösung bereits optimal ist, Stop. Ansonsten Übergang zu einer neuen zulässigen Basislösung mit besserem Zielfunktionswert (oder Erkennen eines unbeschränkten Problems) durch *Austauschen* einer Basisvariable (wird zu Nichtbasisvariable) gegen eine Nichtbasisvariable (wird zu Basisvariable).

Für den Austauschschritt wichtig:

- Wahl der Nichtbasisvariable, die neu in die Basis kommt (abhängig von Zielfunktion, siehe auch Beispiel).
- Wahl der Basisvariable, die die Basis verlässt (abhängig von Restriktionen, neue Basislösung muss zulässig sein).

Auch noch zu klären:

- Finden einer zulässigen Ausgangslösung in Schritt 1?
- Wie stellt man Vorliegen eines unzulässigen Problems fest? (siehe später)
- Wie stellt man Unbeschränktheit fest?
- Effiziente Durchführung von Schritt 2.
- Frage der Endlichkeit des Verfahrens.

Problem:

$$\begin{aligned} Ax &= b \\ A &= (A_B, A_N) \\ x &= (x_B, x_N)^t \end{aligned}$$

Somit

$$A_B x_B + A_N x_n = b.$$

Basislösung:

$$x_N = 0, x_B = \underbrace{A_B^{-1}b}_{=\tilde{b}} - \underbrace{A_B^{-1}A_N}_{=\tilde{A}_N} x_N.$$

Einsetzen in Zielfunktion:

$$c^t x = c_B^t x_B + c_N^t x_N = \underbrace{c_B^t A_B^{-1} b}_{=c_B^t\tilde{b}} + (\underbrace{c_N^t - c_B^t A_B^{-1} A_N}_{\text{Vektor der reduzierten Kostenkoeffizienten}=:\tilde{c}_N^t})x_N,$$

wobei $\tilde{b}$ der augenblickliche Wert der Basisvariablen ist, somit $c_b^t\tilde{b}$ der augenblickliche Zielfunktionswert.

Optimalitätskriterium: x ist optimale Basislösung wenn $\tilde{c}_N \leq 0$.

$\Rightarrow$ liefert Kriterium für die Auswahl einer Nichtbasisvariablen, die neu in die Basis kommen soll, wenn noch keine optimale Lösung vorliegt.

Man kann beliebige Nichtbasisvariable mit $\tilde{c}_{N(j)} > 0$ wählen (feinere Auswahlregel siehe später).

Angenommen die j-te Nichtbasisvariable, $x_{N(j)}$, tritt neu in Basis ein. Welche Variable verlässt die Basis? Bisher

$$\begin{aligned} x_B &= \tilde{b} - \tilde{A}_N x_N \\ x_N &= 0 \end{aligned}$$

Nun $x_{N(j)} = \varepsilon$, $\varepsilon > 0$ und ε so groß wie möglich, unter der Einschränkung, dass neue Basislösung wieder zulässig ist.

$$\tilde{x}_B(\varepsilon) = \tilde{b} - \tilde{A}_N \underbrace{\overline{x}_N}_{\text{neuer } x_N\text{-Vektor}} = \tilde{b} - \tilde{a}_j \cdot \varepsilon,$$

wobei $\tilde{a}_j$ die Spalten von $\tilde{A}_N$ sind, die zur Variable $x_{N(j)}$ gehören.

$$\underbrace{\tilde{b}}_{\geq 0} - \tilde{a}_j \underbrace{\varepsilon}_{>0} \geq 0$$

(da x zulässig), damit $\overline{x}$ zulässig.

$$\varepsilon \leq \frac{\tilde{b}_i}{\tilde{a}_{ij}}$$

für alle i mit $\tilde{a}_{ij} > 0$. Definiere

$$\varepsilon^* = \min\{\frac{\tilde{b}_i}{\tilde{a}_{ij}} \mid \tilde{a}_{ij} > 0\}.$$

Der Fall $\{\tilde{a}_{ij} > 0\} = \emptyset$ wird im Anschluss diskutiert.

Die neue Basislösung erhält man, indem man die neue Basisvariable auf den Wert ε^* setzt.

Als die Basis verlassende Basisvariable kann eine beliebige Basisvariable $x_{B(i)}$ mit $\varepsilon^* = \frac{\tilde{b}_i}{\tilde{a}_{ij}}$ gewählt werden.

Im Fall $\tilde{a}_{ij} \leq 0$ für alle i liegt ein unbeschränktes Problem vor. (ε kann beliebig groß werden.) Hier haben wir also ein *Kriterium für das Vorliegen eines unbeschränkten linearen Programms*!

Nun fehlt noch die Umsetzung des Austauschschritts (*Pivot*-Schritt). Wir brauchen eine Problemdarstellung bzgl. der neuen Basis/Nichtbasis.

Beim Start des Verfahrens (mittels der künstlichen Variablen) haben wir $A_B = I$ (Einheitsmatrix).

Nach einem Austauschschritt:

$$\begin{pmatrix} 1 & & & a_{1s} & & & \\ & \ddots & & \vdots & & & \\ & & 1 & & & & \\ & & & a_{rs} & & & \\ & & & & 1 & & \\ & & & \vdots & & \ddots & \\ & & & a_{ms} & & & 1 \end{pmatrix}$$

In späteren Schritten haben wir

$$I \cdot x_B + \tilde{A}_N x_N = \tilde{b}.$$

Auch hier haben wir die Situation, dass anstelle der Einheitsmatrix nach dem Austauschschritt eine Matrix auftaucht, die sich in einer Spalte von der Einheitsmatrix unterscheidet.

Frage: Wie sieht die Inverse einer Matrix der Form

$$C = \begin{pmatrix} 1 & & & d_1 & & & \\ & \ddots & & \vdots & & & \\ & & 1 & & & & \\ & & & d_r & & & \\ & & & & 1 & & \\ & & & \vdots & & \ddots & \\ & & & d_m & & & 1 \end{pmatrix}$$

$(d_r \neq 0)$ aus?

$$C^{-1} = \frac{1}{d_r} \cdot \begin{pmatrix} d_r & & & -d_1 & & & \\ & \ddots & & \vdots & & & \\ & & d_r & -d_{r-1} & & & \\ & & & 1 & & & \\ & & & -d_{r+1} & d_r & & \\ & & & \vdots & & \ddots & \\ & & & -d_m & & & d_r \end{pmatrix}$$

(Hier ist also auch nur eine Spalte von der Einheitsmatrix verschieden.)

Auf diese Weise lässt sich die neue Matrix $\tilde{A}_{\overline{N}}$ für die neue Nichtbasis $\overline{N}$ in eleganter Weise aus der alten Matrix $\tilde{A}_N$ berechnen.

$$\begin{aligned} \tilde{A}_N &= (t_{ij}) \\ \tilde{A}_{\overline{N}} &= (\bar{t}_{ij}) \end{aligned}$$

Sei s die Spalte von A_N, die verschieden von $A_{\overline{N}}$ ist (d.h. hier spielt sich Wechsel in Nachtbasis ab). Sei r die Zeile, die der Variablen entspricht, die die Basis verlässt.

Es gilt

$$\bar{t}_{ij} = \begin{cases} \frac{1}{t_{rs}} & i = r, j = s \\ \frac{t_{rj}}{t_{rs}} & i = r, j \neq s \\ -\frac{t_{is}}{t_{rs}} & i \neq r, j = s \\ t_{ij} - \frac{t_{is} t_{rj}}{t_{rs}} & i \neq r, j \neq s \end{cases}$$

Zur Illustration anhand unseres Beispiels:

$$\max 5x_1 + 4x_2 + 3x_3$$

s.t.

$$\begin{aligned} 2x_1 + 3x_2 + x_3 + x_4 &= 5 \\ 4x_1 + x_2 + 2x_3 + x_5 &= 11 \\ 3x_1 + 4x_2 + 2x_3 + x_6 &= 8 \end{aligned}$$

Start:

$$\begin{aligned} B &= (4,5,6) \\ N &= (1,2,3) \end{aligned}$$

$$A_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

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

$$b = \begin{pmatrix} 5 \\ 11 \\ 7 \end{pmatrix} = \tilde{b} \text{ zu Beginn}$$

$$c_B = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$c_N = \begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix}$$

x_1 soll neu in Basis,

$$\varepsilon^* = \min\left\{\frac{5}{2}, \frac{11}{4}, \frac{8}{3}\right\} = \frac{5}{2}$$

also verlässt x_4 die Basis.
Neue Basis: $\overline{B} = (1, 5, 6)$
Neue Nichtbasis $\overline{N} = (4, 2, 3)$

$$A_{\overline{B}} = \begin{pmatrix} 2 & 0 & 0 \\ 4 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$$

$$A_{\overline{B}}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ -4 & 2 & 0 \\ -3 & 0 & 2 \end{pmatrix}$$

$$A_{\overline{N}} = \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 2 \\ 0 & 4 & 2 \end{pmatrix}$$

Somit

$$\tilde{A}_{\overline{N}} = A_{\overline{B}}^{-1} \cdot A_{\overline{N}} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ -4 & 2 & 0 \\ -3 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 2 \\ 0 & 4 & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{3}{2} & \frac{1}{2} \\ -2 & -5 & 0 \\ -\frac{3}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

und

$$\tilde{b}_{\text{neu}} = A_{\overline{B}}^{-1} b = \begin{pmatrix} \frac{5}{2} \\ 1 \\ \frac{1}{2} \end{pmatrix}.$$

Zum Beispiel hatten wir $x_1 = \frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3$.

Strukturierte Vorgangsweise in einem *Tableau* (erste Zeile: Zielfunktionszeile; dann: Basisvariablen):

$$\begin{array}{r|rrrl} 0 & 5 & 4 & 3 & \\ \hline 5 & \mathbf{2} & 3 & 1 & x_4 \\ 11 & 4 & 1 & 2 & x_5 \\ 8 & 3 & 4 & 2 & x_6 \end{array},$$

wobei links oben der *negative* Wert der augenblicklichen Zielfunktion steht.

Im nächsten Schritt ($r = 1, s = 1$):

$$\begin{array}{c|cccl} -\frac{25}{2} & -\frac{5}{2} & -\frac{7}{2} & \frac{1}{2} & \\ \hline \frac{5}{2} & \frac{1}{2} & \frac{3}{2} & \frac{1}{2} & x_1 \\ 1 & -2 & -5 & 0 & x_5 \\ \frac{1}{2} & -\frac{3}{2} & -\frac{1}{2} & \frac{1}{2} & x_6 \end{array}$$

Oben stehen die augenblicklich reduzierten Kostenkoeffizienten, links die augenblicklichen Werte der Basisvariablen.

Das Element in Zeile r (*Pivotzeile*) und Spalte s (*Pivotspalte*) heißt *Pivotelement*.

Frage: Endlichkeit des Verfahrens?

2.4 Endlichkeit des Simplexverfahrens

Im Spezialfall, dass sich in jeder Iteration ein verbesserter Zielfunktionswert ergibt, ist die Endlichkeit garantiert. (Es gibt nur endlich viele verschiedene Basen.)

Problemfall: Existieren entartete Basislösungen, so kann der Fall auftreten, dass sich der Zielfunktionswert von einer Iteration zur nächsten nicht ändert. Es kann also passieren, dass man nach einer Abfolge von Austauschschritten wieder bei einer bereits betrachteten Basis ankommt. (Somit taucht ein Kreisen auf, kein endliches Verfahren!)

Es sind Beispiele bekannt, für die ein solches Kreisen wirklich auftritt (etwa Beispiel von Bland).

Es gibt verschiedene Methoden, um Kreisen zu verhindern. Hier 2 Methoden:

1. *Regel von Bland* („Kleinste-Index-Regel“): schränkt sowohl Wahl der Pivot-Zeile als auch Wahl der Pivot-Spalte ein.

    - Wahl der Pivot-Spalte: Wähle erste Nichtbasisvariable (d.h. mit kleinstem Index, Reihenfolge im Tableaut ist gleichgültig), die einen positiven reduzierten Kostenkoeffizienten hat. (Das ist eine eindeutige Festlegung!)

      Beispiel

      $$\begin{array}{c|cccc} & x_1 & x_4 & x_3 & x_4 \\ & -9 & 7 & 9 & 2 \\ \hline & & & & \end{array}$$

      Hier würde man x_3 auswählen.

    - Wahl der Pivot-Zeile: Falls mehrere Basisvariablen bei Bestimmung von ε^* den Minimalwert ergeben, wird jene als Pivotzeile gewählt, die den kleinsten Index hat.

2. *Lexikographische Auswahlregel*: Keine Einschränkung in der Wahl der Pivotspalte, aber höherer Aufwand zur Wahl der Pivotzeile.

    Sei $\tilde{a}_j$ die Pivotspalte des Tableaus. Für jede Zeile i im Tableau, für die $\tilde{a}_{ij} > 0$, berechne den Vektor, der sich ergibt, wenn diese Zeile durch das Pivotelement dividiert wird (inkl. rechter Seite als 0-te Spalte). Wähle jene Zeile, deren zugehöriger Vektor lexikographisch minimal ist.

Beispiel

1	0	5	**3**
2	4	6	-1
3	0	7	9

Hier führt die 1. Zeile auf $\left(\frac{1}{3}, 0, \frac{5}{3}, 1\right)$. *Die 3. Zeile führt auf* $\left(\frac{1}{3}, 0, \frac{7}{3}, 1\right)$. *Die 3. Komponente ergibt die lexikographische Ordnung; wähle hier also die 3. Zeile.*

Man kann folgendes zeigen:

- Jede Zeile des Tableaus (mit Ausnahme der Zielfunktionszeile) bleibt lexikographisch positiv.
- Die Zielfunktionszeile nimmt von Iteration zu Iteration lexikographisch ab.

Definition (lexikographische Ordnung für Vektoren) *Seien* $u, v \in \mathbb{R}^k$.

$$u < v :\Longleftrightarrow u - v < 0,$$

d.h. die erste Nichtnullkomponente von $u - v$ *ist negativ.*

Beispiele

$$(1, 7, 5) < (3, 6, 2)$$
$$(1, 5, 9) < (1, 7, 1)$$

Bemerkung *Ist das Minimum für* ε^* *eindeutig, so stimmen die Ergebnisse mit und ohne lexikographischer Regel überein.*

Für beide Methoden kann bewiesen werden, dass Kreisen verhindert wird. Die Endlichkeit des Simplexverfahrens ist somit garantiert.

Nachteil der Regel von Bland: recht inflexibel, weil auch Wahl der Pivotspalte eingeschränkt.

Noch zu behandeln:

- Bestimmung einer zulässigen Ausgangslösung
- Wie viele Pivotschritte können maximal auftreten?

2.5 Bestimmung einer zulässigen Ausgangsbasislösung

In den typischen Methoden werden Hilfsprobleme (neue Zielfunktion, neue Variable) verwendet.

2.5.1 Methode 1

Gegeben lineares Programm

$$\max c^t x$$

s.t.

$$\begin{aligned} Ax &\leq b \\ x &\geq 0, \end{aligned}$$

wobei $\exists i$ mit $b_i < 0$ (sonst: verwende triviale Startlösung).

An einem Beispiel:

$$\max x_1 - x_2 + x_3$$

s.t.

$$
\begin{aligned}
2x_1 - x_2 + 2x_3 &\leq 4 \\
2x_1 - 3x_2 + x_3 &\leq -5 \\
-x_1 + x_2 - 2x_3 &\leq -1 \\
x_1, x_2, x_3 &\geq 0
\end{aligned}
$$

Idee: Führe eine neue (künstliche) Variable ein; sei diese x_0 genannt. Betrachte folgendes Hilfsproblem:

$$\min x_0 \quad \Leftrightarrow \quad \max -x_0$$

s.t.

$$
\begin{aligned}
2x_1 - x_2 + 2x_3 - x_0 &\leq 4 \\
2x_1 - 3x_2 + x_3 - x_0 &\leq -5 \\
-x_1 + x_2 - 2x_3 - x_0 &\leq -1 \\
x_0, x_1, x_2, x_3 &\geq 0.
\end{aligned}
$$

Das Ausgangsproblem besitzt eine zulässige Lösung genau dann, wenn das Hilfsproblem eine Optimallösung mit Zielfunktionswert 0 ($x_0 = 0$) besitzt.

	x_1	x_2	x_3	x_0	
0	0	0	0	-1	Zielfunktion des Hilfsproblems
0	1	-1	1	0	Zielfunktion des Ausgangsproblems
4	2	-1	2	-1	x_4
-5	2	-3	1	$\mathbf{-1}$	x_5
-1	-1	1	-2	-1	x_6

(kein zulässiges Tableau).

Nimm x_0 in Basis auf und werfe eine Variable, die zu einer Restriktion mit $b_i < 0$ und $|b_i|$ max. gehört, aus der Basis.

	x_1	x_2	x_3	x_5	
5	-2	3	-1	-1	
0	1	-1	1	0	
9	0	2	1	-1	x_4
5	-2	3	-1	-1	x_0
4	-3	**4**	-3	-1	x_6

(zulässiges Tableau).

$$\varepsilon^* = \min\left(\frac{9}{2}, \frac{5}{3}, \frac{4}{4}\right)$$

noch nicht optimal für Hilfsproblem.

	x_1	x_6	x_3	x_5	
2	$\frac{1}{4}$	$-\frac{3}{4}$	$\frac{5}{4}$	$-\frac{1}{4}$	
		$\dots$			
7	$\frac{3}{2}$	$-\frac{1}{2}$	$\frac{5}{2}$	$-\frac{1}{2}$	x_4
2	$\frac{1}{4}$	$-\frac{3}{4}$	$\boldsymbol{\frac{5}{4}}$	$-\frac{1}{4}$	x_0
1	$-\frac{3}{4}$	$\frac{1}{4}$	$-\frac{3}{4}$	$-\frac{1}{4}$	x_2

Dann

$$\begin{array}{c|cccc|l}
 & x_1 & x_6 & x_0 & x_5 & \\
\mathbf{0} & 0 & 0 & -1 & 0 & \text{ab nun ignorieren} \\
\frac{3}{5} & \frac{1}{3} & \frac{2}{5} & & -\frac{1}{5} & \\
\hline
3 & 1 & 1 & -2 & 0 & x_4 \\
\frac{8}{5} & \frac{1}{5} & -\frac{3}{5} & \frac{4}{5} & -\frac{1}{5} & x_3 \\
\frac{11}{5} & -\frac{3}{5} & -\frac{1}{2} & \frac{3}{5} & -\frac{2}{5} & x_2
\end{array}$$

(optimale Lösung des Hilfsproblems erreicht) $\Longrightarrow$ Ausgangsproblem zulässig.

$$x_2 = \frac{11}{5}, x_3 = \frac{8}{5}, x_1 = 0, x_4 = 3, x_5 = x_6 = 0$$

ist zulässige Lösung des Ausgangsproblems.

Nun streichen der Hilfszielfunktionszeile. Wenn x_0 in Nichtbasis ist, kann diese Spalte gestrichen werden. (Lösung des Beispiels noch nicht optimal!)

Hilfsproblem in allgemeiner Form:

$$\min x_0$$

s.t.

$$Ax - x_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \leq b$$
$$x_0, \dots, x_n \geq 0$$

(Beweis der Korrektheit als Übung.)

2.5.2 Methode 2

Eine künstliche Variable pro Restriktion.

$$\max x_1 + x_2 + x_3$$

s.t.

$$\begin{aligned}
x_1 + 2x_2 + 3x_3 &= 3 \\
-x_1 + 2x_2 + 6x_3 &= 2 \\
4x_2 + 3x_3 &= 5 \\
3x_3 + x_4 &= 1 \\
x_1, x_2, x_3, x_4 &\geq 0
\end{aligned}$$

Hilfsproblem

$$\min x_5 + x_6 + x_7 + x_8 \quad \Leftrightarrow \quad \max -x_5 - x_6 - x_7 - x_8$$

s.t.

$$\begin{aligned}
x_1 + 2x_2 + 3x_3 + x_5 &= 3 \\
-x_1 + 2x_2 + 6x_3 + x_6 &= 2 \\
4x_2 + 9x_3 + x_7 &= 5 \\
3x_3 + x_4 + x_8 &= 1 \\
x_1, \dots, x_8 &\geq 0
\end{aligned}$$

Beobachtung: Hilfsproblem hat Optimalwert 0 $\Leftrightarrow$ Ausgangsproblem besitzt zulässige Lösung.

$(5, 6, 7, 8)$ stellt zulässige Basis für Hilfsproblem dar:

$$x_5 = 3, x_6 = 7, x_7 = 5, x_8 = 1.$$

Wir müssen die Hilfszielfunktion durch NBV $x_1, \dots, x_4$ darstellen.

$$\begin{aligned}
-x_5 &= x_1 + 2x_2 + 3x_3 - 3 \\
-x_6 &= -x_1 + 2x_2 + 6x_3 - 2 \\
-x_7 &= 4x_2 + 5x_3 - 5 \\
-x_8 &= 3x_3 + x_4 - 1
\end{aligned}$$

Addieren ergibt $-x_5 - x_6 - x_7 - x_8 = 8x_2 + 21x_3 + x_4 - 11$.

	x_1	x_2	x_3	x_4	
-11	0	8	21	1	Hilfszielfunktion
0	1	1	1	0	Zielfunktion
3	1	2	3	0	x_5
2	-1	2	6	0	x_6
5	0	4	9	0	x_7
1	0	0	3	1	x_8

Nun wird das Hilfsproblem gelöst. Eine mögliche Folge von Austauschschritten:

1. x_4 rein, x_8 raus
2. x_3 rein, x_4 raus
3. x_2 rein, x_6 raus
4. x_1 rein, x_5 raus

	x_5	x_6	x_4	x_8	
0	0	-2	-2	-1	Hilfsfunktionswert
1					x_1
$\frac{1}{2}$					x_2
0					x_7
$\frac{1}{3}$					x_3

Optimaler Wert des Hilfsproblems ist 0, d.h. es gibt eine zulässige Lösung für das Ausgangsproblem.

$$x_1 = 1, x_2 = \frac{1}{2}, x_3 = \frac{1}{3}, x_4 = 0.$$

Zum Weiterrechnen: Hilfszielfunktion streichen, ebenso alle Spalten zu künstlichen Variablen in Nichtbasis.

Bemerkung *Es gibt viele Varianten und Modifikationen dieser Methoden. Die* M-Methode *etwa ist eine Variante von Methode 1 (Hilfszielfunktion eigentliche Zielfunktion $-Mx$, M groß genug).*

2.6 Einige Erweiterungen des Simplexverfahrens

1. Gleichungen
2. Variablen ohne Vorzeichenbeschränkungen
3. Variablen mit oberen Schranken (z.B. $x_4 \leq 39$)

Ziel: direkt behandeln, ohne Umformung (in Fällen 1, 2), bzw. ohne Behandlung als explizite Restriktion (Fall 3).

2.6.1 Gleichungen

Variante 1 Gleichungssystem $Ax = b$ kann nach einer Basis B aufgelöst werden und dann weitermachen.

Nachteile:

- aufwendig
- Es ist nicht offensichtlich zulässig.

Variante 2 Verwende „künstliche" Variablen (eine pro Gleichung).

Beispiel

$$\max x_1 + 2x_2 - x_3$$

s.t.

$$\begin{aligned} x_1 + x_2 &\leq 4 \\ 2x_1 + x_2 + x_3 &= 5 \\ x_2 - x_3 &= -1 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

Hilfszielfunktion:

$$\min y_1 + y_2$$

s.t.

$$\begin{aligned} x_1 + x_2 &\leq 4 \\ 2x_1 + x_2 + x_3 + y_1 &= 5 \\ -x_2 + x_3 + y_2 &= 1 \\ x_1, x_2, x_3, y_1, y_2 &\geq 0 \end{aligned}$$

Beobachtung: Hat Hilfsproblem Optimalwert 0, so existiert eine zulässige Basis für das Ausgangsproblem.

-6	2	0	2	*HZF*
0	1	2	-1	*ZF*
4	1	1	0	x_4
5	2	1	1	y_1
1	0	-1	1	y_2

Zur Übung fertigrechnen. Optimallösung: $x_1 = 2, x_2 = 0, x_3 = 1$.

2.6.2 Variablen ohne Vorzeichenbeschränkung

Optimalitätskriterium abändern: Änderung nur für nicht vorzeichenbeschränkte Variablen. Basislösung $x = (x_B, \dots, x_N)$ ist optimal $\iff \tilde{c}_j \geq 0$ für alle $j \in N$ und x_j ist vorzeichenbeschränkt und $\tilde{c}_j = 0$ für alle $j \in N$ und x_j ist nicht vorzeichenbeschränkt.

Verändertes Spaltenauswahlkriterium: Wähle NBV, die das Optimalitätskriterium nicht erfüllt.

Veränderte Bestimmung der Pivotzeile (Variable, die neu in Basis kommt): Für nicht vorzeichenbeschränkte Basisvariablen ist kein Test auf Positivität erforderlich, ignoriere also solche Zeilen.

Annahme: x_s kommt neu in Basis. Fallunterscheidung:

1. $\tilde{c}_s > 0$, d.h. x_s nicht erhöht. (Anmerkung: Falls x_s vorzeichenbeschränkt, ist dies der einzige Fall.)

$$x_i = \tilde{b}_i - \tilde{a}_{is} x_s$$

(x_i ist die i-te Basisvariable).

a) x_i ist vorzeichenbeschränkt.

$$x_s \leq \frac{\tilde{b}_i}{\tilde{a}_{is}}$$

für $\tilde{a}_{is} > 0$. Keine Einschränkung für x_s für $\tilde{a}_{is} \leq 0$.

b) x_i ist nicht vorzeichenbeschränkt. Keine Einschränkung für Wert von x_s, keine Aktion nötig.

2. x_s wird reduziert, wird < 0. (Dieser Fall tritt nur für nicht vorzeichenbeschr. x_s auf.)

$$x_i = \tilde{b}_i - \tilde{a}_{is} x_s \text{ wie oben.}$$

a) x_i ist vorzeichenbeschränkt.

$$|x_s| \leq \frac{\tilde{b}_i}{|\tilde{a}_{is}|} \text{ für } \tilde{a}_{is} < 0.$$

Keine Einschränkung für $\tilde{a}_{is} \geq 0$.

b) x_i nicht vorzeichenbeschränkt. Keine Aktion nötig.

$$\begin{aligned}\varepsilon^* = \min\Bigg(&\min\left\{\frac{\tilde{b}_i}{\tilde{a}_{is}} : \tilde{a}_{is} > 0 \text{ und } x_i \text{ vorzeichenbeschränkt und } c_s > 0\right\}, \\ &\min\left\{\frac{\tilde{b}_i}{|\tilde{a}_{is}|} : \tilde{a}_{is} < 0 \text{ und } x_i \text{ vorzeichenbeschränkt und } c_s < 0\right\}\Bigg)\end{aligned} \tag{2.1}$$

Eine Zeile, für die das Minimum in ε^* angenommen wird, kann als Pivotzeile gewählt werden.
Verändertes Kriterium für das Vorliegen eines unbeschränkten Problems: Unbeschränktheit liegt vor, wenn es eine NBV (mit $c_s \neq 0$) gibt, für die (2.1) keine Einschränkung ergibt.

Beispiel

$$\max x_1 + 2x_2$$

s.t.

$$\begin{aligned} x_1 - x_2 &\leq 2 \\ x_1 + x_2 &\leq 1 \\ -x_1 + x_2 &\leq 3 \end{aligned}$$

(x_1, x_2 nicht vorzeichenbeschränkt).

	x_1	x_2	
0	1	2	
2	1	−1	x_3
1	1	**1**	x_4
3	−1	1	x_5

Hier ist

$$\varepsilon^* = \min\left\{\frac{3}{1}, \frac{1}{1}\right\} = 1.$$

Wenn beschränkt, wäre hier fertig.

	x_1	x_4	
−2	−1	−2	
3	2	1	x_3
1	1	1	x_2
2	**−2**	−1	x_5

Hier ist

$$\tilde{c}_1 < 0$$
$$\varepsilon^* = \min\left\{\frac{2}{2}\right\} = 1.$$

	x_5	x_4	
-3	$-\frac{1}{2}$	$-\frac{3}{2}$	
5			x_3
2			x_2
-1	$-\frac{1}{2}$		x_1

Optimallösung: $x_1 = -1, x_2 = 2, x_3 = 5$, *Zielfunktionswert: 3.*

2.6.3 Variablen mit oberen Schranken

Beschränkte Variable: $0 \leq x_j \leq d_j$, $d_j > 0$. (In Wirklichkeit: Kann untere Schranke beliebig verschieben. Sonst: Kombiniere 2 und 3.) Das entspricht

$$x_j + \overline{x}_j = d_j \text{ mit } x_j, \overline{x}_j \geq 0. \tag{2.2}$$

$\overline{x}_j$ ist die *Komplementärvariable* zu x_j (ebenso umgekehrt).

Statt (2.2) als Restriktion im Simplexverfahren mitzuführen, führt man nur x_j oder $\overline{x}_j$ als augenblickliche Variable der Komponente j mit.

Wir brauchen neue Auswahlregel: Sei x_s die Variable, die neu in die Basis kommt. Mögliche Situationen:

1. x_s ist eine beschränkte Variable, d.h. $x_s \leq d_s$, also $\varepsilon^* \leq d_s$.
2. Basisvariable $x_i \geq 0$ und x_i nicht nach oben beschränkt.
$$x_s \leq \frac{\tilde{b}_i}{\tilde{a}_{is}} \text{ für } \tilde{a}_{is} > 0,$$
alles wie üblich. D.h.
$$\varepsilon^* \leq \frac{\tilde{b}_i}{\tilde{a}_{is}} \text{ für } \tilde{a}_{is} > 0$$
und x_i nach oben beschränt.
3. Basisvariable $x_i \geq 0$ und x_i nach oben beschränkt ($x_i \leq d_i$).
   a) $\tilde{a}_{is} < 0$: $x_s \leq \frac{\tilde{b}_i - d_i}{\tilde{a}_{is}}$,
   $$\varepsilon^* \leq \frac{\tilde{b}_i - d_i}{\tilde{a}_{is}}.$$
   b) $\tilde{a}_{is} > 0$: $x_s \leq \frac{\tilde{b}_i}{\tilde{a}_{is}}$ wie üblich,
   $$\varepsilon^* \leq \frac{\tilde{b}_i}{\tilde{a}_{is}}.$$

   Noch zu überlegen: Vorgangsweise, wenn Minimum durch $\triangle$ oder $\square$ angenommen wird.

Ad Situation 1: x_s wird auf d_s gesetzt, i.A. wird kein BV zu 0. Was tun?

$$x_s = d_s \Rightarrow \overline{x}_s = 0$$

Idee: Tausche x_s und $\overline{x}_s$ aus. D.h. statt x_s kommt $\overline{x}_s$ ins Problem. (Bzw. wäre $\overline{x}_s$ dort, dann kommt x_s neu hinzu.)

Wann immer Situation 1 auftaucht, wird Variable durch ihr Komplement ersetzt.

$$\tilde{a}_{i1}x_{N(1)} + \cdots + \tilde{a}_{is}x_{N(n)} + \cdots + \tilde{a}_{in}x_{N(n)} + x_{B(i)} = \tilde{b}_i,$$

wobei $N(j)$ Index der j-ten NBV, $B(i)$ Index der i-ten BV.

$$x_{N(s)} + \overline{x}_{N(s)} = d_{N(s)} \Rightarrow x_{N(s)} = d_{N(s)} - \overline{x}_{N(s)}$$

Somit

$$\tilde{a}_{i1}x_{N(1)} + \cdots = \tilde{a}_{is}\overline{x}_{N(s)} + \cdots + \tilde{a}_{in}x_{N(n)} + x_{B(i)} = \tilde{b}_i - \tilde{a}_{is}d_{N(s)}$$

Beobachtung: Übergang zur Komplementärvar. kann in 2 Schritten im Tableau vollzogen werden.

1. Spalte zur Var. $x_{N(s)}$ mit -1 mult.
2. rechte Seite ersetzen durch $\tilde{b}_i - \tilde{a}_{is}d_{N(s)}$

$\overline{x}_{N(s)}$ ist nun 0, ist in Nichtbasis.
Algorithmus kann nun fortgesetzt werden, nächster Pivotschritt.
Ad Situation 3: Fall □
Basisvariable x_r (r: Pivotzeile) erreicht Schranke, muss also durch ihre Komplementärvariable ersetzt werden. Vorgangsweise:

1. Pivotschritt mit Pivotelement $\tilde{a}_{rs} < 0$ ($\varepsilon^* = \frac{\tilde{b}_r - d_r}{\tilde{a}_{rs}}$)
2. Übergang von x_r zu Komplementärvariable (neue NBV). Übergang wie in Situation 1.

Beispiel

$$\max -x_1 + x_2$$

s.t.

$$\begin{aligned} x_1 - x_2 &\leq 2 \\ -x_1 + x_2 &\leq 3 \\ x_2 &\leq 4 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Somit $x_2 + \overline{x}_2 = 4$.

$$\begin{array}{r|rr|l} & x_1 & x_2 & \\ 0 & -1 & 4 & \\ \hline 2 & 1 & -1 & x_3 \\ 3 & -1 & 1 & x_4 \end{array}$$

$$\varepsilon^* = \min\{ \underbrace{4}_{\textit{Situation 1}}, \underbrace{3}_{\textit{Situation 2}} \} = 3.$$

$$\begin{array}{r|rr|l} & x_1 & x_4 & \\ -12 & 3 & -4 & \\ \hline 5 & 0 & 1 & x_3 \\ 3 & -1 & -1 & x_2 \end{array}$$

$$\varepsilon^* = \min \left\{ \underbrace{\frac{3-4}{-1}}_{Situation\ 3} \right\} = 1.$$

Fallunterscheidung, in welcher Situation Minimum für ε^* erreicht wird.

$$\begin{array}{r|rr|l} & x_2 & x_4 & \\ -3 & 3 & -1 & \\ \hline 5 & 0 & 1 & x_3 \\ -3 & -1 & -1 & x_1 \end{array}$$

$$\begin{array}{r|rr|l} & \overline{x}_2 & x_4 & \\ & -3 & -1 & \\ \hline 1 & 0 & 1 & x_3 \\ 1 & 1 & -1 & x_1 \end{array}$$

optimal.

Beispiel

$$\max x_1 - 4x_2$$

s.t.

$$\begin{aligned} x_1 - x - 2 + x_3 &= 2 \\ 0 \le x_2 &\le 4 \\ 0 \le x_3 &\le 5 \\ x_1 &\ge 0 \end{aligned}$$

$$\begin{array}{r|rr|l} & x_1 & x_2 & \\ 0 & -1 & 4 & \\ \hline 2 & 1 & -\mathbf{1} & x_3 \end{array}$$

$$\varepsilon^* = \min \left\{ \underbrace{4}_{Situation\ 1}, \underbrace{\frac{2-5}{-1}}_{Situation\ 3} \right\} = 3.$$

$$\begin{array}{r|rr|l} & x_1 & \overline{x}_3 & \\ -12 & 3 & -4 & \\ \hline 3 & -1 & 1 & x_2 \end{array}$$

$$\varepsilon^* = \min \left\{ \underbrace{\frac{3-4}{-1}}_{Sit.\ 3} \right\} = 1.$$

$$\begin{array}{r|rr|l} & x_2 & \overline{x}_3 & \\ -3 & 3 & -1 & \\ \hline -3 & -1 & -1 & x_1 \end{array}$$

(„Temporär unzulässig.“) Übergang zu $\overline{x}_2$.

$$\begin{array}{r|rr|l} & \overline{x}_2 & \overline{x}_3 & \\ -15 & -3 & -1 & \\ \hline 1 & 1 & -1 & x_1 \end{array}$$

Das ist optimal: $x_1 = 1, x_3 = 5, x_2 = 4$.

2.7 Anmerkungen zur Wahl der Pivotspalte

Es gibt keine dominierende Regel. Verschiedene Regeln im Einsatz:

1. *Regel von Dantzig*: Wähle Nichtbasisvariable mit dem stärksten Anstieg im Raum der augenblicklichen NBV. (Für vorzeichenbeschr. Variablen ist der Anstieg durch $\tilde{c}_s$ gegeben, sonst durch $|\tilde{c}_s|$.)

   $|\tilde{c}_j|$: betragsgrößter red. Kostenkoeff. für Variable, die Opt.krit. erfüllt.

   Vorteil: Recht gut für Verwendung durch Hand.

   Nachteil: für große lineare Programme zu aufwendig!

2. *Erste-Index-Regel*: Wähle die erste NBV, die Optimalitätskriterium verletzt

   Vorteil: schnelle Auswertung

   Nachteil: i.A. größere Anzahl an Pivotschritten als bei anderen Verfahren

3. *stärkster Zuwachs der Zielfunktion*: Berechnet für jede mögliche Wahl der Pivotspalte s das zugehörige $\varepsilon^* = \varepsilon^*(s)$ und wählt jene Pivotspalte, für die

   $$\varepsilon^*(s) \cdot |\tilde{c}_s|$$

   (Anstieg der Zielfkt.) maximal ist.

   Nachteil: sehr, sehr aufwendig.

4. *stärkster Anstieg im Raum aller Variablen*: Änderung von NBV $x_{N(j)}$ um 1, so resultiert Änderung des ZFW um $|\tilde{c}_j|$. Änderung der Werte der BV:

   $$x_B = A_B^{-1} - A_B^{-1} A_N x_N = \tilde{b} - \tilde{a}_j.$$

   Hier $x_{N(j)} = 1$, $x_{N(k)} = 0$ für $k \neq j$. $\tilde{a}_j$ ist die j-te Spalte von $\tilde{A}_N$. $x = (x_B, x_N)$. Änderungsvektor

   $$(-\tilde{a}_{ij}, -\tilde{a}_{2j}, \ldots, -\tilde{a}_{mk}, 0, \ldots, 0, \underbrace{1}_{\text{Stelle von } x_{N(j)}}, 0, \ldots, 0).$$

   Komponente zu $x_{N(j)}$ des Gradienten der Zielfunktion im Raum aller Variablen:

   $$\frac{\tilde{c}_j}{\sqrt{1 + \sum_{i=1}^m \tilde{a}_{ij}^2}} \tag{2.3}$$

   Auswahl jener Variable $x_{N(j)}$, für die (2.3) maximal ist.

   Aufwendiger als 1., liefert meistens aber geringere Anzahl an Pivotschritte. Etwas Rechenaufwand kann durch Rekursionsformeln von Goldfarb und Reid eingespart werden.

5. *Kandidatenlisten-Regeln*: Man baut sich Pool von Kandidatenspalten/-variablen auf und wendet auf diesen Pool eine Auswahlregel (z.B. 1.-4.) an. Meistens kommen nur Variablen in den Pool, die das Optimalitätskriterium verletzen. Verschiedene Methoden für Poolmanagement.

2.8 Kurze Anmerkung zur Effizienz des Simplexverfahrens

2 Einflussgrößen:

1. Zahl der Pivotschritte: Kernproblem des Simplexverfahrens,
2. Aufwand pro Pivotschritt: ist polynomial, mit Standardmethoden aus numerischen linearen Algebra gut im Griff.

2.8.1 Zahl der Pivotschritte

- Man kennt keine Spaltenauswahlregel, für die das Simplexverfahren auch im schlechtesten Fall nur eine polynomiale Anzahl von Pivotschritten benötigt.
- Für die meisten bekannten Spaltenauswahlregeln kennt man Beispiele für lineare Programme, für die das resultierende Simplexverfahren exponentiell viele Pivotschritte benötigt (z.B. Klee, Minty Beispiele).
- In der Praxis ist die Anzahl der Pivotschritte i.A. vernünftig klein, oft sogar linear in der Anzahl der Restriktionen (empirische Aussage).
- *Average case Analyse* (Borgwardt, 1977, später Adler und Megiddo): Im durchschnittlichen Fall unter Gleichverteilungsannahme der Daten ist die Anzahl der Pivotschritte polynomiell beschränkt (Borgwardt). Später: $O\left(\min\{n^2, m^2\}\right)$.
- *Smoothed analysis* („geglättete Analyse“; Spielman): Die sich schlecht verhaltenden Instanzen sind recht dünn verteilt.

3 Konvexe Mengen, Polyeder und Zusammenhang zur linearen Optimierung

Definition *$C \subseteq \mathbb{R}^n$ heißt* konvex*, wenn*

$$\forall x, y \in C \ \forall \lambda \in [0,1] : \lambda x + (1-\lambda) y \in C.$$

Lemma *Seien C, D konvex. Dann gilt:*

1. *$\lambda C := \{\lambda x \mid x \in C\}$ ist konvex für alle λ.*
2. *$C \cap D$ ist konvex.*
3. *Der Durchschnitt von endlichen vielen konvexen Mengen ist konvex.*
4. *$C + D = \{x + y \mid x \in C, y \in D\}$ ist konvex.*

Definition (konvexe Hülle) *Sei $S \subseteq \mathbb{R}^n$. Die kleinste (im Sinne von $\subseteq$) konvexe Menge, die S enthält, heißt* konvexe Hülle *von S,* $\operatorname{conv}(S)$*.*

Bemerkung *Sei $S = \{x_1, \ldots, x_k\}$. Dann ist*

$$\operatorname{conv}(S) = \left\{ \sum_{i=1}^{k} \lambda_i x_i \mid x_i \in S, \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

(Dieser Darstellungssatz gilt auch für nicht endliche Mengen.)

Definition (Kegel) *Sei $S \subseteq \mathbb{R}^n$ mit $0 \in S$. S heißt* Kegel *(engl.* cone*), wenn*

$$x \in S \to \forall \alpha > 0 : \alpha x \in S.$$

Beispiele *Skizzen.*

Definition (Hyperebene, Halbraum) *Sei $a \in \mathbb{R}^n, \alpha \in \mathbb{R}$.*

1. *Die Menge*
$$H := \{x \in \mathbb{R}^n \mid a^t x = \alpha\} \subseteq \mathbb{R}^n$$
heißt Hyperebene*.*
(Spezialfälle: $n = 2$ Gerade, $n = 3$ Ebene.)
2. *Die Menge*
$$H_1 = \{x \in \mathbb{R}^n \mid a^t x \geq \alpha\}$$
heißt (abgeschlossener) Halbraum*.*

Definition (Polyeder, Polytop)

1. *Ein* Polyeder *ist der Durchschnitt von endlich vielen Halbräumen (engl.* polyhedron*).*
2. *Ein* Polytop *ist ein nicht leerer, beschränkter Polyeder.*

Beispiele *Skizzen.*

Lemma *Polyeder sind konvexe Mengen.*

Beweis Als Übung. ■

Bemerkung *Die Restriktionenmenge eines linearen Programms beschreibt einen Polyeder.*

Korollar *Die Restriktionenmenge eines linearen Programms ist konvex.*

Frage: Zusammenhang zwischen Polyeder und Simplexverfahren (insbesondere Basislösungen)?

Definition (Extremalpunkt, Ecke, Eckpunkt) *Sei $C \subseteq \mathbb{R}^n$ konvex. $x \in C$ heißt* Extremalpunkt *(oder* Ecke*) von C, falls es keine 2 verschiedenen Punkte $x_1, x_2 \in C$ $(x_1 \neq x_2)$ gibt, sodass*

$$x = \alpha x_1 + (1-\alpha)x_2 \qquad \text{mit } \alpha \in (0,1).$$

Beispiele

- *Kreis: Jeder Randpunkt ist Ecke.*
- *Halbraum: keine Ecken.*

Beispiel

$$\begin{aligned} x_1 + x_2 &\leq 3 \\ x_1, x_2 &\geq 0 \end{aligned}$$

In Standardform:

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

2 NBV, 1 BV.

1. $B = (1), x_1 = 3, x_2 = x_3 = 0$
2. $B = (2), x_1 = 3, x_1 = x_3 = 0$
3. $B = (3), x_3 = 3, x_1 = x_2 = 0$

Skizze.

3.1 Zusammenhang zwischen Basislösungen und Ecken

Satz 3.1 (Zusammenhang zwischen Basislösungen und Ecken) *Sei A eine $m \times n$ Matrix,* $\operatorname{rank} A = m, b \in \mathbb{R}^m$*. Betrachte Polyeder*

$$k := \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}.$$

Dann ist x genau dann eine (zulässige) Ecke von K, wenn x eine (zulässige) Basislösung des linearen Programms mit den Restriktionen $Ax = b, x \geq 0$ ist.

Beweis „$\Longrightarrow$“: Sei

$$x = (x_1, \ldots, x_m, 0, \ldots, 0)$$

eine zulässige Basislösung, d.h. $x_i \geq 0$ für alle $i \in \{1, \ldots, m\}$. $\sum_{i=1}^m a_i x_i = b$, wobei a_i die i-te Spalte von A ist. $\{a_1, \ldots, a_m\}$ ist linear unabhängig.

Annahme: x ist keine Ecke.

Dann existieren y, z $(y \neq z)$ und $\alpha \in (0,1)$, sodass

$$x = \alpha y + (1-\alpha)z$$

und $Ay = b, y \geq 0$ und $Az = b, z \geq 0$.

Da $x_j = 0$ für $j > m$ und $y \geq 0, z \geq 0, \alpha \in (0,1)$, gilt $y_j = 0$ und $z_j = 0$ für alle $j > m$. Insgesamt also

$$\sum_{i=1}^{m} a_i y_i = b \quad \text{und} \quad \sum_{i=1}^{m} a_i z_i = b.$$

Da $\{a_1, \ldots, a_m\}$ linear unabhängig, folgt $y = z$ (lineares Gleichungssystem eindeutig lösbar), Widerspruch.

„$\Longleftarrow$“: Sei $x \in K$ eine (zulässige) Ecke,

$$x = (\underbrace{x_1, \ldots, x_q}_{>0}, 0, \ldots, 0)$$

(ggf. Komponenten umnummerieren). Aus $x \in K$ folgt

$$\sum_{i=1}^{q} x_i a_i = b.$$

Zu zeigen bleibt noch, dass $\{a_1, \ldots, a_q\}$ linear unabhängig ist.

Annahme: $\{a_1, \ldots, a_q\}$ linear abhängig. Dann existieren $y_1, \ldots, y_q \in \mathbb{R}$ mit $y = (y_1, \ldots, y_q) \neq 0$ und $\sum_{i=1}^{q} y_i a_i = 0$.

$$\tilde{y} = (y_1, \ldots, y_q, 0, \ldots, 0) \in \mathbb{R}^n$$

mit $A\tilde{y} = 0$. Da $x_i > 0$ für alle $i \in \{1, \ldots, q\}$, gibt es ein $\varepsilon > 0$ mit

$$u := x + \varepsilon\tilde{y} \geq 0$$
$$v := x - \varepsilon\tilde{y} \geq 0$$

Somit

$$Au = Ax + \varepsilon A\tilde{y} = b + 0 = b$$

(analog $Av = b$) und somit $u \in K, v \in K$ und $u \neq v$ (da $\tilde{y} \neq 0$).

Nun gilt aber

$$x = \frac{1}{2}u + \frac{1}{2}v,$$

Widerspruch dazu, dass x Ecke ist. ∎

$\operatorname{rg} A = m$. $x \in \mathbb{R}^n$ ist Ecke von $\{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\} \Leftrightarrow x$ ist zulässige Lösung.

Konsequenz und Beobachtungen $K = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}, \operatorname{rg} A = m$ (vgl. Satz 3.1).

1. $K \neq \emptyset \Longrightarrow K$ besitzt mindestens eine Ecke.
2. $K \neq \emptyset$ und beschränkt bzw. es liegt beschränktes lineares Programm vor $\Longrightarrow$ Es gibt optimale Ecke (d.h. Ecke von K, die einer Optimallösung des linearen Programms entspricht).
3. Es gibt endlich viele Ecken.

Definition (entartete Ecke) *Sei $K = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ mit $\operatorname{rg} A = m$. x heißt* entartete *(*degenerierte*) Ecke von K, wenn sich mehr als $n - m$ Hyperebenen in x schneiden (d.h. mehr als $n - m$ Komponenten von x 0 sind).*

Bemerkung *Das ist gleichbedeutend damit, dass x eine entartete (zulässige) Basislösung ist.*

Konsequenz Wenn Entartung vorliegt, dann gibt es keine eindeutige Zuordnung zwischen Ecken und Basen. Eine entartete Ecke korrespondiert zu mehreren Basen. Tritt im Simplexverfahren Entartung auf, so steckt man in einer entarteten Ecke fest.

3.2 Geometrische Interpretation des Simplexverfahrens

1. Starte in (zulässiger) Ecke von K.
2. Teste, ob Ecke optimal ist (lokaler Optimumstest). (Das entspricht dem Test, ob es eine benachbarte/adjazente Ecke mit besserem Zielfunktionswert gibt. Die zugehörigen Basislösungen entstehen durch Tausch einer BV gegen eine NBV.)
3. Wenn eine Ecke noch nicht optimal ist, geht man zur Ecke mit besserem (nicht schlechterem im Entartungsfall) ZWF, sonst stoppe.

4 Dualität

4.1 Motivation und Einführung

Beispiel

$$\min x^2 + y^2$$

s.t.

$$2x + y = 5.$$

Lagrangefunktion:

$$L(x, y, \lambda) = x^2 + y^2 + \lambda \cdot (5 - y - 2x).$$

Nun muss gelten:

$$\begin{aligned} L_x &= 0 & \qquad L_x &= 2x - 2\lambda \\ L_y &= 0 & L_y &= 2y - \lambda \\ L_\lambda &= 0 & L_\lambda &= 5 - y - 2x \end{aligned}$$

Gegeben: lineares Programm (*primal*).

Gesucht ist ein anderes lineares Programm (*dual*) mit dem Ziel eines Optimalitätsnachweises à la Lagrange-Multiplikatoren.

Definition (Duales lineares Programm) *Sei*

$$\max c^t x$$

s.t.

$$\begin{aligned} Ax &\le b \\ x &\ge 0 \end{aligned}$$

das primale Problem*. Das dazu* duale lineare Programm *hat folgende Gestalt:*

$$\min b^t y$$

s.t.

$$\begin{aligned} A^t y &\ge c \\ y &\ge 0 \end{aligned}$$

(y: Vektor der dualen Variablen).

Bemerkung *y spielt dabei die Rolle der Lagrangemultiplikatoren für das primale Problem.*

Beispiel *Primales Problem:*

$$\max 4x_1 + 2x_2 - x_3$$

s.t.

$$\begin{aligned} x_1 - 2x_2 + x_3 &\le 3 \\ 2x_1 + 3x_3 &\le 5 \\ -x_2 &\ge -4 \\ x_1, x_2, x_3 &\ge 0. \end{aligned}$$

Duales Problem:

$$\min 3y_1 + 5y_2 + 4y_3$$

s.t.

$$\begin{aligned} y_1 + 2y_2 &\geq 4 \\ -2y_1 + y_3 &\geq 2 \\ y_1 + 3y_2 &\geq -1 \\ y_1, y_2, y_3 &\geq 0 \end{aligned}$$

4.1.1 Direkte Behandlung von Gleichungen

Seien $\alpha \in \mathbb{R}^n, \beta \in \mathbb{R}$.

$$\alpha^t x = \beta$$

entspricht

$$\begin{aligned} \alpha^t x &\leq \beta & (\rightsquigarrow y_1) \\ -\alpha^t x &\leq -\beta & (\rightsquigarrow y_2). \end{aligned}$$

Somit lautet das duale Problem:

$$\min \beta y_1 - \beta y_2$$

s.t.

$$\begin{aligned} \alpha_1 y_1 - \alpha_1 y_2 &\geq c_1 \\ \alpha_2 y_1 - \alpha_2 y_2 &\geq c_2 \\ &\vdots \\ \alpha_n y_1 - \alpha_n y_2 &\geq c_n \\ y_1, y_2 &\geq 0. \end{aligned}$$

Das ist äquivalent zu

$$\min \beta \tilde{y}$$

s.t.

$$\begin{aligned} \alpha_1 \tilde{y} &\geq c_1 \\ &\vdots \\ \alpha_n \tilde{y} &\geq c_n \end{aligned}$$

($\tilde{y}$ nicht vorzeichenbeschränkt).

Eine Gleichung im primalen Problem entspricht also einer nicht vorzeichenbeschränkten Variablen im dualen Problem.

4.1.2 Direkte Behandlung von nicht vorzeichenbeschränkten Variablen

x_j nicht vorzeichenbeschränkt bedeutet

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

Seien $\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$. Betrachte das Problem

$$\max cz$$

s.t.

$$\alpha z \leq \beta$$

(z nicht vorzeichenbeschränkt). Das ist äquivalent zu

$$\max c \cdot (z^+ - z^-)$$

s.t.

$$\begin{aligned} \alpha \cdot (z^+ - z^-) &\leq \beta \\ z^+, z^- &\geq 0. \end{aligned}$$

Das dazu duale Problem lautet:

$$\min \sum_{j=1}^{n} \beta_j y_j$$

s.t.

$$\begin{aligned} \sum \alpha_j y_j &\geq c \\ \sum -\alpha_j y_j &\geq -c \\ y_j &\geq 0. \end{aligned}$$

Eine nicht vorzeichenbeschränkte Variable im primalen Problem entspricht also einer Gleichungsrestriktion als zugehöriger dualer Restriktion.

Beispiel *Das primale Problem*

$$\max x_1 + x_2 + 3x_3$$

s.t.

$$\begin{aligned} 2x_1 + x_2 - x_3 &\leq 5 \\ x_1 - x_2 &\leq 3 \\ x_1 - 2x_2 + 4x_3 &= 1 \\ x_1, x_2 &\geq 0 \end{aligned}$$

entspricht dem dualen Problem

$$\min 5y_1 + 3y_2 + y_3$$

s.t.

$$\begin{aligned} 2y_1 + y_2 + y_3 &\geq 1 \\ y_1 - y_2 - 2y_3 &\geq 1 \\ -y_1 + 4y_3 &= 3 \\ y_1, y_2 &\geq 0. \end{aligned}$$

4.1.3 Zusammenfassung

P (primal)	D (dual)
Zielfunktion $\max c^t x$	rechte Seite c
rechte Seite b	Zielfunktion $\min b^t y$
Koeffizientenmatrix A	Koeffizientenmatrix A^t
i-te Restriktion ist Gleichung	i-te duale Variable ist nicht vorzeichenbeschr.
i-te Restriktion ist $\leq$-Ungleichung	i-te duale Variable ist vorzeichenbeschränkt
j-te Variable ist vorzeichenbeschränkt	j-te duale Restriktion ist $\geq$-Ungleichung
j-te Variable ist nicht vorzeichenbeschr.	j-te duale Restriktion ist Gleichung

Bemerkung *Das duale lineare Programm von (D) ist wieder (P).*

4.2 Interpretation des dualen Problems zum Transportproblem

m Fabriken; a_i: erzeugte Menge in Fabrik i.
n Abnehmer; b_j: Bedarf von Abnehmer j.
Annahme:

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j.$$

c_{ij}: Transportkosten pro Einheit von Fabrik i zu Abnehmer j.
Ziel: minimiere Transportkosten.
Variable x_{ij}: Transportmenge von Fabrik i zu Abnehmer j.
Das Problem lautet also

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

s.t.

$$\begin{aligned} \sum_{j=1}^{n} x_{ij} &= a_i & j &= 1, \ldots, m & &\rightsquigarrow \text{Dualvariable } u_i \\ \sum_{i=1}^{m} x_{ij} &= b_j & j &= 1, \ldots, n & &\rightsquigarrow \text{Dualvariable } v_i \\ x_{ij} &\geq 0 \end{aligned}$$

Duales Problem:

$$\max \sum_{i=1}^{m} a_i u_i + \sum_{j=1}^{n} b_j v_j$$

s.t.

$$u_i + v_j \leq c_{ij}$$

(u, v, nicht vorzeichenbeschränkt).

Bemerkung *x_{ij} kommt in zwei Restriktionen in (P) vor.*

Frage: Bedeutung bzw. Interpretation des Dualproblems?
Setze $\overline{u}_i := -u_i$. Das duale Problem lautet dann

$$\max \sum_{j=1}^{n} b_j v_j - \sum_{i=1}^{m} a_i \overline{u}_i$$

s.t.

$$v_j - \overline{u}_i \leq c_{ij} \quad i = 1, \ldots, m;\ j = 1, \ldots, n.$$

Betrachte (externen) Transporteur: Kauft Waren in Fabriken auf, verkauft sie bei Abnehmern wieder und möchte Gewinn maximieren.
v_j: Preis, den der Transporteur beim Verkauf einer Einheit beim Abnehmer j erhält.
$\overline{u}_i$: Preis, der beim Kauf in Fabrik i zu zahlen ist.
Dann ist

$$\sum_{j=1}^{n} b_j v_j - \sum_{i=1}^{m} a_i \overline{u}_i$$

genau der Gewinn des Transporteurs und

$$v_j - \overline{u}_i \leq c_{ij}$$

eine ökonomische Bedingung, damit der Produzent einwilligt.

4.3 Der Dualitätssatz der linearen Optimierung

Satz 4.1 (Dualitätssatz) *Besitzt eines von zwei zueinander dualen linearen Programmen eine endliche Lösung, so auch das andere, und die optimalen Zielfunktionswerte stimmen überein.*

Satz 4.2 (schwacher Dualitätssatz) *Sei (P)*

$$\max c^t x$$

s.t.

$$\begin{aligned} Ax &\leq b \\ x &\geq 0 \end{aligned}$$

und (D)

$$\min b^t y$$

s.t.

$$\begin{aligned} A^t y &\geq c \\ y &\geq 0. \end{aligned}$$

Seien weiters

$$\begin{aligned} M_P &= \{x \in \mathbb{R}^n \mid Ax \leq b, x \geq 0\}, \\ M_D &= \{y \in \mathbb{R}^m \mid A^t y \geq c, y \geq 0\}. \end{aligned}$$

Dann gilt

$$\forall x \in M_P, y \in M_D : c^t x \leq b^t y.$$

Beweis Sei $x \in M_P$ und $y \in M_D$. Dann gilt

$$Ax \leq b \quad \text{und} \quad x \geq 0$$

und

$$A^t y \geq c \quad \text{und} \quad y \geq 0.$$

Daraus folgt

$$c^t x \leq (A^t y)^t x = y^t A x \leq y^t b = b^t y.$$

■

Satz 4.3 (Korollar) *Sei $x \in M_P$ und $y \in M_D$ mit $c^t x = b^t y$. Dann ist x Optimallösung für (P) und y Optimallösung für (D).*

Satz 4.4 (Existenzsatz)

1. *Haben zwei zueinander duale lineare Programme beide zulässige Lösungen, so haben beide eine endliche Optimallösung und die optimalen Zielfunktionswerte stimmen überein.*
2. *Wenn nur eines der zwei zueinander dualen linearen Programme eine zulässige Lösung hat, dann ist dieses Problem unbeschränkt, d.h. besitzt keine endliche Optimallösung.*
3. *Hat ein Problem zulässige Lösungen, aber keine endliche Optimallösung (also unbeschränkt), so besitzt das duale Problem keine zulässige Lösung.*

Beweis

1. $x \in M_P, y \in M_D$. Dann gilt wegen Satz 4.2
$$c^t x \leq b^t y.$$
Somit ist $\{c^t x \mid x \in M_P\}$ eine nach oben beschränkte Menge. M_P ist abgeschlossen, also wird das Maximum angenommen. Somit besitzt (P) eine endliche Optimallösung. Aus Satz 4.1 folgt, dass (D) eine endliche Optimallösung besitzt und die optimalen Zielfunktionswerte übereinstimmen.

2. Sei $M_P \neq \emptyset$, also $M_D = \emptyset$. Annahme: (P) besitzt eine endliche Optimallösung. Wegen Satz 4.1 besitzt dann auch (D) eine endliche Optimallösung, Widerspruch zu (D) unzulässig.

3. Sei $M_P \neq \emptyset$, (P) unbeschränkt. Annahme: $M_D \neq \emptyset$. Dann folgt aus 1, dass beide Probleme eine endliche Optimallösung bestizen, Widerspruch zu (P) unbeschränkt. ■

Satz 4.5 (Komplementaritätssatz, Satz vom komplementären Schlupf)
Sei $x \in M_P$ und $y \in M_D$. Dann sind folgende beiden Aussagen äquivalent:

1. *x ist optimal für (P) und y ist optimal für (D).*

2.
$$x^t(A^t y - c) = 0 \tag{4.1}$$
$$y^t(Ax - b) = 0 \tag{4.2}$$

Interpretation von (4.1) und (4.2):
$$\underbrace{x^t}_{\geq 0}\underbrace{(A^t y - c)}_{\geq 0} = 0$$
somit gilt für alle $j = 1, \ldots, n$
$$x_j = 0 \quad \text{oder} \quad (A^t y - c)_j = 0,$$
also
$$\sum_{i=1}^{m} a_{ij} y_i - c_j = 0 \Leftrightarrow \sum_{i=1}^{m} a_{ij} y_j = c_j$$
(j-te Restriktion im dualen Problem ist mit Gleichheit erfüllt).

D.h. (4.1) gilt genau dann, wenn für alle $j = 1, \ldots, n$ gilt: Entweder ist j-te primale Variable $x_j = 0$ oder j-te duale Restriktion ist mit Gleichheit erfüllt.

Analog interpretiert man (4.2): $y^t(Ax - b) = 0 \Longleftrightarrow$ Entweder ist die i-te duale Variable $y_i = 0$ oder die i-te primale Restriktion ist mit Gleichheit erfüllt.

Beweis „$\Longrightarrow$“: Sei
$$x \in M_P, \text{ d.h. } Ax \leq b, x \geq 0$$
$$y \in M_D, \text{ d.h. } A^t y \geq c, y \geq 0.$$

Wegen $A^t y \geq c$ und $x \geq 0$ gilt
$$(A^t y)^t x \geq c^t x = b^t y \tag{4.3}$$
(Dualitätssatz, x und y optimal).
$$(A^t y)^t x = y^t A x \leq y^t b = b^t y$$
(da $x \in M_P$, $y \geq 0$).

Insgesamt $y^t(Ax - b) \leq 0$. Aus (4.3) folgt außerdem $y^t(Ax - b) \leq 0$. Somit folgt (4.2).

Sei $x \in M_P, y \in M_D$ mit (4.1).

$$Ax \leq b \Rightarrow (Ax)^t y \leq b^t y = c^t x.$$

Erhalte einmal ≥ 0 und einmal ≤ 0, also $x^t(A^t y - c) = 0$ (4.1).
„$\Longleftarrow$“: Sei $x \in M_P, y \in M_D$, (4.1) und (4.2) erfüllt.

$$x^t(A^t y - c) = y^t(Ax - b) = 0.$$

Also

$$\begin{aligned} x^t A^t y - x^t c &= 0 & \text{und } y^t Ax - y^t b &= 0 \\ \underbrace{x^t A^t y}_{=y^t Ax} &= x^t c = c^t x & y^t Ax &= y^t b = b^t y \end{aligned}$$

und somit

$$c^t x = c^t y,$$

also ist (wegen $x \in M_P, y \in M_D$) x optimal für (P) und y optimal für (D). ∎

Beispiel *Betrachte das primale Problem (P)*

$$\max 7x_1 + 6x_2 + 5x_3 - 2x_4 + 3x_5$$

s.t.

$$\begin{aligned} x_1 + 3x_2 + 5x_3 - 2x_4 + 2x_5 &\leq 4 \\ 4x_1 + 2x_2 - 2x_3 + x_4 + x_5 &\leq 3 \\ 2x_1 + 4x_2 + 4x_3 - 2x_4 + 5x_5 &\leq 5 \\ 3x_1 + x_2 + 2x_3 - x_4 - 2x_5 &\leq 1 \\ x_1, x_2, x_3, x_4, x_5 &\geq 0 \end{aligned}$$

und das dazu duale Problem (D)

$$\min 4y_1 + 3y_2 + 5y_3 + y_4$$

s.t.

$$\begin{aligned} y_1 + 4y_2 1 + 2y_3 + 3y_4 &\geq 7 \\ 3y_1 + 2y_2 + 4y_3 + y_4 &\geq 6 \\ 5y_1 - 2y_2 + 4y_3 + 2y_4 &\geq 5 \\ -2y_1 + y_2 - 2y_3 - y_4 &\geq -2 \\ 2y_1 + y_2 + 5y_3 - 2y_4 &\geq 3 \\ y_1, y_2, y_3, y_4 &\geq 0. \end{aligned}$$

Teste, ob $x = \begin{pmatrix} 0 & \frac{4}{3} & \frac{2}{3} & \frac{5}{3} & 0 \end{pmatrix}^t$ *optimal für (P) ist. Setze dazu x in die primale Restriktion ein:*

$$\begin{aligned} 0 + 4 + \frac{10}{3} - \frac{10}{3} + 0 &= 4 \\ \frac{8}{3} - \frac{4}{3} + \frac{5}{3} &= 3 \\ \frac{16}{3} + \frac{8}{3} - \frac{10}{3} &< 5 \\ \frac{4}{3} + \frac{4}{3} - \frac{5}{3} &= 1, \end{aligned}$$

also $y_3 = 0$.

Betrachte nun (4.1)*:* $x_2, x_3, x_4 > 0$*, daher müssen die 2., 3. und 4. duale Restriktion mit Gleichheit erfültt sein:*

$$\begin{aligned} 3y_1 + 2y_2 + 4y_3 + y_4 &= 6 \\ 5y_1 - 2y_2 + 4y_3 + 2y_4 &= 5 \\ -2y_1 + y_2 - 2y_3 - y_4 &= -2, \end{aligned}$$

also $y_1 = 1, y_2 = 1, y_3 = 0, y_4 = 1$.

Es steht noch der Test aus, ob $y \in M_D$*. Betrachte dazu die 1. und 5. duale Restriktion. Es gilt zwar*

$$1 + 4 + 3 \geq 7,$$

nicht aber

$$2 + 1 - 2 \geq 3.$$

Die Restriktion ist damit verletzt und y *nicht zulässig. Daher ist* x *nicht optimal.*

Satz 4.1 besitzt mehrere Beweismöglichkeiten. Die einfachste und kürzeste benutzt das Simplexverfahren. Wir gehen hier einen längeren Umweg.

4.4 Trennungssätze im $\mathbb{R}^n$

Definition (Trennung) *Die Hyperebene* $H = \{x \mid a^t x = a_0\}$ trennt *die nicht leere Menge* A *von der nicht leeren Menge* B*, wenn*

$$\forall x \in A, y \in B : a^t x \leq a_0 \leq a^t y$$

gilt. Sie trennt echt*, wenn es ein* $\alpha \in \mathbb{R}$ *gibt mit*

$$\forall x \in A, y \in B : a^t x \leq \alpha < a^t y.$$

Satz 4.6 (Trennung eines Punktes von einer konvexen Menge) *Sei* $C \neq \emptyset$*,* $C \subseteq \mathbb{R}^n$*,* C *konvex,* $y \notin \overline{C}$ *(topologischer Abschluss von* C*). Dann gibt es eine Hyperebene* H*, die den Punkt* y *echt von* C *trennt, d.h.*

$$\exists a \in \mathbb{R}^n, a_0 \in \mathbb{R}\ \forall x \in C : a^t x \leq a_0 < a^t y.$$

Beweisskizze Betrachte

$$\inf_{x \in C} \|x - y\| = \|x_0 - y\| =: \delta.$$

Sei $x \in C$ beliebig. Weil C konvex: $x_0 + \lambda(x - x_0) \in \overline{C}$ für ein $0 \leq \lambda \leq 1$.

$$\|x_0 + \lambda(x - x_0) - y\| > \lambda = \|x_0 - y\|,$$

d.h.

$$\|x_0 - y\|^2 + 2\lambda(x - x_0)^t(x_0 - y) + \lambda^2 \|x - x_0\|^2 \geq \|x_0 - y\|^2.$$

Sei $\lambda \neq 0$. Dann folgt

$$2(x - x_0)^t(x_0 - y) + \lambda \underbrace{\|x - x_0\|^2}_{=(x-x_0)^t(x-x_0)} \geq 0$$

für alle λ mit $0 < \lambda \leq 1$.

$$(x - x_0)^t(x_0 - y) \geq 0,$$

d.h.

$$(x_0 - y)^t x \geq (x_0 - y)^t x_0 = (x_0 - y)^t y + (x_0 - y)^t(x_0 - y),$$

somit

$$(x_0 - y)^t x > (x_0 - y)^t y \qquad \forall x \in C.$$

Verwende $x_0 - y$ als Vektor a für die Hyperebene H:

$$H = \{x \mid (x_0 - y)^t = \underbrace{(x_0 - y)^t y}_{=a_0}\}$$

trennt y echt von C. □

Definition *Eine Hyperebene H heißt* Stützhyperebene *(engl. „supporting plain") von C, wenn gilt:*

1. *$H \cap C \neq \emptyset$,*
2. *C liegt ganz auf einer Seite von H.*

Satz 4.7 *Sei $C \subseteq \mathbb{R}^n$ nicht leer und konvex, y ein Randpunkt von C. Dann gibt es eine Stützhyperebene H für C durch y.*

Beweisidee Durch Grenzprozesss aus Satz 4.6. Wähle Folge $\langle y_k \rangle$ mit $y_k \to y$, $y_k \notin \overline{C}$. Sei

$$a_k := x_0(y_k) - y_k$$

die Folge von (normierten) Normalvektoren der wie im Beweis von Satz 4.6 konstruierten trennenden Hyperebenen. Es gilt

$$a_k^t y_k < \inf_{x \in C} a_k^t x.$$

Da $\{a_k\}$ beschränkt ist, gibt es eine konvergente Teilfolge von a_k, deren Grenzwert sei a. Dann gilt

$$a^t y = \lim_{k \to \infty} a_k^t y_k \leq \lim_{k \to \infty} a_k^t x = a^t x$$

für alle $x \in C$. □

Satz 4.8 *Seien $C_1, C_2 \subseteq \mathbb{R}^n$, $C_1, C_2 \neq \emptyset$ und konvex, $C_1 \cap C_2 = \emptyset$, C_2 offen. Dann gibt es eine Hyperebene, die C_1 und C_2 trennt, d.h.*

$$\exists a \in \mathbb{R}^n, \beta \in \mathbb{R}\ \forall x \in C_1, y \in C_2 : a^t x \leq \beta < a^t y.$$

Beweis Sei

$$C_3 = \{y - x \mid y \in C_2, x \in C_1\}.$$

$0 \notin C_3$, da $C_1 \cap C_2 = \emptyset$. C_3 ist konvex.

C_3 ist offen, da

$$C_3 = \bigcup_{x \in C_1} \underbrace{\{y - x \mid y \in C_2\}}_{\text{offen, da } C_2 \text{ offen}}$$

eine Vereinigung von offenen Mengen, also offen ist.

Es gibt also eine Hyperebene H, sodass C_3 ganz in einem von H erzeugten Halbraum liegt, d.h.

$$\exists a \in \mathbb{R}^n : a^t(y - x) \geq 0 = a^t \cdot 0.$$

Setze

$$\beta := \inf_{y \in C_2} a^t y.$$

Dann gilt

$$a^t x \leq \beta < a^t y$$

($<$ weil C_2 offen). ■

4.5 Alternativsätze

Satz 4.9 *Entweder besitzt das Gleichungssystem*

$$Ax = b \tag{4.4}$$

eine Lösung, oder das Gleichungssystem

$$\begin{aligned} A^t y &= 0 \\ b^t y &= 0 \end{aligned} \tag{4.5}$$

besitzt eine Lösung (aber nicht beide gleichzeitig).

BEWEIS

1. Wir zeigen, dass (4.4) und (4.5) nicht zugleich lösbar sind. Beweis durch Widerspruch. Angenommen, es gibt x und y mit

   $$\begin{aligned} Ax &= b \\ A^t y &= 0 \\ b^t y &= 1 \end{aligned}$$

   Daraus folgt
   $$1 = b^t y = (Ax)^t y = x^t A^t y = 0,$$
   Widerspruch.

2. Wenn $Ax = b$ nicht lösbar ist, so ist $A^t y = 0, b^t y = 1$ lösbar: $Ax = b$ nicht lösbar heißt, dass b linear unabhängig von den Spalten von A ist. Sei $\operatorname{rank} A = r$. Dann gilt

   $$\begin{aligned} \operatorname{rank} \begin{pmatrix} Ab \end{pmatrix} &= r + 1 \\ \operatorname{rank} \begin{pmatrix} A^t \\ b^t \end{pmatrix} &= r + 1 \\ \operatorname{rank} \begin{pmatrix} A^t & 0 \\ b^t & 1 \end{pmatrix} &= r + 1. \end{aligned}$$

   Letztere ist genau die erweiterte Koeffizientenmatrix von (4.5).

3. Analog sieht man, dass (4.4) lösbar ist, wenn (4.5) nicht lösbar ist. ∎

Satz 4.10 *Entweder besitzt das System*

$$\begin{aligned} Ax &= b \\ x &\geq 0 \end{aligned} \tag{4.6}$$

eine Lösung, oder das System

$$\begin{aligned} A^t y &\geq 0 \\ b^t y &< 0 \end{aligned} \tag{4.7}$$

besitzt eine Lösung, aber nie beide gleichzeitig.

Bemerkung *Geometrische Interpretation:*

$$K = \{z \in \mathbb{R}^n \mid z = Ax, z \geq 0\}$$

ist Kegel (der von den Spalten von A aufgespannte Kegel).

- (4.6) *hat Lösung* $\Leftrightarrow$ *b liegt im Kegel K*
- (4.7) *hat Lösung* $\Leftrightarrow$ *es gibt eine Hyperebene durch 0, die den Kegel von b trennt*

BEWEIS

1. (4.6) und (4.7) sind nicht zugleich lösbar. Beweis durch Wiederspruch. Angenommen es gibt x und y mit
$$Ax = b, x \geq 0 \quad \text{und} \quad A^t y \geq 0, b^t < 0.$$
Dann gilt
$$0 > b^t y = (Ax)^t y = \underbrace{x^t}_{\geq 0} \underbrace{A^t y}_{\geq 0} \geq 0,$$
Widerspruch.

2. Wir zeigen, dass wenn $Ax = b$ keine Lösung hat (somit (4.6) auch keine Lösung hat), (4.7) $A^t Y \geq 0, b^t y < 0$ eine Lösung hat.

   $Ax = b$ hat keine Lösung, somit hat
$$A^t y = 0, b^t y = 1$$
lt. Satz 4.9 eine Lösung. Sei $\hat{y}$ eine solche Lösung, also $A^t \hat{y} = 0, b^t \hat{y} = 1$. Betrachte $\hat{\hat{y}} = -\hat{y}$. Es gilt
$$A^t \hat{\hat{y}} = 0 \geq 0 \quad \text{und} \quad b^t \hat{\hat{y}} = -1 < 0,$$
$\hat{\hat{y}}$ ist also eine Lösung von (4.7).

3. (4.6) hat keine Lösung, aber $Ax = b$ ist lösbar (sonst siehe Fall 2), d.h. jede Lösung x von $Ax = b$ hat mindestens eine negative Komponente.

   Sei
$$K := \{z \in \mathbb{R}^n \mid z = Ax, x \geq 0\},$$
$b \notin K$ (da (4.6) nicht lösbar). K ist abgeschlossen und konvex (nicht leer).

   Idee: Verwende Trennungssatz (Satz 4.8). Setze
$$C_1 := K, C_2 := \{\alpha u \mid \|a - b\| < \varepsilon, \alpha > 0\}.$$
$\varepsilon > 0$ wird hinreichend klein gewählt. C_2 ist konvex und offen, $C_1 \cap C_2 = \emptyset$. Es gibt also eine Hyperebene H, die C_1 und C_2 trennt. Sei
$$H := \{y \mid a^t y = \beta\}.$$
Sei $z_1 \in C_1$ und $z_2 \in C_2$. Dann gilt
$$a^t z_1 \leq \beta \leq a^t z_2 \qquad \text{für alle } z_1 \in C_1, z_2 \in C_2 \tag{4.8}$$
(Trennungseigenschaft). $a^t z_2 = a^t \alpha u$, wobei α beliebig klein werden kann. Daher muss $\beta = 0$ sein. Dann folgt aus (4.8)
$$\forall z_1 \in C_1 : a^t z_1 \leq 0.$$
Wähle $y := -a$. Dann gilt $y^t z \geq 0$ für alle $z \in C_1$.
$$b^t y = y^t b < 0,$$
da b nicht im Kegel K liegt. Insgesamt heißt das, dass y eine Lösung für
$$A^t y \geq 0$$
$$b^t y < 0$$
ist. (4.7) besitzt also eine Lösung. ∎

Satz 4.11 (Lemma von Farkas) *Folgende zwei Aussagen sind äquivalent:*

1. *Für alle y mit $A^t y > 0$ gilt $b^t y \geq 0$.*
2. *Es gibt ein $x \geq 0$ mit $Ax = b$.*

Beweis Wir zeigen, dass aus Satz 4.10 Satz 4.11 folgt.
Wenn (2) gilt, hat (4.6) laut Satz 4.10 eine Lösung und (4.7) nicht. Somit gibt es keine Lösung für

$$A^t y \geq 0, \quad b^t y < 0.$$

Daher gilt für alle y mit $A^t y \geq 0$

$$b^t y \geq 0.$$

Also gilt (1).
Wenn (2) nicht gilt, hat (4.6) laut Satz 4.10 keine Lösung. Somit hat (4.7) eine Lösung, also gibt es ein y mit

$$A^t y \geq 0, \quad b^t y < 0.$$

Somit gilt (1) nicht.
Unter Verwendung von Satz 4.10 haben wir also gezeigt, dass (1) $\Leftrightarrow$ (2). ■

Bemerkung *Satz 4.10 und Satz 4.11 sind äquivalent.*

Beweis Noch zu zeigen ist, dass aus Satz 4.11 Satz 4.10 folgt.
Wenn (2) gilt, gilt auch (1), also ist (4.7)

$$A^t y \geq 0, \quad b^t y < 0$$

nicht lösbar.
Gilt (2), d.h. (4.7)

$$Ax = b, \quad x \geq 0$$

hat keine Lösung, gilt (1) auch nicht. Somit gibt es ein y mit

$$A^t y \geq 0, \quad b^t y < 0,$$

d.h. (4.7) ist lösbar. ■

Satz 4.12 (Tucker) *Die Systeme $A^t y \geq 0$ und $Ax = 0, x \geq 0$ besitzen Lösungen $\tilde{x}$ und $\tilde{y}$ mit*

$$A^t \tilde{y} + \tilde{x} > 0 \qquad \textit{(komponentenweise).}$$

Beweis Bezeichne mit $a^{(j)}$ die j-te Spalte von A.

$$Ax = 0, x \geq 0 \Rightarrow \sum_{j \neq k} a^{(j)} x_j = -a^{(k)} x_k, x_k \geq 0, k = 1, \ldots, n.$$

Betrachte folgende Systeme:

$$\begin{aligned} \sum_{j \neq k} a^{(j)} x_j &= \underbrace{-a^{(k)}}_{\hat{b}} \qquad x_j \geq 0, j = 1, \ldots, n, j \neq k \\ \hat{A}\hat{x} &= \hat{b} \end{aligned} \tag{4.9}$$

$$\begin{aligned} \left(a^{(j)}\right)^t y \geq 0 &\Leftrightarrow \hat{A}^t y \geq 0 \qquad j = 1, \ldots, n, j \neq k \\ \left(a^{(k)}\right)^t y > 0 &\Leftrightarrow -\hat{b}^t y < 0 \end{aligned} \tag{4.10}$$

Wende Satz 4.10 an. Somit ist entweder das System (4.9) oder das System (4.10) lösbar. Sei nun $Z_1 \subseteq \{1, \ldots, n\}$ die Menge aller k, für die (4.9) eine Lösung hat (analog Z_2 für (4.10)). Es gilt

$$Z_1 \cup Z_2 = \{1, \ldots, n\}, \quad Z_1 \cap Z_2 = \emptyset.$$

Sei (4.9) lösbar für $k \in \{1, \ldots, n\}$, also $k \in Z_1$. Es gibt ein $\tilde{x}^{(k)}$ mit

$$\begin{aligned} \sum_{j \neq k} a^{(j)} \hat{x}_j^{(k)} &= -a^{(k)} \qquad j = 1, \ldots, n, j \neq k \\ \hat{x}_j^{(k)} &\geq 0. \end{aligned}$$

Setze $\hat{x}_k^{(k)} = 1$ (≥ 0), somit

$$A\hat{x}^{(k)} = 0, \hat{x}^{(k)} \geq 0.$$

Sei nun (4.10) lösbar für k, also $k \in Z_2$. Dann existiert ein $\hat{y}^{(k)}$ mit

$$\begin{aligned} A^t \hat{y}^{(k)} &\geq 0 \\ \left(a^{(k)}\right)^t \hat{y}^{(k)} &> 0. \end{aligned}$$

Setze

$$\tilde{x} := \begin{cases} 0 & \text{falls } Z_1 = \emptyset \\ \sum_{k \in Z_1} \hat{x}^{(k)} & \text{sonst} \end{cases}$$

und

$$\tilde{y} := \begin{cases} 0 & \text{falls } Z_2 = \emptyset \\ \sum_{k \in Z_2} \hat{y}^{(k)} & \text{sonst} \end{cases}.$$

Es bleibt zu zeigen:

1. $A^t \tilde{y} \geq 0$
2. $A\tilde{x} = 0$, $\tilde{x} \geq 0$.
3. $A^t \tilde{y} + \tilde{x} > 0$

Beweis:

1.

$$A^t \tilde{y} = \begin{cases} 0 \geq 0 & Z_2 = \emptyset \\ \sum_{k \in Z_2} A^t \tilde{y}^{(k)} \geq 0 & \text{sonst} \end{cases}$$

aufgrund der Konstruktion.

2.

$$A^t \tilde{y} = \begin{cases} 0 \geq 0 & Z_1 = \emptyset \\ \sum_{k \in Z_1} A\tilde{x}^{(k)} \geq 0 & \text{sonst} \end{cases} = 0.$$

$\tilde{x} \geq 0$ analog.

3. Annahme: $A^t \tilde{y} + \tilde{x}$ habe Nullkomponente. Es gibt also eine Komponente q von $A^t \tilde{y} + \tilde{x}$, sodass die q-te Komponente von $\tilde{x}$ und die q-te Komponente von $A^t \tilde{y}$ gleich 0 sind. Widerspruch zur Konstruktion von Z_1 von Z_2. ■

Satz 4.13 (Tucker) *Sei A eine reelle schiefsymmetrische Matrix ($A^t = -A$). Dann gibt es ein $w \in \mathbb{R}^n$, sodass*

$$\begin{aligned} Aw &\geq 0 \\ w &\geq 0 \\ Aw + w &> 0 \quad \textit{(komponentenweise).} \end{aligned}$$

Beweis Betrachte

$$\begin{aligned} \begin{pmatrix} I \\ A \end{pmatrix} y &\geq 0 \\ \begin{pmatrix} I & -A \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} &= 0 \\ \begin{pmatrix} x \\ z \end{pmatrix} &\geq 0. \end{aligned}$$

Wende Satz 4.12 an: Es gibt $\overline{y}, \overline{x}, \overline{z}$ mit

$$\begin{aligned} \overline{y} &\geq 0 \\ A\overline{y} &\geq 0 \\ \overline{x} - A\overline{z} &= 0 \\ \overline{x} &\geq 0 \\ \overline{z} &\geq 0 \end{aligned}$$

und

$$\begin{pmatrix} I \\ A \end{pmatrix} \overline{y} + \begin{pmatrix} \overline{x} \\ \overline{z} \end{pmatrix} > 0.$$

Das heißt

$$\begin{aligned} \overline{y} + \overline{x} &> 0 \\ A\overline{y} + \overline{z} &> 0. \end{aligned}$$

Daraus folgt

$$\overline{x} + \overline{y} + A\overline{y} + \overline{z} > 0. \tag{4.11}$$

Wegen $\overline{x} = A\overline{z}$ folgt

$$A\overline{z} + \overline{y} + A\overline{y} + \overline{z} > 0,$$

also

$$(\overline{y} + \overline{z}) + A(\overline{y} + \overline{z}) > 0.$$

Setze $w := \overline{y} + \overline{z}$. Dann gilt

$$Aw + w > 0.$$

Ferner gilt $Aw \geq 0$ ($A\overline{y} \geq 0, A\overline{z} = \overline{x} \geq 0$), $w \geq 0$ ($\overline{y}, \overline{z} \geq 0$). ∎

4.6 Beweis des Dualitätssatzes

$$\max\{c^t x \text{ s.t. } Ax \leq b, x \geq 0\} \tag{4.12}$$

$$\min\{b^t y \text{ s.t. } A^t y \geq c, y \geq 0\} \tag{4.13}$$

Aus schwachem Dualitätssatz folgt

$$c^t x \leq b^t y \quad \text{für alle} x \in M_P, y \in M_D.$$

Definiere

$$D := \begin{pmatrix} 0_n & A^t & -c \\ -A & 0_m & b \\ c^t & -b^t & 0 \end{pmatrix}.$$

D ist schiefsymmetrisch. Weiters

$$w = \begin{pmatrix} \overline{x} \\ \overline{y} \\ t \end{pmatrix}.$$

Wende Satz von Tucker (4.13) an: Es gibt also eine Lösung w mit

1.
$$\begin{aligned} \overline{x} &\geq 0 \\ \overline{y} &\geq 0 \\ t &\geq 0 \end{aligned}$$
(aus $w \geq 0$)

2.
$$\begin{aligned} A^t\overline{y} - tc &\geq 0 \\ -A\overline{x} + tb &\geq 0 \\ c^t\overline{x} - b^t\overline{y} &\geq 0 \end{aligned}$$
(aus $Dw \geq 0$)

3.
$$\begin{aligned} A^t\overline{y} - tc + \overline{x} &> 0 \\ -A\overline{x} + tb + \overline{y} &> 0 \\ c^t\overline{x} - bt\overline{y} + t &> 0 \end{aligned}$$
(aus $Dw + w > 0$)

Fallunterscheidung:

- $t > 0$: Satz 4.14 bzw.
- $t = 0$: Satz 4.15

Satz 4.14 *Ist $t > 0$, so gibt es Optimallösungen x^* von* (4.12) *und y^* von* (4.13) *mit*

$$\begin{aligned} C^t x^* &= b^t y^* \\ A^t y^* + x^* &> c \\ Ax^* - y^* &< b. \end{aligned}$$

Beweis Sei

$$\begin{aligned} x^* &= \frac{\overline{x}}{t} \\ y^* &= \frac{\overline{y}}{t}. \end{aligned}$$

Da $\overline{x}, \overline{y} \geq 0$ und $t > 0$ folgt

$$x^* \geq 0$$
$$y^* \geq 0.$$

Es gilt

$$A^t\overline{y} - tc \geq 0$$

(vgl. 2), somit

$$A^t y^* \geq c.$$

Ebenso

$$-A\overline{x} + tb \geq 0 \Rightarrow Ax^* \leq b.$$

Also $x^* \in M_P$ und $y^* \in M_D$. Ferner gilt

$$c^t\overline{x} - b^t\overline{y} \geq 0$$
$$c^t x^* - b^t y^* \geq 0, \text{ d.h. } c^t x^* \geq b^t y^*,$$

somit

$$c^t x^* = b^t y^*$$

(schwacher Dualitätssatz). x^* ist also optimal für (4.12) und y^* ist optimal für (4.13).

Die beiden anderen Aussagen folgen aus 3:

$$A^t\overline{y} - tc + \overline{x} > 0 \Rightarrow A^t y^* - c + x^* > 0 \Leftrightarrow A^t y^* + x^* > c$$

bzw.

$$-A\overline{x} + tb + \overline{y} > 0 \Rightarrow -Ax^* + b + y^* > 0 \Leftrightarrow Ax^* - y^* < b.$$

■

Satz 4.15 *Ist $t = 0$, so gilt*

1. (4.12) *oder* (4.13) *besitzt keine zulässigen Lösungen.*
2. *Ist $M_P \neq \emptyset$, so ist* (4.12) *unbeschränkt.*
   Ist $M_D \neq \emptyset$, so ist (4.13) *unbeschränkt.*
3. *Keines der beiden Probleme besitzt eine Optimallösung.*

Beweis 1. Annahme $M_P \neq \emptyset, M_D = \emptyset$. Es gibt also $x \in M_P, y \in M_D$. Aus (3) folgt

$$c^t\overline{x} - b^t y + \underbrace{t}_{=0} > 0 \Rightarrow x^t\overline{x} > b^t\overline{y}.$$

Weiters gilt

$$c^t\overline{x} \leq \left(A^t y\right)^t \overline{x} = \underbrace{y^t}_{\geq 0}\underbrace{A\overline{x}}_{\leq 0},$$

somit

$$b^t\overline{y} < c^t\overline{x} \leq 0.$$

Andererseits gilt

$$\underbrace{x^t}_{\geq 0 \text{ da } x \in M_P} \underbrace{A^t\overline{y}}_{\geq 0} \geq 0. \tag{4.14}$$

Außerdem

$$x^t A^t \overline{y} = \underbrace{(Ax)}_{\leq b}{}^t\, \overline{y} \leq b^t\overline{y},$$

also

$$b^t y \geq 0,$$

Widerspruch.

2. Sei $M_P \neq \emptyset$, also gibt es ein $x \in M_P$.
   Behauptung:
   $$\tilde{x}(\lambda) := x + \lambda \overline{x} \in M_P$$
   für beliebige $\lambda \geq 0$. („Wollen unendliche Menge, damit unbeschränkt.")
   Beweis:
   $$A\tilde{x}(\lambda) = \underbrace{Ax}_{\leq b} + \lambda \underbrace{A\overline{x}}_{\leq 0} \leq b$$
   und
   $$\tilde{x}(\lambda) = \underbrace{x}_{\geq 0} + \underbrace{\lambda}_{\geq 0} \underbrace{\overline{x}}_{\geq 0} \geq 0.$$
   Betrachte Zielfunktionswert von $\tilde{x}(\lambda)$:
   $$c^t \tilde{x}(\lambda) = c^t x + \underbrace{\lambda}_{\geq 0} c^t \overline{x}.$$
   Aus (3) folgt
   $$c^t \overline{x} > b^t \overline{y}.$$
   Verwende (4.14) zusammen mit $M_P \neq \emptyset$:
   $$\overline{b}\overline{y} \geq 0,$$
   somit $c^t \overline{x} > 0$, also wird $c^t \tilde{x}(\lambda)$ beliebig groß, wenn λ wächst. Also ist (4.12) unbeschränkt. Analog für (4.13).
3. Folgt aus Rest. ■

Die Aussagen aus den Sätzen 4.14 und 4.15 vervollständigen den Beweis von Satz 4.1.

5 Duales Simplexverfahren

$(P) : \max\{c^t x : Ax \leq b, x \geq 0\}$
$(D) : \min\{b^t y : A^t y \geq c, y \geq 0\}$
In Normalform: $\max\{-b^t y : -A^t y \leq -c, y \geq 0\}$.
Tableau:

$(-1) \cdot$ ZFW	$\bar{c}$: reduzierte Kostenkoeffizienten
$\tilde{b}$: Werte der Basisvar.	$\tilde{A}_N$

- Tableau (primal) zulässig: $\tilde{b} \geq 0$
- (primal) optimal: $\bar{c} \leq 0$
- dual zulässig: $-\bar{c} \geq 0 \Leftrightarrow \bar{c} \leq 0$
- dual optimal: $\tilde{b} \geq 0$

x ist optimal für (P), wenn x primal und dual zulässig.
Primales Simplexverfahren:

- Starte mit primal zulässiger Basislösung.
- Schreite von (primal zulässiger) Basislösung zu (primal zulässiger) Basislösung, solange Verbesserung möglich (d.h. nicht primal optimal).

Duales Simplexverfahren:

- Ersetze *primal* durch *dual*
- Umsetzung?

Beispiel

$$\max -x_1 - 2x_2$$

s.t.

$$\begin{aligned} x_1 + x_2 &\geq 3 \\ x_2 &\geq 2 \end{aligned}$$

$$\begin{aligned} -x_1 + x_2 &\leq 3 \\ x_1 - x_2 &\leq 3 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Starttableau

	x_1	x_2	
0	-1	-2	
-3	-1	-1	x_3
-2	0	-1	x_4
3	-1	1	x_5
3	1	-1	x_6

nicht primal zulässig, aber dual zulässig!
Gegeben: dual zulässiges Tableau ($\bar{c} < 0$). 2 Möglichkeiten:

1. $\tilde{b} \geq 0$: *Tableau optimal (primal und dual optimal), Stop.*
2. $\tilde{b} \ngeq 0$: *Es gibt* i *mit* $\tilde{b}_i < 0$, *Kandidaten für Pivotzeile (vgl. Auswahlregeln im primalen Simplexverfahren). Wähle so eine Pivotzeile* r *mit* $\tilde{b}_r < 0$.

Wahl der Pivotspalte:

- *Unbeschränktes Problem liegt vor, wenn es eine Zeile* r *mit* $\tilde{b}_r < 0$ *im Tableau gibt mit* $\tilde{a}_{rj} \geq 0$ *für alle* j.
- *Andernfalls berechne*

$$\varepsilon^* := \min\left\{\frac{-\overline{c}_j}{-\tilde{a}_{rj}} : -\tilde{a}_{rj} > 0\right\} = \min\left\{\frac{\overline{c}_j}{\tilde{a}_{rj}} : \tilde{a}_{rj} < 0\right\}.$$

Wähle als Pivotspalte eine Spalte s *mit* $\frac{\overline{c}_s}{\tilde{a}_{rs}} = \varepsilon^*$.

Im Beispiel:

$$\varepsilon^* = \left\{\frac{-1}{-1}, \frac{-2}{-1}\right\} = 1,$$

$r = 1, s = 1$.
Neues Tableau:

$$\overline{a}_i := \begin{cases} \frac{1}{\tilde{a}_{rs}} & i = r, j = s \\ \frac{\tilde{a}_{rj}}{\tilde{a}_{rs}} & i = r, j \neq s \text{ (Pivotzeile)} \\ -\frac{\tilde{a}_{is}}{\tilde{a}_{rs}} & i \neq r, j = s \text{ Pivotspalte)} \\ \tilde{a}_{ij} - \frac{\tilde{a}_{is}\tilde{a}_{rj}}{\tilde{a}_{rs}} & i \neq r, j \neq s \end{cases}$$

	x_3	x_2	
3	-1	-1	
3	-1	1	x_1
-2	0	-1	x_4
0	-1	2	x_5
0	1	-2	x_6

Typische Anwendungen:

- Hinzunehmen neuer Restriktionen (z.B. in Schnittebenenverfahren, „cutting plane methods“)
- Behandlung von Variablen mit oberen Schranken

6 Innere Punkte Methoden (interior point methods)

Hier: eine Variante der primal-dualen Pfadverfolgungsmethode.

- Simplexverfahren:
  - Vorteil: in Praxis recht gut
  - Nachteil: keine Pivotregel bekannt mit polynomialem Gesamtaufwand
- Ellipsoidverfahren:
  - Vorteil: polynomialer Laufzeit für lineare Programme im worst case
  - Nachteil: praktisch unbrauchbar
- Innere Punkte Verfahren: versuchen Vorteile der beiden obigen Verfahren zu kombinieren

Philosophie hinter inneren Punkteverfahren: arbeiten im Inneren des zulässigen Bereichs.
Betrachte das primale Problem (P)

$$\min\{c^t x : Ax = b, x \geq 0\}$$

und das dazu duale Problem (D)

$$\max\{b^t y : A^t y \leq c\} = \min\{b^t y : A^t y + z = c, z \geq 0\}.$$

Inneres: $x_j > 0\ \forall j$ und $z_j > 0\ \forall j$.
Idee: Einführung eines Strafterms. Für (P) definiere

$$H(x) = \underbrace{c^t x}_{\text{ZF von (P)}} - \underbrace{\underbrace{\mu}_{>0} \underbrace{\sum_{j=1}^{n} \ln x_j}_{\ln \prod_{j=1}^{n} x_j}}_{\text{Strafterm}}$$

und für (D)

$$b^t + \mu \sum_{i=1}^{n} \ln z_i.$$

Beobachtung: $x_j \to 0 \Longrightarrow H(x) \to \infty$.
Betrachte also die Probleme

$$\min c^t x - \mu \sum_{j=1}^{n} \ln x_j$$

s.t.

$$Ax = b$$

und

$$\max b^t y + \mu \sum_{i=1}^{n} \ln z_i$$

s.t.

$$A^t y + z = c.$$

Stelle Lagrangefunktionen auf:

$$L_P(x, y, \mu) := c^t x - \mu \sum_{i=1}^{n} \ln x_i + \underbrace{y^t (Ax - b)}_{\text{Vektor der Lagrangemult.}}$$

$$L_D(x, y, \mu) := b^t y + \mu \sum_{i=1}^{n} \ln z_i - x^t (A^t y + z - c)$$

Bemerkung *Bei einem Paar von zueinander dualen linearen Programmen entsprechen die Lagrange-Multiplikatoren des einen Programms den Variablen des dazu dualen Programms.*

Betrachte die partiellen Ableitungen der Lagrangefunktionen:

$$\frac{\partial L_P}{\partial x_i} = c_i - \frac{\mu}{x_i} - y^t \underbrace{A_i}_{i\text{-te Spalte}} \stackrel{!}{=} 0 \qquad i = 1, \ldots, n \tag{6.1}$$

$$\frac{\partial L_P}{\partial y} = (Ax - b) \stackrel{!}{=} 0 \tag{6.2}$$

$$\frac{\partial L_D}{\partial z_i} = \frac{\mu}{z_i} - x_i \stackrel{!}{=} 0 \tag{6.3}$$

$$\frac{\partial L_D}{\partial x} = (A^t y + z - c) \stackrel{!}{=} 0. \tag{6.4}$$

Aus (6.4) folgt

$$c = A^t y + z. \tag{6.5}$$

Aus (6.3) folgt

$$x_i z_i = \mu \quad \forall i = 1, \ldots, n.$$

Sei

$$X = \operatorname{diag}(x_1, \ldots, x_n), \quad Z = \operatorname{diag}(z_1, \ldots, z_n), \quad Z^{-1} = \operatorname{diag}(\tfrac{1}{z_1}, \ldots, \tfrac{1}{z_1})$$

und

$$e = \begin{pmatrix} 1 & \ldots & 1 \end{pmatrix}^t,$$

also

$$Z^{-1} \cdot e = \begin{pmatrix} \frac{1}{z_1} \\ \vdots \frac{1}{z_n} \end{pmatrix}.$$

(6.1) in Vektorform heißt dann

$$c - \mu X^{-1} e - A^t y = 0. \tag{6.6}$$

Einsetzen von (6.5) in (6.6) ergibt

$$A^t y + z - \mu X^{-1} e - A^t y = 0,$$

also

$$z = \mu X^{-1} e \Leftrightarrow XZ = \mu I \Leftrightarrow x_i z_i = \mu \quad \text{für } i = 1, \ldots, n.$$

Es ergeben sich also 3 resultierende Bedingungen (S):

$$Ax = b \qquad (x \geq 0) \tag{6.7}$$

$$A^t y + z = c \qquad (z \geq 0) \tag{6.8}$$

$$x_j z_j = \mu \qquad \forall j \tag{6.9}$$

(6.9) mit $\mu = 0$ heißt *verallgemeinerte Komplementaritätsbedingung.*

Bemerkung *Für $\mu = 0$ beschreibt (S) die Bedingungen des Satzes vom komplementären Schlupf.*

Für $\mu \to 0$ zeigt sich, dass eine Lösung (x, y, z) von (S) gegen Optimallösung x von (P) und Optimallösung y von (D) strebt.

Wir interessieren uns für (S) für beliebige $\mu > 0$.
Behauptung: Für festes μ kann (S) nach x, y und z aufgelöst werden.

Definition (zentraler Pfad)

$$\Gamma := \{(x(\mu), y(\mu), z(\mu)) : x(\mu), y(\mu), z(\mu) \text{ löst (S) für } \mu, \mu > 0\}.$$

Satz 6.1 *Es existieren x, y, z mit $x > 0, Ax = b$ und $z > 0, A^t y + z = c$. Dann hat das System*

$$\begin{aligned} Ax &= b \\ A^t y + z &= c \\ XZe &= \mu e \end{aligned}$$

für jedes $\mu > 0$ eine Lösung $(x(\mu), y(\mu), z(\mu))$, wobei $x(\mu)$ und $z(\mu)$ eindeutig bestimmt sind. Hat A vollen Rang, so ist auch $y(\mu)$ eindeutig.

Beweis FEHLT. ■

6.1 Grundidee der primal-dualen Pfadverfolgungsmethode

Geg. Sei (x, y, z) mit $x > 0, z > 0$ (beliebiger innerer Punkt). (Anmerkung: Zulässigkeit nicht erforderlich, $Ax \neq b$ erlaubt und $Ay + z \neq c$ erlaubt.)

1. Wähle μ „geeignet“ (siehe später).

2. Berechne $(\Delta x, \Delta y, \Delta z)$, sodass $(x + t\Delta x, y + t\Delta y, z + t\Delta z)$ in der Nähe des zentralen Pfads liegt.

   $(x(\mu), y(\mu), z(\mu))$

   t: Schrittweite so wählen, dass $x + t\Delta x > 0, z + t\Delta z > 0$.

3. $$\begin{aligned} x_{\text{neu}} &:= x + t\Delta x \\ y_{\text{neu}} &:= y + t\Delta y \\ z_{\text{neu}} &:= z + t\Delta z \end{aligned}$$

   und gehe wieder an Anfang.

6.1.1 Details

1. Am zentralen Pfad gilt $x_j z_j = \mu$ für alle $j = 1, \ldots, n$.

   $$x^t z = n\mu.$$

   Am zentralen Pfad gilt $\mu = \frac{x^t z}{n}$.

   Damit Konvergenz von μ gegen 0 schneller fortschreitet, wird Faktor $0 < \delta < 1$ verwendet. Wähle $\mu := \delta \cdot \frac{x^t z}{n}$ (häufig wird $\delta = \frac{1}{10}$ verwendet).

2. Geg. $x, y, z; \mu$.
   Ges. $\Delta x, \Delta y, \Delta z$ sodass

$$
\begin{aligned}
A(x+\Delta x) &= b \\
A^t(y+\Delta y) + (z+\Delta z) &= c(x_j+\Delta x_j)(z_j+\Delta z_j) & &= \mu \quad \text{für alle } j,
\end{aligned}
$$

   d.h.

$$(X+\Delta X)(Z+\Delta Z)e = \mu e$$

   mit $\Delta X = \operatorname{diag}(\Delta x_1, \dots, \Delta x_n)$, $\Delta Z = \operatorname{diag}(\Delta z_1, \dots, \Delta z_n)$.
   Aus (1) folgt

$$A\Delta x = \underbrace{b - Ax}_{\rho}.$$

   Aus (2) folgt

$$A^t\Delta y + \Delta z = \underbrace{c - A^t y - z}_{\sigma}.$$

   In (3) wird nichtlinearer Term weggelassen: $x_j z_j + z_j \Delta x_j + x_j \Delta z_j = \mu$ für alle j, d.h.

$$Z\Delta x + X\Delta z = \mu e - Xz.$$

   Resultierendes lineares Gleichungssystem (in Blockform):

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} \rho \\ \sigma \\ \mu e - Xz \end{pmatrix}.$$

   Einsetzen in 1. Block:

$$A\Delta x = \mu A Z^{-1} e - Ax - AXZ^{-1}\sigma + AXZ^{-1}A^t\Delta y \overset{!}{=} \rho.$$

   (Zugrundeliegend: Iteration des Newton-Verfahrens.)
   Schrittweitenwahl:

$$
\begin{aligned}
t &< \min\left\{-\frac{x_i}{\Delta x_i} \mid \Delta x_i < 0\right\} \\
t &< \min\left\{-\frac{z_i}{\Delta z_i} \mid \Delta z_i < 0\right\}
\end{aligned}
$$

   Typischerweise wählt man Verkleinerungsfaktor β mit $0 < \beta < 1$.

6.1.2 Verfahren im Detail

1. Start mit (x, y, z) beliebig mit $x > 0, z > 0$.
2. Setze

$$
\begin{aligned}
\rho &:= b - Ax \\
\sigma &:= c - A^t y - z \\
\mu &:= \rho \frac{x^t z}{h} \qquad \left(0 < \delta < 1, \text{ z.B. } \delta = \frac{1}{10}\right)
\end{aligned}
$$

3. Berechne Δy als Lösung von

$$(AXZ^{-1}A^t)\delta y = b - \mu A Z^{-1} e + AXZ^{-1}\sigma.$$

4. Berechne $\Delta z = \sigma - A^t \Delta y$.
5. Berechne Δx als
$$\Delta x = \mu Z^{-1} e - x - X Z^{-1} \Delta z.$$
6. Bestimme
$$\alpha_P := \min\{-\frac{x_i}{\Delta x_i} \mid \Delta x_i < 0\}$$
$$\alpha_D := \min\{-\frac{z_i}{\Delta z_i} \mid \Delta z_i < 0\}.$$
Setze $\alpha_P := 1$ (bzw. $\alpha_D := 1$), falls Auswahlmenge leer.
7. Setze
$$x_{\text{neu}} := x + \beta\alpha_P \Delta x$$
$$y := y + \beta\alpha_D \Delta y$$
$$z := z + \beta\alpha_D \Delta z$$
8. Gilt $\|X\|_\infty \geq K$ (K geeignet groß): Abbruch ((P) wird als unbeschränkt erklärt).
Gilt $\|Y\|_\infty \geq K$: Abbruch ((D) wird als unbeschränkt erklärt).
Gilt $\|\rho\|_1 < \varepsilon$ (Maß für primale Unzulässigkeit), $\|\sigma\|_1 < \varepsilon$ (Maß für duale Unzulässigkeit) und $x^t z < \varepsilon$ (Maß für Entfernung von Optimallösung für (P) und (D)).

Beispiel

$$\min x_1 + 2x_2 + 4x_3$$

s.t.

$$2x_1 + x_2 + x_3 \geq 2$$
$$x_1, x_2, x_3 \geq 0$$

Start: $x_1 = 1, x_2 = 1, x_3 = 1, z_1 = 1, z_2 = 1, z_3 = 1, y = 0, b = \frac{1}{10}$. $\beta = 0.995$.

x_1	x_2	x_3	z_1	z_2	z_3	y	μ
1	1	1	1	1	1	0	0.1
1.3255	0.6745	0.0010	0.3714	1.0224	1.6959	0.3143	0.0395
1.1981	0.0007	0.0156	0.1717	1.3796	1.6959	0.4142	0.0084
$\vdots$							
0.9999	0.0001	0.0000	0.0002	1.5001	3.5001	0.4999	
$\downarrow$	$\downarrow$						
1	0	0	0	1.5	3.5	0.5	

Anmerkungen zur Konvergenz und praktische Effizienz:

- Man kann zeigen, dass bei geeigneter Wahl von δ und der Schrittweitenstrategie der Algorithmus $O\left(\sqrt{n}\log\left(\frac{\varepsilon_0}{\varepsilon}\right)\right)$ Iterationen benötigt, um die Dualitätslücke von ε_0 auf ε zu reduzieren. D.h. polynomielle Laufzeit bei geeigneter Implementierung.
- In empirischen Versuchen zeigt sich typischerweise ein Schrittzahlverhalten von
$$O\left(\log n \cdot \log\left(\frac{\varepsilon_0}{\varepsilon}\right)\right).$$
- Methode wird gerne auch in Praxis eingesetzt, insbesondere für große lineare Programme.

Bemerkung *Es gibt viele weitere innere Punkte Verfahren und spezielle Schrittweiten-Steuerungsmethoden (z.B. Prädiktor-Korrektur-Verfahren).*

Teil II

Ganzzahlige Optimierung

7 Vollständig unimodulare Matrizen

Allgemeine Fragestellung:

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

Frage: Wann besitzt dieses lineare Programm (nur) ganzzahlige Optimallösungen?

Definition *Eine $n \times n$-Matrix heißt* unimodular, *wenn*

$$\det A = 1 \quad \textit{oder} \quad \det A = -1.$$

Eine $n \times n$-Matrix heißt vollständig unimodular *(„totally unimodular"), wenn für jede quadratische Untermatrix B von A*

$$\det B = 0 \quad \textit{oder} \quad \det B = 1 \quad \textit{oder} \quad \det B = -1$$

gilt.

Beispiel

$$A = \begin{pmatrix} 3 & 2 \\ 7 & 5 \end{pmatrix}$$

ist unimodular, aber nicht vollständig unimodular.

Bemerkung *A vollständig unimodular $\Longrightarrow a_{ij} \in \{0, \pm 1\}$.*

Beispiel

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

Untersuche, ob A vollständig modular ist.

Satz 7.1 *Sei A vollständig unimodular. Dann besitzt*

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

für jedes ganzzahlige b eine ganzzahlige Optimallösung.

Beweis Aufgrund des Hauptsatzes der linearen Optimierung können wir uns auf Basislösungen einschränken. Betrachte Basislösung:

$$\begin{aligned} x_B &= A_B^{-1} b \\ x_N &= 0 \quad \text{(ganzzahlig)}, \end{aligned}$$

wobei

$$A_B^{-1} = \frac{1}{\det A_B} A_B^{+}$$

(Matrix der Kofaktoren ist ganzzahlig, $\det A_B \in \{+1, -1\}$). Somit ist x_B ganzzahlig. Also ist jede Basislösung (Ecke) ganzzahlig, wenn b ganzzahlig ist. ∎

Satz 7.2 (Hoffman, Kruskal, Gale) *Sei A eine ganzzahlige $n \times n$ Matrix. Dann sind folgende Aussagen äquivalent:*

1. *A ist vollständig unimodular.*
2. *Für jedes beliebige ganzzahlige b besitzt die Menge*
$$S(b) := \{x : Ax \leq b, x \geq 0\}$$
nur ganzzahlige Ecken.
3. *Jede quadratische, nicht singuläre Untermatrix von A besitzt eine ganzzahlige Inverse.*

Beweis $1 \Longrightarrow 2$: Ecke von $S(b)$ entspricht einer Basislösung

$$Ax \leq b \Rightarrow Ax + Iy = b,$$

erhalte also das System

$$\begin{pmatrix} A & I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = b.$$

Basismatrix zur Basislösung ist Untermatrix von $\begin{pmatrix} A & I \end{pmatrix}$ mit $\det \neq 0$.

Sei A_B von der Gestalt

$$\begin{pmatrix} B_k & 0 \\ * & I_{m-k} \end{pmatrix}.$$

Basislösung $x_B = A_B^{-1}b$. Wie sieht also die Inverse aus?

$$A_B^{-1} = \begin{pmatrix} B_k^{-1} & 0 \\ * & I_{m-k} \end{pmatrix}$$

B_k ist quadratische Untermatrix von A. Lt. Voraussetzung $\det B_k \in \{+1, -1\}$.

Mit dem selben Argument wie im Beweis von Satz 7.1 folgt, dass B_k^{-1} ganzzahlig ist.

$$\det A_B = \underbrace{\det(B_k)}_{+1,-1} \cdot \underbrace{\det(I_{r-k})}_{=1} \in \{\pm 1\},$$

also ist A_B^{-1} ganzzahlig. Somit ist auch $A_B^{-1}b$ ganzzahlig.

$2 \Longrightarrow 3$: Sei A_B beliebige Baismatrix. Zu zeigen: A_B^{-1} ganzzahlig. Sei $\overline{b_i}$ die i-te Spalte von A_B^{-1}. Sei weiters z ein beliebiger ganzzahliger Vektor, sodass $\overline{b_i} + z \geq 0$ für alle i.

$$S^{(i)}(b(z)) := \{x \mid Ax \leq \underbrace{A_B \cdot z + e_i}_{\text{ganzzahlig}}\},$$

wobei e_i der i-te Einheitsvektor ist. Alle $S^{(i)}(b(z))$ haben nur ganzzahlige Ecken.

$$x_B = A_B^{-1}(A_B z + e_i) = z + \overline{b_i}$$

ist ganzzahlig. Daher ist auch z ganzzahlig, und somit auch A_B^{-1}.

Sei F eine beliebige reguläre, quadratische Untermatrix von A, die keine Basismatrix ist. Dann lässt sich F zu einer Basismatrix von

$$Ax + Iy = b$$

ergänzen.

$$A_B = \begin{pmatrix} F & 0 \\ * & I_{m-k} \end{pmatrix},$$

wobei F eine $k \times k$ Matrix ist.

$$A_B^{-1} = \begin{pmatrix} F^{-1} & 0 \\ * & I_{m-k} \end{pmatrix}$$

ist ganzzahlig (vorher bewiesen), damit auch F^{-1}.

$3 \Longrightarrow 1$: Sei F eine beliebige reguläre, quadratische Untermatrix von A. Laut Voraussetzung sind F und F^{-1} ganzzahlig.

$$F \cdot F^{-1} = I$$
$$\det F \cdot \det F^{-1} = \det I = 1$$

Weil F und F^{-1} ganzzahlig sind, sind auch ihre Determinanten ganzzahlig, und daher

$$\det F, \det F^{-1} \in \{\pm 1\}.$$

Somit ist A vollständig unimodular. ■

Fragestellungen:

- Wie schauen vollständig unimodulare Matrizen aus?
- Beispiele bzw. Beispielklassen?
- Charakterisierung bzw. Erkennung

Bemerkung *Sei A vollständig unimodular. Dann sind auch die Matrizen*

$$A^t, \begin{pmatrix} A & A \end{pmatrix}, -A, \begin{pmatrix} A & I \end{pmatrix}, \begin{pmatrix} A & -A \end{pmatrix}$$

vollständig unimodular.

Aus A, B vollständig unimodular folgt allerdings nicht*, dass $\begin{pmatrix} A & B \end{pmatrix}$ vollständig unimodular ist.*

Satz 7.3 (Heller, Tompkins, Gale) *Sei A eine $m \times n$ Matrix mit Einträgen $\in \{0, \pm 1\}$. A ist vollständig unimodular, wenn folgendes gilt:*

1. *Jede Spalte von A hat höchstens 2 von 0 verschiedene Einträge.*
2. *Die Zeilen von A lassen sich in zwei disjunkte Klassen Z_1 und Z_2 einteilen, sodass gilt:*
   a) *Hat eine Spalte zwei Einträge ungleich 0 mit verschiedenen Vorzeichen, so liegen die zugehörigen Zeilen in derselben Klasse.*
   b) *Hat eine Spalte zwei Einträge mit gleichem Vorzeichen, so liegen die zugehörigen Zeilen in verschiedenen Klassen.*

Beispiel

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

Bedingung 1 ist erfüllt.
Bedingung 2: Klasseneinteilung $\langle Z_1, Z_1, Z_2, Z_2 \rangle$.
A ist also vollständig unimodular.

Bemerkung *Bedingung ist hinreichend, aber nicht notwendig! Wenn keine Klasseneinteilung existiert, kann A sowohl unimodular als auch nicht unimodular sein.*

Beweis Vollständige Induktion nach Größe der Matrix:
1×1 Matrix: klar.

Wir greifen nun aus A eine beliebige $p \times p$ Matrix B heraus. Induktionsannahme: Behauptung für alle $(p-1) \times (p-1)$ Untermatrizen bereits bewiesen. Fallunterscheidung:

1. B enthält 0-Spalte.

   Dann ist $\det B = 0$, fertig.

2. B enthält eine Spalte mit genau einem Eintrag $\neq 0$.

   Entwickle $\det B$ nach einer solchen Spalte j.

   $$|\det(B)| = \left| b_{ij} \cdot \det \tilde{B} \right|,$$

   wobei b_{ij} der Eintrag $\neq 0$ in der Spalte j ist und $\tilde{B}$ die $(p-1) \times (p-1)$ Untermatrix von B, A. Die Determinante ist also $\in \{0, \pm 1\}$ lt. Induktionsannahme.

3. Jede Spalte von B habe genau 2 Einträge $\neq 0$.

   Es gilt

   $$\sum_{i \in Z_1} a_{ij} = \sum_{i \in Z_2} a_{ij}$$

   für alle Spalten j von B, also

   $$\sum_{i \in Z_1} a_{ij} - \sum_{i \in Z_2} a_{ij} = 0.$$

   Es ergibt sich also eine lineare Abhängigkeit, weswegen $\det B = 0$ gilt. ■

7.1 Beispielklassen für vollständig unimodulare Matrizen

7.1.1 Transportproblem

n Fabriken, n Kunden, Transporte kosten c_{ij}.

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

s.t.

$$\begin{aligned} \sum_{j=1}^{n} x_{ij} &= a_i \quad \forall i = 1, \ldots, m \\ \sum_{i=1}^{m} x_{ij} &= b_j \quad \forall j = 1, \ldots, n \\ x_{ij} &\geq 0. \end{aligned}$$

Jede Spalte hat zwei von 0 verschiedene Einträge, Klasseneinteilung wie oben. Die Restriktionsmatrix eines Transportproblems ist also vollständig unimodular.

7.1.2 Knoten-Kanten-Inzidenzmatrix eines gerichteten Graphen

Knoten-Kanten-Inzidenzmatrizen sind vollständig unimodular. Sie treten als Restriktionsmatrix in Flussproblemen auf (Flusserhaltung).

7.1.3 Network matrices

Man kann zeigen, dass network matrices vollständig unimodular sind.

Korollar *Seien A, b, c ganzzahlig. A sei vollständig unimodular.*

1. *Dann hat sowohl das primale Problem*
$$\max\{c^t x : Ax \leq b, x \geq 0\}$$
also auch das duale
$$\min\{b^t y : A^t y \geq c, y \geq 0\}$$
nur ganzzahlige Ecken (Basislösungen).

2. *Die Ecken von*
$$S := \{x : \tilde{b} \leq Ax \leq \hat{b}, 0 \leq x \leq \tilde{u}\}$$
sind für alle ganzzahligen $\tilde{b}, \hat{b}, \tilde{u}$ ganzzahlig.

Beweis

1. A vollständig unimodular $\Longrightarrow$ A^t vollständig unimodular.
2. $$\begin{pmatrix} Ax \\ -Ax \\ Ix \end{pmatrix} \leq \begin{pmatrix} \hat{b} \\ -\tilde{b} \\ \tilde{u} \end{pmatrix},$$

$\begin{pmatrix} A \\ -A \\ I \end{pmatrix}$ vollständig unimodular. ■

7.2 Anmerkung zur Erkennung vollständig unimodularer Matrizen

- Klasseneinteilungskriterium erkennt nur Teilklasse.
- Trivialer Algorithmus hat exponentielle Laufzeit und ist für größere Matrizen unbrauchbar.
- 1980 hat Seymour ein Zerlegungsresultat für die Klasse der vollständig unimodularen Matrizen gezeigt. (Vollständig unimodulare Matrizen lassen sich zerlegen über der Klasse der *network matrices* und 2 speziellen Matrizen.) Das Zerlegungsresultat kann für polynomielle Erkennungsalgorithmen eingesetzt werden.

8 Dynamische Programmierung (dynamische Optimierung)

8.1 Binäres Rucksackproblem

Das binäre Rucksackproblem: n Gegenstände. Gegenstand i habe Gewicht a_i und Wert c_i. Rucksackkapazität b.

$$\max \sum_{i=1}^{n} c_i x_i$$

s.t.

$$\sum_{i=1}^{n} a_i x_i \leq b$$
$$x_i \in \{0,1\}.$$

Dieses Problem gehört zur Klasse der NP-schweren Probleme! Naive Methode: Alle Möglichkeiten durchprobieren (Zeitbedarf $O(2^n)$, unbrauchbar).

Idee: Verwende nur Teile dieses Graphen. Zentrale Idee: rekursives Vorgehen/Stufeneinteilung. Bestimmung der besten Lösung auf Stufe k aufbauend auf der besten Lösung auf Stufe $k-1$.

Bei Rucksackproblem: Neue Stufe ist neuer Gegenstand.
Stufe 1: Nur Gegenstand 1 zur Verfügung.
Stufe 2: Nur Gegenstände 1 und 2 zur Verfügung.
etc.

$$F(k,y) := \max\{\sum_{i=1}^{k} c_i x_i : \sum_{i=1}^{n} a_i x_i \leq y, x_i \in \{0,1\}, i = 1,\dots,k\}, \qquad k = 1,\dots,n, \quad y = 0,\dots,b$$

ist das Rucksackproblem mit Kapazität y und den Gegenständen $1,\dots,k$. $F(n,b)$ ergibt den Optimalwert von (RP). Das Ziel ist also, $F(n,b)$ zu berechnen.

Zentrale Beobachtung:

$$F(k+1,y) = \max\{\underbrace{F(k, y - a_{k+1}) + c_{k+1}}_{\text{Gegenstand } k+1 \text{ kommt mit}}, \underbrace{F(k,y)}_{\text{Gegenstand } k+1 \text{ kommt nicht mit}}\}.$$

Initialisierung:

$$F(0,y) = 0 \qquad \text{für } y \geq 0$$
$$F(k,y) = -\infty \qquad \text{für } y < 0, k = 0,\dots,n.$$

$F(k+1,y)$ hängt nur von $F(k,\cdot)$-Werten ab $\rightarrow$ stufenweise Berechnung: zuerst $k=1$, dann $k=2$ etc.

Beispiel $n = 4, b = 11$ *und*

i	1	2	3	4
a_i	2	2	4	6
c_i	8	6	10	12

.

Berechne schrittweise $F(k, y)$ und erhalte so

$y\backslash k$	0	1	2	3	4
0	0	0	0	0	0
1	0	0	0	0	0
2	0	8	8	8	8
3	0	8	8	8	8
4	0	8	14	14	14
5	0	8	14	14	14
6	0	8	14	18	18
7	0	8	14	18	18
8	0	8	14	24	24
9	0	8	14	24	24
10	0	8	14	24	26
11	0	8	14	24	26

Der Optimalwert ist also 26.
Bestimmung der optimalen Packung:

$$\begin{aligned} F(4,11) &> F(3,11) && \Rightarrow x_4^* = 1 \\ F(3,5) &= F(2,5) && \Rightarrow x_3^* = 0 \\ F(2,5) &= F(21, 5-2) + 6 && \Rightarrow x_2^* = 1 \\ F(1,3) &> F(0,3) && \Rightarrow x_1^* = 1. \end{aligned}$$

Zeitaufwand: $O(n{\cdot}b)$ Variablen (Tabellenaufwand), $O(1)$ pro Auswertung von $F(k, y)$. Insgesamt also $O(n \cdot b)$ (pseudopolynomial, nicht polynomialer Algorithmus).

8.2 Matrixmultiplikation

Folge von Matrizen $M_1, \ldots, M_q$, Matrix M_i hat Dimension $n_i \times n_{i+1}$, $i = 1, \ldots, q$. Wir wollen das Produkt $M_1 M_2 \cdots M_q$ berechnen und die Anzahl der Elementaroperationen soll minimiert weren. Welche Klammerung soll gewählt werden?

Aufwand für Multiplikation einer $r \times s$ mit einer $s \times t$ Matrix: $O(rst)$.

Beispiel *4 Matrizen, $n_1 = 10, n_2 = 5, n_3 = 100, n_4 = 10$.*
Beobachtung: Unterschiedliche Klammerungen führen auf unterschiedlichen Gesamtaufwand.

Idee: Sei t_{ij} der minimale Aufwand (im obigen Sinn) für die Berechnung von $M_i M_{i+1} \cdots M_j$.

$$\begin{aligned} t_{ii} &= 0 && \text{für } i = 1, \ldots, q \\ t_{i,i+1} &= n_i n_{i+1} n_{i+2} && \text{für } i = 1, \ldots, q-1 \\ t_{i,i+s} &= \min_{k=1,\ldots,i+s-1} \{t_{ik} + t_{k+1,i+s} + n_i \cdot n_{k+1} \cdot n_{i+s+1}\} && \text{für } s \geq 2 \text{ und } i = 1, \ldots, r-s. \end{aligned}$$

Ergebnis: t_{1q}.
Rechenaufwand: $O(q^2)$ Variablen, $O(q)$ pro Auswertung, also $O(q^3)$ insgesamt.

9 Die Branch and Bound Methode

Betrache das Rucksackproblem

$$z_{\mathrm{RP}} = \max \sum_{i=1}^{n} c_i x_i$$

s.t.

$$\sum_{i=1}^{n} a_i x_i 6 \leq b$$
$$x_i \in \{0,1\} \qquad quadi = 1, \ldots, n.$$

Jedes Blatt entspricht einer Packung. Naiver Lösungsansatz: Alle Blätter des Baumes durchprobieren. (Explizite Enumeration: 2^n Blätter.)

Verbesserungsidee: Implizite Enumeration des Lösungsbaums, lasse also Teile des Lösungsbaums weg, für die erwiesen ist, dass sie die Optimallösung nicht enthalten.

Frage: Wie lassen sich Teile des Lösungsbaums ausschließen?

Beispiel $n = 4, b = 11$ *und*

	1	2	3	4
c_j	8	6	10	12
a_j	2	2	4	6
$\frac{c_j}{a_j}$	4	3	2.5	2

.

Hilfsproblem (LP):

$$z_{LP} = \max \sum_{i=1}^{n} c_i x_i$$

s.t.

$$\sum_{i=1}^{n} a_i x_i \leq b$$
$$0 \leq x_i \leq 1 \quad i = 1, \ldots, n.$$

Beobachtung: $z_{LP} \geq z_{RP}$ *(bzw.* $\lfloor z_{LP} \rfloor \geq z_{RP}$ *bei ganzzahligen* a_i, c_i*). (LP), die sogenannte* lineare Programm Relaxation *von (RP), liefert eine obere Schranke für den Optimalwert von (RP) (vgl. Ü68).*

$$x_1^* = 1, x_2^* = 1, x_3^* = 1, x_4^* = \frac{1}{2}.$$

Der optimale Zielfunktionswert von (LP) ist also $8 + 6 + 10 + \frac{1}{2} \cdot 12 = 30$. *(Es gibt also keine Lösung von (RP) mit Wert* > 30*.) Außerdem liefert dies eine zulässige Lösung von (RP):*

$$x_1^* = 1, x_2^* = 1, x_3^* = 1, x_4^* = 0$$

mit Zielfunktionswert 24.

Allgemeine Vorgehensweise Versuche schwieries Problem (hier Rucksackproblem) durch ein „einfacheres" relaxiertes Problem zu ersetzen, sodass

- die zulässige Menge des relaxierten Problems eine Obermenge der zulässigen Menge des Ausgangsproblems ist und
- das relaxierte Problem möglichst hinreichend einfach zu lösen ist, aber immer noch eine hinreichend große Beziehung zum Ausgangsproblem hat (*Trade-off*).

Für das Rucksackproblem ist (LP) die typischerweise gewählte Relaxation. Die Relaxation bildet eine obere Schranke für Maximierungsprobleme und eine untere Schranke für Minimierungsprobleme.

Beispiel *Verzweigung nach kritischer Variable x_4 ($x_4^* \notin \{0,1\}$).*
Zweige brauchen nicht enumeriert werden, wenn bereits eine bessere zulässige Lösung in einem andere Zweig gefunden wurde!

Hauptbestandteile des Branch and Bound Verfahrens

- Relaxation (führt für Maximierungsproblem auf obere Schranke)
- Aufspaltungsregel in Teilprobleme (Branching). Notwendig: „Vereinigung" der Teilprobleme muss aufzuspaltendes Problem ergeben. Nach Möglichkeit strebt man Disjunktheit der Probleme an. Wichtiges Kriterium bei der Wahl der Aufteilung ist, dass die Relaxationen der Teilprobleme „vernünftig" lösbar sind.
- Auswahlregel für das als nächstes zu behandelnde Teilproblem. Trade-off zwischen Speicherplatzbedarf und der Größe des resultierenden Suchbaums.
  - Best-Bound: eher schlecht für Speicherplatzbedarf
  - LILO (last in last out): entspricht Tiefensuche, eher günstig für Speicherplatzbedarf, eher schlecht für Baumgröße
  - FIFO (first in first out)

  und viele Kombinationen. Wahl typischerweise problemabhängig.
- Optional: Von Vorteil stellt sich heraus, wenn eine Methode zur Bestimmung guter zulässiger Lösungen für die Teilprobleme bekannt ist. Daraus ergibt sich eine untere Schranke für Maximierungsprobleme (bezieht sich auf entsprechende Teilprobleme).

  Für Rucksackproblem: zulässige Lösung durch *Greedy-Lösung* (erhält man aus Relaxationslösung durch Setzen der kritischen Variable auf 0).

Es gibt weitere Verfeinerungen von B&B-Verfahren, z.B. die *Dominanzregel*.

9.1 Gemischt-ganzzahlige lineare Programme (mixed integer programs)

$$\max c^t x$$

s.t.

$$\begin{aligned} Ax &\leq b \\ x &\geq 0 \end{aligned}$$

mit

$$x_1, \ldots, x_k \in \mathbb{Z}, x_{k+1}, \ldots, x_n \in \mathbb{R}.$$

Beispiel

$$\max x_1 + 2x_2$$

s.t.

$$\begin{aligned} 2x_1 + x_2 &\leq 10 \\ -x_1 + x_2 &\leq 5 \\ x_1 &\leq 4 \\ x_1, x_2 &\in \mathbb{Z} \end{aligned}$$

Relaxiertes lineares Programm durch Streichen der Bedingung $x_1, x_2 \in \mathbb{Z}$. *Löse lineares Programm: Optimallösung* $x_1^* = \frac{5}{3}, x_2^* = \frac{20}{3}$ *mit ZFW 15.*

Verzweigung nach Variablen mit nicht ganzzahligem Wert in der Optimallösung der Relaxation. Verzweige also nach x_1*:*

$$\left\lfloor \frac{5}{3} \right\rfloor = 1, \quad \left\lceil \frac{5}{3} \right\rceil = 2.$$

Bemerkung *Diese Methode funktioniert in der Praxis für größere Probleme schlecht, weil der Suchbaum typischerweise sehr groß ist. Dies führt auf das* Schnittebenenverfahren *(„cutting plane“),* Branch and Cut.

10 Einige Beispiele für die Modellierung mit ganzzahligen Variablen

Typische Anwendungen

- naheliegend: Anzahl etc.
- für 0/1-Variablen: Indikator für Ja/Nein-Entscheidung (z.B: Rucksackproblem: $x_j = 0$: j-ter Gegenstand kommt nicht mit, $x_j = 1$: j-ter Gegenstand kommt mit.)

Transportproblem Sei x_{ij} die transportierte Menge einer Ware von einem potenziellen Fabrikstandort i zu einem Abnehmer j. Es gibt fixe Errichtungskosten und Betriebskosten f_i für Fabrik i und eine obere Schranke u_i für die produzierte Menge. Die Summe der Errichtungskosten und Transportkosten ist zu minimieren.

Führe die Variablen

$$y_i = \begin{cases} 0 & \text{keine Fabrik in Standort } i \\ 1 & \text{sonst} \end{cases}$$

ein. Das Problem lautet dann

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} + \sum_{i=1}^{m} f_i y_i$$

s.t.

$$\sum_{i=1}^{m} x_{ij} = b_i, \quad j = 1, \ldots, n.$$

Kopplung zwischen x_{ij} und y_i: Fabrik i nicht errichtet ($y_i = 0$) $\Longrightarrow$ kein Transport von i aus $\Leftrightarrow x_{ij} = 0 \ \forall j$.

Frage: Wie lässt sich $y_i = 0 \Longrightarrow x_{ij} = 0 \ \forall j$ mittels einer linearen Restriktion formulieren?

Eine Möglichkeit ist

$$x_{ij} \leq u_i \cdot y_i \ \forall j.$$

Für $y_i = 0$ bedeutet das $x_{ij} \leq 0 \Rightarrow x_{ij} = 0$, für $y_i = 1$ ergibt sich mit $x_{ij} \leq u_i$ keine neue Einschränkung.

Weiters muss

$$\begin{aligned} x_{ij} &\geq 0 && \forall i, j \\ y_i &\in \{0, 1\} && \forall i \end{aligned}$$

gefordert werden.

10.1 Disjunkte Nebenbedingungen (Restriktionen)

Beispiele

- $|x| \geq 7$
- $x_1 + x_2 \leq 4$ *oder* $x_1 + 3x_2 \geq 7$

Der Bereich ist jeweils nicht konvex, daher mit linearer Optimierung nicht formulierbar!
Idee: *Indikatorvariable* für „Erfülltheit“ der Restriktionen.

Beispiel

$$\delta_1 = \begin{cases} 1 & x_1 + x_2 \leq 4 \\ 0 & \textit{sonst} \end{cases} \quad \textit{und} \quad \delta_2 = \begin{cases} 1 & x_1 + x_2 \geq 9 \\ 0 & \textit{sonst} \end{cases}.$$

Neue Restriktionen:

$$\delta_1 + \delta_2 \geq 1$$

(Gleichheit bei ausschließendem Oder*).*

Sei weiters M_1 eine obere Schranke für $x_1 + x_2 - 4$, d.h. $x_1 + x_2 - 4 \leq M_1$ für alle zulässigen Kombinationen von x_1 und x_2. Somit

$$x_1 + x_2 - 4 \leq M_1(1 - \delta_1).$$

Wenn $x_1 + x_2 > 4$ (also $>$ statt $\geq$) vorkommt:

- *Bei ganzzahligen Variablen: $x_1 + x_2 \geq 5$.*
- *Einführen von $\varepsilon > 0$ geeignet klein („Maschinenepsilon“, problemabhängig gewählt). Statt mit $x_1 + x_2 > 4$ wird mit $x_1 + x_2 \geq 4 + \varepsilon$ gearbeitet.*

Analog für 2. Restriktion und δ_2.

Bemerkung *Oft lässt sich auf einen Teil der Koppelung verzichten mittels Optimalitätsüberlegungen.*

10.2 Formulierung stückweise linearer Zielfunktionen

Beispiel

$$\min f(x) + 8y$$

s.t.

$$\begin{aligned} x + y &\geq 5 \\ y &\leq 4 \\ 0 \leq x &\leq 8 \end{aligned}$$

mit

$$f(x) = \begin{cases} 10x & 0 \leq x \leq 2 \\ 8x + 4 & 2 \leq x \leq 4 \\ 6x + 12 & 4 \leq x \leq 6 \\ 4x + 24 & 6 \leq x \leq 8 \end{cases}.$$

Führe Indikatorvariablen ein:

$$z_1 = \begin{cases} 1 & 0 \leq x \leq 2 \\ 0 & \textit{sonst} \end{cases}, \quad z_2 = \begin{cases} 1 & 2 \leq x \leq 4 \\ 0 & \textit{sonst} \end{cases}, \ \textit{etc.}$$

Restriktionen:

$$\begin{aligned} z_1 + z_2 + z_3 + z_4 &= 1 \\ z_i &\in \{0, 1\}, \quad i = 1, \dots, 4 \end{aligned}$$

Definiere

$$x_i = \begin{cases} x & x \text{ im } i\text{-ten Intervall } (z_i = 1) \\ 0 & \text{sonst} \end{cases}.$$

Es muss gelten: $z_i \Longrightarrow x_i = 0$. *Dies kann mittels*

$$\begin{array}{llll} x_1 \geq 0 & x_2 \geq 2z_2 & x_3 \geq 4z_3 & x_4 \geq 6z_4 \\ x_1 \leq 2z_1 & x_2 \leq 4z_2 & x_3 \leq 6z_3 & x_4 \leq 8z_4 \end{array}$$

und $x = x_1 + x_2 + x_3 + x_4$ *erreicht werden.*

Bemerkung *Wenn Zielfunktion zu minimieren und die Funktion stückweise linear und konvex ist, dann kann man ohne ganzzahlige Variablen auskommen.*

Analoge Technik funktioniert auch für fixkostenartige Funktionen (Sprungstelle dabei):

$$g(z) = \begin{cases} 0 & z = 0 \\ a + bz & z > 0 \end{cases}$$

$(a \neq 0)$. Indikator $a \cdot \delta + bz$ mit

$$\delta = \begin{cases} 1 & z > 0 \\ 0 & \text{sonst} \end{cases}.$$

10.3 Funktionen mit N möglichen Werten

$$f(x_1, \ldots, x_n) \in \{d_1, \ldots, d_N\},$$

z.B.

$$f(x_1, \ldots, x_n) = \sum_{j=1}^{n} a_j x_j.$$

Indikatorvariablen:

$$y_i = \begin{cases} 1 & f(x_1, \ldots, x_n) = d_i \\ 0 & \text{sonst} \end{cases}$$

$$\sum_{i=1}^{N} y_i = 1, \quad y_i \in \{0, 1\}.$$

Dann gilt

$$f(x_1, \ldots, x_n) = d_1 \cdot y_1 + d_2 \cdot y_2 + \cdots + d_N y_N.$$

Beispiel $3x_1 + 2x_2 \in \{6, 12, 18\}$.
$N = 3, d_1 = 6, d_2 = 12, d_3 = 18$.

$$\begin{aligned} 3x_1 + 2x_2 &= 6y_1 + 12y_2 + 18y_3 \\ y_1 + y_2 + y_3 &= 1 \\ y_i &\in \{0, 1\} \quad i = 1, 2, 3 \end{aligned}$$

10.4 Transportproblem

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

s.t.

$$\begin{aligned} \sum_j x_{ij} &= a_i & i &= 1, \dots, m \\ \sum_i x_{ij} &= b_j & j &= 1, \dots, n \\ x_{ij} &\geq 0, \end{aligned}$$

o.B.d.A. mit

$$\sum_i a_i = \sum_j b_j.$$

Nordwesteckenregel zur Bestimmung einer Ausgangslösung: Setze vorerst

$$\begin{aligned} \tilde{a}_i &:= a_i & i &= 1, \dots, m \\ \tilde{b}_j &:= b_j & j &= 1, \dots, n. \end{aligned}$$

Starte in linker oberer Ecke: $i := 1, j := 1$. Betrachte x_{ij} („nordwestlich“, nicht gestrichen):

$$x_{ij} := \min\{\tilde{a}_i, \tilde{b}_j\}$$

und

$$\tilde{a}_i := \tilde{a}_i - x_{ij}, \qquad \tilde{b}_j := \tilde{b}_j - x_{ij}.$$

Falls $\tilde{a}_i = 0, \tilde{b}_j > 0$, streiche Zeile i.
Falls $\tilde{a}_i > 0, \tilde{b}_j = 0$, streiche Spalte j.
Falls $\tilde{a}_i = 0, \tilde{b}_j = 0$, streiche Zeile i und Spalte j (degenerierter Fall, relevant für Basis).

Beispiel $a = (6, 25, 20, 13)$ *und* $b = (4, 6, 10, 9, 7, 18)$.

$$x = \begin{array}{cccccc|c} 4 & 16 & 10 & 9 & 7 & 18 & \\ \hline 4 & 2 & & & & & 6 \\ & 14 & 10 & 1 & & & 25 \\ & & & 8 & 7 & 5 & 20 \\ & & & & & 13 & 13 \end{array}.$$

Es gilt

$$\operatorname{rg} A = m + n - 1.$$

Beobachtung A hat nicht vollen Zeilenrang $\Longrightarrow \operatorname{rg}(A) \leq m + n - 1$. Der zugehörige Graph ist kreisfrei und zusammenhängend (ein spannender Baum).

Beobachtung Eine Kantenauswahl, die einen Kreis im Transportproblem enthält, liefert keine Basis (lineare Abhängigkeit).

Behauptung: $\overline{A}_B$ ist genau dann Basismatrix von $\overline{A}$, wenn die Spalten von $\overline{A}_B$ den Kanten eines spannenden Baumes entsprechen.

Beweis: Bereits gezeigt: Kreis enthalten $\Longrightarrow$ keine Basis. Andere Richtung konstruktiv, Vorgangsweise startet bei Blättern.

- Nordwestregel
- spannende Bäume: $m+n-1$ Kanten im Basisgraph
- Wahl Variable neu in Basis?

  Verwende reduzierte Kostenkoeffizienten:

b_j	2	4	6	3
4	2	2		
7		2	5	
4			1	3

(nicht degenerierte Basislösung).

$$\tilde{c}_N^t = c_N^t - c_B^t \underbrace{A_B^{-1} A_N}_{\tilde{A}_N}.$$

$(A_B, A_N) = \overline{A}$ ist die Transportmatrix mit einer (beliebigen) Zeile gestrichen.

$$\pi^t = c_B^t A_B^{-1},$$

also

$$\pi^t A_B = c_B^t.$$

Somit ist π Lösung dieses Gleichungssystems.

$$\pi^t = (u_1, \dots, u_m, v_1, \dots, v_{n-1})$$

(n-te Abnehmerbedingung gestrichen, setze also $v_n := 0$).

Transportproblem:

$$\min \sum_i \sum_j c_{ij} x_{ij}$$

s.t.

$$\begin{aligned} \sum_j x_{ij} &= a_i \quad \forall i \\ \sum_i x_{ij} &= b_j \quad \forall j \\ x_{ij} &\geq 0. \end{aligned}$$

Duales Problem:

$$\max \sum a_i u_i + \sum b_j v_j$$

s.t.

$$u_i + v_j \leq c_{ij} \quad \forall i, j.$$

Es gilt nun

$$u_i + v_j = c_{ij} \quad \forall (i,j) \in B,$$

wobei B eine Basis ist.

Beispiel

$$C = \begin{pmatrix} 2 & 3 & 4 & 1 \\ 5 & 4 & 2 & 3 \\ 5 & 2 & 8 & 6 \end{pmatrix}.$$

Hier gilt also

$$\begin{aligned} u_1 + v_1 &= 2 \\ u_1 + v_2 &= 3 \\ &\vdots \\ u_3 + v_4 &= 6. \end{aligned}$$

Vorgehensweise:

1. *Bestimme π, also die u_i, v_j.*
2. *Berechne den reduzierten Kostenkoeffizienten $\overline{c}_{ij}$ mittels*

$$\overline{c}_{ij} = c_{ij} - u_i - v_j.$$

Wende das rekursiv auf den Baum an, startend bei den Blättern.

$$\begin{aligned} \overline{c}_{13} &= 4 - u_1 - v_3 = 4 + 1 - 2 = 3 \\ \overline{c}_{14} &= 1 - u_1 - v_4 = 2 \end{aligned}$$

etc. Somit

$$\overline{C} = \begin{pmatrix} & & 3 & 2 \\ 2 & & & 3 \\ -5 & -8 & & \end{pmatrix}.$$

Optimalitätskriterium: Basis optimal dann, wenn $\overline{c}_{ij} \geq 0$ für alle (i, j).

Antwort: Wähle Nichtbasisvariable (also nicht im Baum) mit $\overline{c}_{ij} < 0$.

- Wahl Variable raus aus Basis?

  Durch das Hinzufügen einer neuen Kante zum Basisbaum entsteht genau ein Kreis.

  Beobachtung: Entlang des Kreises werden abwechselnd δ mehr bzw. um δ weniger Einträge verschickt.

  Sei Q die Menge der Kreiskanten und Q^+ die Kanten mt $+\delta$, Q^- jene mit $-\delta$.

  Zulässigkeitsbedingung:

  $$x_{ij} - \delta \geq 0$$

  für alle Kanten im Kreis, für die neuer Flusswert $x_j - \delta$ ist. Somit

  $$\delta = \min\{x_{ij} \mid (i, j) \in Q^-\}.$$

Teil III

Kurzeinführung in Nichtlineare Optimierung

11 Einführung

Die Nichtlineare Optimierung beschäftigt sich mit dem Finden lokaler Optima.
Es gibt zwei große Klassen:

1. Probleme ohne Nebenbedingungen („unconstrained“)
2. mit Nebenbedingungen („constrained“)

Hier:

- Optimalitätsbedingungen für 2 (und damit auch für 1
- Kurze Übersicht über einige Methoden zur Lösung von 1

Wir wollen folgendes Problem betrachten:

$$\min f(x)$$

s.t.

$$\begin{aligned} g_i(x) &\leq 0 \qquad i = 1, \ldots, p \\ h_j(x) &= 0 \qquad j = 1, \ldots, m \end{aligned}$$

(*Ungleichungsrestriktionen* bzw. *Gleichungsrestriktionen*) mit $x \in \mathbb{R}^n$ und $f, g_i, h_j : \mathbb{R}^n \to \mathbb{R}$ (nicht der allgemeinste Fall). $p =$ und/oder $m = 0$ sind möglich.
Spezielle Klassen von Problemen ergeben sich für spezielle Eigenschaften von f, g_i, h_j.

Definition (Zulässigkeit)

$$\Omega = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0 \textit{ für } i = 1, \ldots, p, h_j(x) = 0 \textit{ für } j = 1, \ldots, m\}$$

heißt zulässige Menge. x *heißt* zulässiger Punkt

Definition (lokales, globales Minimum)

1. $x^* \in \Omega$ *heißt* lokales Minimum *von* f *über* Ω *genau dann, wenn*
$$\exists \varepsilon > 0 \ \forall x \in \Omega : \|x - x^*\| < \varepsilon \to f(x) \geq f(x^*).$$
2. $x^* \in \Omega$ *heißt* streng lokales Minimum *von* f *über* Ω *genau dann, wenn*
$$\exists \varepsilon > 0 \ \forall x \in \Omega : x \neq x^* \textit{ und } \|x - x^*\| < \varepsilon \to f(x) > f(x^*).$$
3. $x^* \in \Omega$ *heißt* (globales) Minimum *von* f *über* Ω *genau dann, wenn*
$$\forall x \in \Omega : f(x^*) \leq f(x).$$

Definition (zulässige Richtung („feasable direction“)) *Sei* $x_0 \in \Omega$. $d \in \mathbb{R}^n$ *heißt* zulässige Richtung *in* x_0, *wenn*

$$\exists \overline{\lambda} > 0 \ \forall \lambda \textit{ mit } 0 \leq \lambda \leq \overline{\lambda} : x_0 + \lambda d \in \Omega.$$

Beispiele

1. $\Omega = [0,1]^2$ *(Quadrat)*, $x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

$$d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \text{ zulässig in } x_0 \Leftrightarrow d_1 \geq 0 \text{ und } d_2 \geq 0.$$

2. *Betrachte*
$$\Omega = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$$
(Einheitskreisscheibe), $x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

   Beobachtung: Es muss $d_1 \leq 0$ *gelten. Aus* $d_1 = 0$ *folgt* $d_2 = 0$. *Eine zulässige Richtung liegt also genau für* $d_1 < 0$ *vor.*

12 Verfahren zur Minimierung von Funktionen einer Variablen

Zwei Klassen von Verfahren:

1. solche, die Ableitungsinformationen verwenden (siehe Analysis und numerische Mathematik, z.B. Newton-Verfahren), und
2. solche, die keine Ableitungsinformationen verwenden (unser Schwerpunkt, z.B. Intervallschachtelung). (Typischerweise sind Ableitungsinformationen schwer oder gar nicht zu beschaffen!)

12.1 Verfahren ohne Ableitungsinformation

Definition *Sei* $f : I \to \mathbb{R}$, $I \subseteq \mathbb{R}$, *heißt* unimodal*, wenn* f *ein Minimum* $x^* \in I$ *besitzt und es gilt*

$$\begin{aligned} f(x) &\geq f(x^*) \quad \forall x \in I \text{ mit } x \leq x^* \\ f(x) &\geq f(x^*) \quad \forall x \in I \text{ mit } x \geq x^*. \end{aligned}$$

Definition $f : I \to \mathbb{R}$ *heißt* quasi konvex *auf* I*, wenn*

$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}$$

für alle $x, y \in I$ *und* $0 \leq \lambda \leq 1$ *gilt.*

Bemerkung *Quasikonvexität ist eine Verallgemeinerung der Konvexität.*

Lemma *Sei* $f : I \to \mathbb{R}$, $I = [a, b]$*, quasikonvex. Sei* $u, v \in I$ *mit* $u < v$.
Falls $f(u) > f(v)$*, dann gilt* $f(x) \geq f(v)$ *für alle* $x \in [a, u]$*. (Das heißt das Teilintervall* $[a, u]$ *kann eliminiert werden.)*
Falls $f(u) < f(v)$*, dann gilt* $f(x) \geq f(v)$ *für alle* $x \in [v, b]$.

1. Naive Vorgehensweise: Vorauswahl der Kandidatenpunkte (äquidistante Unterteilung)
2. Drittelungsmethode
3. Methode des goldenen Schnitts
4. Fibonacci-Suche („endliche Variante“ der goldenen Schnittmethode)
5. Dichotomie-Methode

12.2 Verfahren mit Ableitungsinformation

1. Bisektionsverfahren
2. Newton-Verfahren
3. Sekantenverfahren („regula falsi“)

13 Mehrdimensionale nichtlineare Optimierung ohne Nebenbedingungen

13.1 Wiederholung

13.2 Grundlagen von Lösungsverfahren

13.3 Allgemeines Abstiegsverfahren

13.4 Schrittweiten

1. Armijo-Regel
2. Wolfe-Powell-Strategie

13.5 Steilstes Abstiegsverfahren (Gradientenverfahren)

13.6 Newton-Verfahren

13.7 Quasi-Newtonverfaren

Versuch: Newton-artige Verfahren, ohne Verwendung von Hesse-Matrix.

Newton-Iteration:

$$x_{k+1} = x_k - (\nabla^2 f(x_k))^{-1} \cdot \underbrace{\nabla f(x_k)}_{g(x_k)}$$

Bezeichne $g(x_k) = \nabla f(x_k)$.

Bei Quasi-Newton:

$$x_{k+1} = x_k - t_k \cdot G_k \cdot g(x_k)$$

(t_k bezeichnet die *Schrittweite*), wobei G_k eine symmetrische $n \times n$-Matrix ist.

Ziel: G_k soll $(\nabla^2 f(x_k))^{-1}$ „geeignet“ annähern.

Für $G_k = (\nabla^2 f(x_k))^{-1}$: Newton Newtonverfahren.
Für $G_k = I$: schnellstes Abstiegsverfahren.

Es gilt

$$\underbrace{g(x_{k+1}) - g(x_k)}_{q_k} \approx \nabla^2 f(x_k) \underbrace{(x_{k+1} - x_k)}_{p_k}$$

(Annäherung 1. Ordnung, resultiert aus Taylorentwicklung). Setze

$$\begin{aligned} q_k &:= g(x_{k+1}) - g(x_k) \\ p_k &:= x_{k+1} - x_k. \end{aligned}$$

G_{k+1} wird als „gute“ Approximation für $(\nabla^2 f(x_{k+1}))^{-1}$ bezeichnet, wenn

$$G_{k+1} \cdot q_k = p_k$$

(*Quasi-Newton-Bedingung*) gilt. Wir haben also $\frac{n^2+n}{2}$ Freiheitsgrade (Variablen), aber nur n Gleichungen, also ein stark unterbestimmtes Gleichungssystem.

Unterschiedliche Ansätze unterscheiden sich durch die Zusatzforderungen an G_k.

Als recht erfolgreich hat sich folgender Ansatz herausgestellt (*rank 2 update*):

$$G_{k+1} = G_k + a \cdot uu^t + b \cdot vv^t,$$

wobei $a, b \in \mathbb{R}$ und $u, v \in \mathbb{R}^n$. (uu^t und vv^t sind also Matrizen vom Rang 1, die Summe eine Matrix vom Rang 2.) Wir haben also $2n+2$ Freiheitsgrade. (Spezialfall $b = 0$: *rank 1 update*, nicht sehr zielführend.)

Frage: Wie wählt man a, b, u, v? Quasi-Newton-Bedingung:

$$\begin{aligned} G_{k+1}q_k &= p_k \\ (G_k + auu^t + bvv^t)q_k &= p_k \\ G_kq_k + auu^tq_k + bvv^tq_k &= p_k \\ G_kq_k + a(u^tq_k)u + b(v^tq_k)v &= p_k \end{aligned}$$

Eine mögliche Lösung dieser Gleichung ist durch

$$\begin{aligned} u &= p_k, \\ v &= G_kq_k, \\ a &= \frac{1}{p_k^tq_k}, \\ b &= -\frac{1}{q_k^tG_k^tq_k} \end{aligned}$$

gegeben (*Daridon, Fletcher, Powell*, DFP).

DFP-Algorithmus Start $G_0 := I$, $x_0 \in \mathbb{R}^n$ Startpunkt, $k := 0$.
Iteration k:

- Berechne $g(x_k)$.
- Berechne Suchrichtung $d_k = -G_kg(x_k)$.
- Bestimme Schrittweite t_k (mit geeigneter Schrittweitenstrategie).
- Setze $x_{k+1} := x_k + t_kd_k$.
- Bestimme $g(x_k + 1)$:

$$\begin{aligned} p_k &= x_{k+1} - x_k = -t_kG_kg(x_k) \\ q_k &= g(x_{k+1}) - g(x_k). \end{aligned}$$

- Bestimmte

$$G_{k+1} = G_k + \frac{1}{p_k^tq_k}p_kp_k^t - \frac{1}{q_k^tG_k^tq_k} \cdot G_kq_kq_k^tG_k.$$

Bemerkung *G_k positiv definit $\Longrightarrow$ G_{k+1} positiv definit, $p_k^tp_k > 0$.*

Für den Spezialfall quadratischer Funktionen ist das Verfahren stabil.

Alternative Vorgangsweise Statt $(\nabla^2 f(x_k))^{-1}$ approximiert man $\nabla^2 f(x_k)$. Bisher

$$G_{k+1} q_k = q_k,$$

nun

$$q_k = H_{k+1} p_x.$$

(H_{k+1} ist Approximation für $\nabla^2 f(x_{k+1})$.)

Wieder Rang 2 Ansatz für H_{k+1}. Es resultiert

$$H_{k+1} = H_k + \frac{q_k q_k^t}{p_x^t p_x} - \frac{1}{p_k^t H_k p_k} (H_k p_k)(H_k p_k)^t$$

(*Broyoten, Fletcher, Goldfarb, Shanno*, BFGS).

Im BFGS-Verfahren ist für die Berechnung der Suchrichtung das Lösen eines Gleichungssystems nötig:

$$H_k d_k = -g(x_k).$$

Für das BFGS-Verfahren sind bessere Konvergenzresultate bekannt.

14 Optimalitätskriteria für nichtlineare Optimierungsaufgaben mit Nebenbedingungen

Uns interessieren hier Kriterien für *lokale* Extrema (Kandidatenpunkte).

Problemform:

$$\min f(x)$$

s.t.

$$\begin{aligned} h_i(x) &= 0 \qquad & i &= 1, \dots, m \\ g_i(x) &\leq 0 \qquad & i &= 1, \dots, p \end{aligned}$$

mit $x \in \mathbb{R}^n$.

Bezeichne mit

$$\Omega = \{x \in \mathbb{R}^n \mid h_i(x) = 0 \text{ für } i = 1, \dots, m, g_i(x) \leq 0 \text{ für } i = 1, \dots, p\}$$

die Menge der zulässigen Punkte.

Spezialfälle des gesuchten Optimalitätskriteriums:

- $m = p = 0$: $\nabla f(x) = 0$
- $p = 0, m > 0$: Lagrangefunktion

  $$L(x, \lambda) = f(x) + \lambda^t h(x)$$

  mit

  $$\begin{aligned} L_x &= 0 \\ L_\lambda &= 0. \end{aligned}$$

- f, g, h linear: Satz vom komplementären Schlupf

2 einfache Beispiele zur einführenden Betrachtung:

Beispiel

$$\min x_1 + x_2$$

s.t.

$$x_1^2 + x_2^2 - 2 = 0.$$

Für jeden Punkt auf dem Kreisrand $\neq (-1, -1)$ *ist es möglich, zu einem anderen zulässigen Punkt (also auf Kreisrand) mit kleinerem Wert von* f *überzugehen.*

$$\nabla f = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\nabla h = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

Beobachtung: In x^* (und $\tilde{x}$, das ist Lsg. von $\max x_1 + x_2$) sind die Richtungen von ∇f und ∇h parallel zueinander, d.h.

$$\exists \lambda \in \mathbb{R} : \nabla f(x^*) = \lambda \cdot \nabla h(x^*)$$

(ebenso für $\tilde{x}$).

Betrachte die Lagrangefunktion

$$L(x_1, x_2, \mu) = x_1 + x_2 + \mu(x_1^2 + x_2^2 - 2),$$

also

$$\begin{aligned} L_{x_1} &= 1 + 2 \cdot \mu x_1 \stackrel{!}{=} 0 \\ L_{x_2} &= 1 + 2 \cdot \mu x_2 \stackrel{!}{=} 0 \\ L_\mu &= x_1^2 + x_2^2 - 2 \stackrel{!}{=} 0. \end{aligned}$$

Daraus ergibt sich $x_1 = -1, x_2 = -1, \mu = \frac{1}{2}$.

Beispiel

$$\min x_1 + x_2$$

s.t.

$$x_1^2 + x_2^2 - 2 \leq 0.$$

Hier bleibt die Lösung aus dem vorherigen Beispiel erhalten.

$$x_1^2 + x_2^2 - 2 \leq 0$$

ist eine sogenannte aktive Restriktion *(Ungleichung).*

x ist nicht optimal, wenn wir eine Richtung d finden können, sodass die Zulässigkeit bei Fortbewegung entlang d erhalten bleibt und f abnimmt.

Abnahme von f (Kriterium erster Ordnung):

$$(\nabla f(x))^t d < 0. \tag{14.1}$$

Zulässigkeit im Ungleichungsfall ($g(x) \leq 0$):

$$g(x + d) \approx g(x) + (\nabla h(x))^t d \leq 0. \tag{14.2}$$

Suchen Richtung d, die (14.1) und (14.2) erfüllt.

Fall 1 $g(x) < 0$.

Im zweiten Beispiel ist jede Richtung d für x im Kreisinneren zulässigkeitserhaltend.

Die Richtung

$$d = g(x) \frac{\nabla f(x)}{\|\nabla f(x)\|}$$

erfüllt (14.1) und (14.2) für $\nabla f(x) \neq 0$.

Fall 2 $g(x) = 0$.

Erste Bedingung:

$$(\nabla f(x))^t d < 0$$

(offener Halbraum). Zweite Bedingung:

$$(\nabla g(x))^t d \leq 0$$

(abgeschlossener Halbraum).

Es gibt keinen Schnitt der beiden Halbräume, wenn ∇f und ∇g parallel sind.

Beispiel

$$\min x_1 + x_2$$

s.t.

$$\begin{aligned} x_1^2 + x_2^2 - 2 &\leq 0 \\ -x_2 &\leq 0 \end{aligned}$$

Die Lagrangefunktion ist hier

$$L(x_1, x_2, \mu_1, \mu_2) = x_1 + x_2 + \mu_1(x_1^2 + x_2^2 - 2) - \mu_2 x_2$$

und hat die Ableitungen

$$\begin{aligned} L_{x_1} &= 1 + 2\mu_1 x_1 \\ L_{x_2} &= 1 + 2\mu_1 x_2 - \mu_2 \\ L_{\mu_1} &= x_1^2 + x_2^2 - 2 \\ L_{\mu_2} &= -x_2. \end{aligned}$$

Wie im Fall $p = 0$ bauen wir also eine Lagrangefunktion auf:

$$L(x, \lambda, \mu) = f(x) + \lambda^t h(x) + \mu^t g(x) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{p} \mu_j g_j(x).$$

Zunächst Spezialfall mit Ungleichungen, also $m = 0$.

Lemma *Sei $f : \mathbb{R}^n \to \mathbb{R}$, $S \subseteq R^n, S \neq \emptyset$. f sei für $\overline{x} \in S$ differenzierbar.*

1. *Sei $\overline{x}$ lokales Minimum von f über S. Dann gilt*

$$F_0 \cap D = \emptyset,$$

   wobei

$$\begin{aligned} F_0 &= \{d \in \mathbb{R}^n \mid (\nabla f(\overline{x}))^t d < 0\} \\ D &= \{d \in \mathbb{R}^n \mid d \neq 0, \overline{x} + \alpha \cdot d \in S \ \forall \alpha \in (0, \delta) \text{ für ein } \delta > 0\}. \end{aligned}$$

2. *Sei $F_0 \cap D = \emptyset$, f pseudokonvex in $\overline{x}$ (bedeutet* nicht *das gleiche wie quasikonvex!). Es gebe eine ε-Umgebung $U_\varepsilon(\overline{x})$ von $\overline{x}$, sodass $d = x - \overline{x} \in D$ für alle $x \in S \cap U_\varepsilon(\overline{x})$. Dann ist $\overline{x}$ ein lokales Minimum.*

Beweisskizze Angenommen es gibt ein $d \in F_0 \cap D$. Dann gibt es ein $\delta_1 > 0$ mit

$$f(\overline{x} + \alpha d) < f(\overline{x}) \quad \forall \alpha \in (0, \delta_1).$$

Da $d \in D$, gibt es weiters ein $d_2 > 0$ mit

$$\overline{x} + \beta \cdot d \in S \quad \forall \beta \in (0, \delta_2),$$

Widerspruch zu $\overline{x}$ lokales Minimum über S. □

Lemma *Sei Ω die zulässige Menge*

$$\Omega = \{x \in \mathbb{R}^n \mid g(x) \leq 0\}.$$

Sei $\overline{x} \in \Omega$ und sei

$$I := \{i \in \{1, \ldots, p\} \mid g_i(\overline{x}) = 0\}$$

die Menge der aktiven Indizes. *($g_i(\overline{x})$ mit $i \in I$ sind die* aktiven Restriktionen.*)*
Seien die $g_i, i \in I$, in $\overline{x}$ differenzierbar. Die restlichen g_i seien in $\overline{x}$ stetig. Sei

$$\begin{aligned} G_0 &:= \{d \in \mathbb{R}^n \mid (\nabla g_i(\overline{x}))^t d < 0 \ \forall i \in I\} \\ G_0' &:= \{d \in \mathbb{R}^n \mid d \neq 0, (\nabla g_i(\overline{x}))^t d \leq 0 \ \forall i \in I\}. \end{aligned}$$

Dann gilt

$$G_0 \subseteq D \subseteq G_0'.$$

Satz 14.1 *Betrachte*

$$\min f(x)$$

s.t.

$$g(x) \leq 0.$$

Sei Ω die Menge der zulässigen Punkte, $\overline{x} \in \Omega$, I die Menge der aktiven Indizes bzgl. $\overline{x}$, f differenzierbar in $\overline{x}$, $g_i, i \in I$, differenzierbar in $\overline{x}$, $g_i, i \neq I$, stetig in $\overline{x}$.

1. *Sei $\overline{x}$ ein lokales Minimum von f. Dann gilt*

   $$F_0 \cap G_0 = \emptyset,$$

   wobei

   $$\begin{aligned} F_0 &= \{d \in \mathbb{R}^n \mid (\nabla f(\overline{x}))^t d < 0\} \\ G_0 &= \{d \in \mathbb{R}^n \mid (\nabla g_i(\overline{x}))^t d < 0 \ \forall i \in I\}. \end{aligned}$$

2. *Sei $F_0 \cap G_0 = \emptyset$, f pseudokonvex in $\overline{x}$, $g_i, i \in I$ pseudokonvex in einer Umgebung von $\overline{x}$. Dann ist $\overline{x}$ ein lokales Minimum.*

Beispiel

$$\min f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2$$

s.t.

$$\begin{aligned} x_1^2 + x_2^2 &\leq 5 \\ x_1 + x_2 &\leq 3 \\ x_1, x_2 &\geq 0. \end{aligned}$$

Hier ist

$$\begin{aligned} g_1(x_1, x_2) &= x_1^2 + x_2^2 - 5 \\ g_2(x_1, x_2) &= x_1 + x_2 - 3 \\ g_3(x_1, x_2) &= -x_1 \\ g_4(x_1, x_2) &= -x_2. \end{aligned}$$

$\overline{x} = \begin{pmatrix} \frac{9}{5} \\ \frac{6}{5} \end{pmatrix}$ *ist ein zulässiger Punkt. $I = \{2\}$, da nur 2. Ungleichung aktiv in $\overline{x}$. Es gilt*

$$\begin{aligned} \nabla f(x_1, x_2) &= \begin{pmatrix} 2(x_1 - 3) \\ 2(x_2 - 2) \end{pmatrix} \\ \nabla g_2(x_1, x_2) &= \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}, \nabla g_2(x_1, x_2) &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \end{aligned}$$

also $\nabla(\overline{x}) = \begin{pmatrix} -\frac{12}{5} \\ -\frac{8}{3} \end{pmatrix}$. *Die zu betrachtenden Mengen sind also*

$$F_0 = \{d \in \mathbb{R}^2 \mid -\frac{12}{3}d_1 - \frac{8}{3}d_2 < 0\}$$
$$G_0 = \{d \in \mathbb{R}^2 \mid d_1 + d_2 < 0\},$$

somit $F_0 \cap G_0 \neq \emptyset$. *(Übung: Bestimme* $d \in F_0 \cap G_0$.*) Also ist* $\overline{x}$ *kein lokales Minimum.*

Betrachte zweiten Punkt $\hat{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. *Hier ist* $I = \{1, 2\}$ *und* $\nabla f(\hat{x}) = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$. *Die Mengen sind*

$$F_0 = \{d \in \mathbb{R}^2 \mid -d_1 - d_2 < 0\}$$
$$G_0 = \{d \in \mathbb{R}^2 \mid 4d_1 + 2d_2 < 0, d_1 + d_2 < 0\}.$$

Offensichtlich gilt $F_0 \cap G_0 = \emptyset$. *Da* f, g *konvex sind, ist* $\hat{x}$ *ein lokales Minimum.*

14.1 Karush-Kuhn-Tucker Bedingungen (KKT)

Satz 14.4 *Betrachte*

$$\min f(x)$$

s.t.

$$g(x) \leq 0,$$

wobei $f : \mathbb{R}^n \to \mathbb{R}, g : \mathbb{R}^n \to \mathbb{R}, g = (g_1, \ldots, g_p)$. Ω *sei die zulässige Menge,* $\overline{x} \in \Omega$. I *sei die Menge der aktiven Indizes bzgl.* $\overline{x}$. $f, g_i, i \in I$, *seien in* $\overline{x}$ *differenzierbar.* $g_i, i \notin I$, *stetig in* $\overline{x}$.

Ferner seien die Vektoren $\nabla g_i(\overline{x}), i \in I$, *linear unabhängig.*

Ist $\overline{x}$ *lokales Minimum von* f, *so gibt es Skalare* $\lambda_i, i \in I$, *(Lagrange-Multiplikatoren), mit*

$$\nabla f(\overline{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\overline{x}) = 0 \tag{14.3}$$

und

$$\lambda_i \leq 0 \quad \text{für alle } i \in I. \tag{14.4}$$

Bemerkung (14.3) *bedeutet, dass* $\nabla f(\overline{x})$ *als Linearkombination von* $\nabla g_i(\overline{x}), i \in I$, *darstellbar ist.*

Die linke Seite von (14.3) *entspricht* ∇L *(*L *die Lagrangefunktion, siehe auch folgende Umformulierung von* (14.3) *und* (14.4)*).*

Bemerkung *Formulierung von* (14.3) *und* (14.4) *ohne* I*:*

Voraussetzung: $g_i, i = 1, \ldots, p$, *differenzierbar in* $\overline{x}$, $\nabla g_i(\overline{x}), i = 1, \ldots, p$, *linear unabhängig.*

Es gibt $\lambda_i, i = 1, \ldots, p$, *sodass*

$$\nabla f(\overline{x}) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\overline{x}) = 0 \tag{14.5}$$

und

$$\lambda_i \geq 0 \quad \text{für } i = 1, \ldots, p. \tag{14.6}$$

Es muss erzwungen werden, dass $\lambda_i = 0$, *wenn* $i \notin I$, *also*

$$i \in I \Leftrightarrow g_i(\overline{x}) = 0.$$

Somit

$$\lambda_i \cdot g_i(\overline{x}) = 0 \quad \text{für } i = 1, \ldots, p \tag{14.7}$$

(Komplementaritätsbedingung).

Definition *Ein Punkt* $\overline{x}$*, der* (14.3) *und* (14.4) *(bzw.* (14.5), (14.6) *und* (14.7)*) erfüllt, heißt* KKT-Punkt*.*

Satz 14.4 liefert notwendige Bedingung für das Vorliegen eines lokalen Minimums. Ein hinreichendes Kriterium 1. Ordnung liefert der folgende Satz.

Definition (pseudokonvex) $f : S \to \mathbb{R}, S \subseteq \mathbb{R}^n$ *heißt* pseudokonvex*, wenn für alle* $x_1, x_2 \in S$

$$\nabla f(x_1)^t (x_2 - x_1) \geq 0 \to f(x_2) \geq f(x_1)$$

gilt.

Satz 14.5 (Hinreichendes Optimalitätskriterium 1. Ordnung) *Betrachte*

$$\min f(x)$$

s.t.

$$g(x) \leq 0.$$

Sei $\overline{x}$ *ein KKT-Punkt, I die Menge der aktiven Indizes bzgl.* $\overline{x}$*,*

$$\overline{\Omega} = \{x : g_i(x) \leq 0, i \in I\}.$$

Wenn es eine ε*-Umgebung* $U_\varepsilon(\overline{x})$ *mit* $\varepsilon > 0$ *gibt, sodass* f *pseudokonvex für* $x \in U_\varepsilon(\overline{x} \cap \overline{\Omega}$ *und* $g_i, i \in I$ *differenzierbar in* $\overline{x}$ *und quasikonvex für* $x \in U_\varepsilon(\overline{x}) \cap \overline{\Omega}$*, dann ist* $\overline{x}$ *ein lokales Minimum.*

Betrachte (P)

$$\min f(x)$$

s.t.

$$g(x) \leq 0$$
$$h(x) = 0.$$

Satz 14.6 (KKT-Bedingungen) *Sei* $\overline{x} \in \Omega$ *(d.h.* $g(\overline{x}) \leq 0$*,* $h(\overline{x}) = 0$*). Sei weiters I die Menge der aktiven Indizes (*$g_i(\overline{x}) = 0$ *für* $i \in I$*).* $f, g_i, i \in I$*, seien differenzierbar in* $\overline{x}$ *und* $g_i, i \notin I$*, stetig in* $\overline{x}$*.* $h_i, i = 1, \ldots, m$*, sei differenzierbar in* $\overline{x}$*. Außerdem seien* $\{g_i(\overline{x}), i \in I\}$ *linear unabhängig, ebenso* $\{h_i(\overline{x}), i \in I\}$*.*

Wenn $\overline{x}$ *ein lokales Minimum ist, so gibt es Skalare* $\lambda_i, i \in I$*, und* $\mu_j, j = 1, \ldots, m$*, mit*

$$\nabla f(\overline{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\overline{x}) + \sum_{j=1}^{m} \mu_j \nabla h_j(\overline{x}) = 0 \tag{14.8}$$

und

$$\lambda_i \geq 0 \quad \text{für alle } i \in I. \tag{14.9}$$

Bemerkung *Ohne die Menge I heißt das*

$$\nabla f(\overline{x}) + \sum_{i=1}^{p} \lambda_i g_i(\overline{x}) + \sum_{j=1}^{m} \mu_j \nabla h_j(\overline{x}) = 0 \tag{14.10}$$

und

$$\lambda_i \geq 0 \qquad i = 1, \ldots, p \tag{14.11}$$
$$\lambda_i \cdot g_i(\overline{x}) = 0 \qquad i = 1, \ldots, p. \tag{14.12}$$

(Außerdem muss die Zulässigkeit von $\overline{x}$ *gefordert werden, d.h.* $g_i(\overline{x}) \leq 0, h_i(\overline{x}) = 0$*.)*

Bemerkung *Der Satz 14.6 verallgemeinert die Bedingungen an den Gradienten, Lagrangemultiplikatoren und den Satz vom komplementären Schlupf.*

Satz 14.7 (Hinreichende Optimalitätskriterium 1. Ordnung) *Sei (P)*

$$\min f(x)$$

s.t.

$$\begin{aligned} g(x) &\leq 0 \\ h(x) &= 0, \end{aligned}$$

$\overline{x} \in \Omega$. *Weters sei*

$$\begin{aligned} I &= \{i \in \{1, \ldots, p\} \mid g_i(\overline{x}) = 0\} \\ J &= \{j \in \{1, \ldots, m\} \mid \mu_j \geq 0\} \\ K &= \{j \in \{1, \ldots, m\} \mid \mu_j < 0\}. \end{aligned}$$

Seien f pseudokonvex in $\overline{x}$ und $g_i, i \in I$, $h_j, j \in J$, $h_j, j \in K$, quasikonvex in $\overline{x}$.

Dann ist $\overline{x}$ ein lokales Minimum, wenn obige Bedingungen nicht nur in $\overline{x}$, sondern auch in einer ε-Umgebung um $\overline{x}$ gelten.

Beispiel

$$\min 2x_1^2 - 2x_1x_2 + x_2^2 - 15x_1 - 15x_2$$

s.t.

$$\begin{aligned} x_1 + x_2 &\leq 30 \\ x_1 &\geq 0 \\ x_2 &\geq 0. \end{aligned}$$

Forme um auf

$$\begin{aligned} g_1(x_1, x_2) &= x_1 + x_2 - 30 \leq 0 \\ g_2(x_1, x_2) &= -x_1 \leq 0 \\ g_3(x_1, x_2) &= -x_2 \leq 0. \end{aligned}$$

Berechne Gradienten:

$$\nabla f(x_1, x_2) = \begin{pmatrix} 4x_1 - 2x_2 - 15 \\ 2x_2 - 2x_1 - 15 \end{pmatrix} \qquad \nabla g_1(x_1, x_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$\nabla g_2(x_1, x_2) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \qquad \nabla g_3(x_1, x_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Die KKT-Bedingungen lauten hier:

$$\begin{aligned} 4x_1 - 2x_2 - 15 + \lambda_1 - \lambda_2 &= 0 \\ 2x_2 - 2x_1 - 15 + \lambda_1 - \lambda_3 &= 0 \\ \lambda_1(x_1 + x_2 - 30) &= 0 \\ -\lambda_2 x_1 &= 0 \\ -\lambda_3 x_2 &= 0 \\ \lambda_1, \lambda_2, \lambda_3 &\geq 0 \end{aligned}$$

und die Zulässigkeitsbedingungen

$$\begin{aligned} x_1 + x_2 &\leq 30 \\ x_1, x_2 &\geq 0. \end{aligned}$$

Typische zwei Fragestellungen:

1. *Bestimme alle KKT-Punkte (geht nur für kleine Probleme, eher theoretisch).*
2. *Gegeben ein $\overline{x}$, teste ob $\overline{x}$ ein KKT-Punkt ist (Zulässigkeitsproblem für lineares System, einfach lösbar).*

1. *Fallunterscheidung: Jedes λ_i kann 0 oder ≥ 0 sein. ($\lambda_i > 0 \Rightarrow g_i(x) = 0$.)*

1. Fall $x_1 > 0, x_2 > 0$.
Daraus folgt $\lambda_2 = 0$ und $\lambda_3 = 0$.

$$\begin{aligned} 4x_1 - 2x_2 - 15 + \lambda_1 &= 0 \\ -2x_1 + 2x_2 - 15 + \lambda_1 &= 0 \\ \lambda_1 \cdot (x_1 + x_2 - 30) &= 0. \end{aligned}$$

Fall 1a $\lambda_1 = 0$.

$$\begin{aligned} 4x_1 - 2x_3 - 15 &= 0 \\ -2x_1 + 2x_2 - 15 &= 0, \end{aligned}$$

somit $x_1 = 15 > 0$ und $x_2 = \frac{45}{2} > 0$. Hier gilt aber $x_1 + x_2 > 30$, der Punkt ist also nicht zulässig und damit kein KKT-Punkt.

Fall 1b $\lambda_1 \neq 0$.
Daraus folgt $x_1 + x_2 = 30$ und es gibt sich $x_1 = 12 > 0$ und $x_2 = 18 > 0$. $P_1(12, 18)$ ist also ein KKT-Punkt.

2. Fall $x_1 = x_2 = 0$.
Daraus folgt $x_1 + x_2 < 30$ und somit $\lambda_1 = 0$. Es soll

$$\begin{aligned} -15 - \lambda_2 &= 0 \\ -15 - \lambda_3 &= 0 \end{aligned}$$

gelten, also $\lambda_2 = -15 < 0$. Es handelt sich also um keinen KKT-Punkt.

3. Fall $x_1 > 0, x_2 = 0$.
Daraus folgt $\lambda_2 = 0$ und

$$\begin{aligned} 4x_1 - 15 + \lambda_1 &= 0 \\ -2x_1 - 15 + \lambda_1 - \lambda_3 &= 0 \\ \lambda_1 \cdot (x_1 - 30) &= 0. \end{aligned}$$

Fall 3a $\lambda_1 = 0$.
Daraus folgt $x_1 = \frac{15}{4}$ und $x_3 = -\frac{45}{2} < 0$, also kein KKT-Punkt.

Fall 3b $\lambda_1 \neq 0$.
Somit $x_1 = 30, \lambda_1 = -105 < 0$, also auch kein KKT-Punkt.

4. Fall $x_1 = 0, x_2 > 0$.
Daraus folgt $\lambda_3 = 0$ und

$$\begin{aligned} -2x_1 - 15 + \lambda_1 - \lambda_2 &= 0 \\ 2x_2 - 15 + \lambda_1 &= 0. \end{aligned}$$

Fall 4a $\lambda_1 \neq 0$.
Somit $x_2 = 30, \lambda_1 = -45 < 0$, *kein KKT-Punkt.*
Fall 4b $\lambda_1 = 0$.
Somit $x_2 = \frac{15}{2}, \lambda_2 = -30 < 0$, *kein KKT-Punkt.*

$P_1(12, 18)$ *ist also der einzige KKT-Punkt.*

2. *Test, ob* $(15, 15)$ *ein KKT-Punkt ist:*

$$\begin{aligned} 60 - 30 - 15 + \lambda_1 - \lambda_2 &= 0 \\ 30 - 30 - 15 + \lambda_1 - \lambda_3 &= 0 \\ \lambda_3 &= 0 \\ \lambda_3 &= 0. \end{aligned}$$

Gleichungssystem nicht lösbar, also kein KKT-Punkt.

A Programmpakete

Solver für lineare Programme

- *CPLEX* zu bevorzugen.
- *Lpsolve*: nicht kommerziell, deutlich schlechter.
- *Minos*: schlecht geeignet!

Sätzeverzeichnis

Index