

Kapitel 9

Dualitätssätze der linearen Optimierung

Definition 9.1 **Duales lineares Programm.** Sei

$$\begin{aligned} z = \mathbf{c}^T \mathbf{x} &\rightarrow \min ! \\ \mathbf{A} \mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned} \tag{9.1}$$

mit $\mathbf{c}, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ ein lineares Programm.

Das lineare Programm

$$\begin{aligned} \tilde{z} = \mathbf{b}^T \mathbf{y} &\rightarrow \max ! \\ \mathbf{A}^T \mathbf{y} &\leq \mathbf{c} \end{aligned} \tag{9.2}$$

mit $\mathbf{y} \in \mathbb{R}^m$ wird das zu (9.1) duale lineare Programm genannt. Man nennt (9.1) primal. \square

Bemerkung 9.2 Ziel. Die Ziele dieses Abschnitts bestehen darin, die Existenz zulässiger Lösungen des dualen linearen Programms, die Relationen der zulässigen Lösungen des primalen und dualen linearen Programms, die Relationen zwischen den Optimallösungen und die Verbesserung der numerischen Verfahren zu untersuchen. Ein duales Analogon zur Simplexmethode soll entwickelt werden. \square

Satz 9.3 *Ist \mathbf{x} eine zulässige Lösung von (9.1) und ist \mathbf{y} eine zulässige Lösung von (9.2), dann gilt $z(\mathbf{x}) \geq \tilde{z}(\mathbf{y})$.*

Beweis: Es gelten $\mathbf{x} \geq \mathbf{0}$ und $\mathbf{c}^T \geq \mathbf{y}^T A$. Damit folgt

$$z(\mathbf{x}) = \mathbf{c}^T \mathbf{x} \geq \mathbf{y}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{b} = \tilde{z}(\mathbf{y}).$$

■

Man nennt diesen Satz auch schwachen Dualitätssatz.

Folgerung 9.4 *Sind \mathbf{x}_0 eine zulässige Lösung von (9.1) und \mathbf{y}_0 eine zulässige Lösung von (9.2) und gilt $z(\mathbf{x}_0) = \tilde{z}(\mathbf{y}_0)$, dann ist \mathbf{x}_0 eine Optimallösung von (9.1) und \mathbf{y}_0 ist eine Optimallösung von (9.2).*

Satz 9.5 Starker Dualitätssatz. *Das primale Problem (9.1) besitzt genau dann eine endliche Optimallösung, wenn das duale Problem (9.2) eine endliche Optimallösung besitzt. In diesem Fall gilt $z_{\min} = \tilde{z}_{\max}$.*

Beweis: 1.) Es existiere ein endliches Minimum des primalen Problems (9.1) und dieses Minimum werde von $\mathbf{x}_0 = (x_1^{(0)}, \dots, x_m^{(0)}, 0, \dots, 0)^T$ angenommen. Dann sind erklärt:

- Die zugehörigen Basisvektoren seien $\mathbf{a}_1, \dots, \mathbf{a}_m$.
- Mit $X = (x_{ij})_{i=1, \dots, m, j=1, \dots, n}$ werden die Darstellungskoeffizienten für alle Spalten von A bezüglich dieser Basisvektoren bezeichnet.
- $\mathbf{z} = (z_1, \dots, z_n)^T$ sei der Vektor, der durch $z_j = \sum_{i=1}^m c_i x_{ij}$, $j = 1, \dots, n$, erzeugt wird.
- $A_0 = (\mathbf{a}_1, \dots, \mathbf{a}_m)$,
- $\mathbf{c}_0 = (c_1, \dots, c_m)^T$.

Wegen des Optimalitätskriteriums der Simplexmethode gilt für \mathbf{x}_0

$$\mathbf{z} \leq \mathbf{c}, \quad (9.3)$$

wegen $z_k - c_k \leq 0$, $k = 1, \dots, n$. Aus der Nebenbedingung und der Definition der Darstellungskoeffizienten folgt

$$A_0 \mathbf{x}_0 = \mathbf{b}, \quad A_0 X = A.$$

Daraus ergibt sich

$$\mathbf{x}_0 = A_0^{-1} \mathbf{b}, \quad X = A_0^{-1} A. \quad (9.4)$$

Weiter erhalten wir aus der Definition von \mathbf{z}

$$\mathbf{c}_0^T X = \mathbf{z}^T \leq \mathbf{c}^T. \quad (9.5)$$

Jetzt setzen wir $\mathbf{y}_0 := A_0^{-T} \mathbf{c}_0$ und zeigen, dass \mathbf{y}_0 eine Optimallösung des dualen Problems (9.2) ist. Die Zulässigkeit von \mathbf{y}_0 folgt aus (9.4) und (9.5)

$$\mathbf{y}_0^T A = \mathbf{c}_0^T A_0^{-1} A = \mathbf{c}_0^T X \leq \mathbf{c}^T.$$

Die Lösung \mathbf{y}_0 ist optimal wegen

$$\tilde{z}(\mathbf{y}_0) = \mathbf{b}^T \mathbf{y}_0 = \mathbf{y}_0^T \mathbf{b} = \mathbf{c}_0^T A_0^{-1} \mathbf{b} = \mathbf{c}_0^T \mathbf{x}_0 = z(\mathbf{x}_0),$$

wobei (9.4) verwendet wurde. Damit liefert \mathbf{y}_0 einen Zielfunktionswert, der mit dem von \mathbf{x}_0 übereinstimmt. Da \mathbf{x}_0 Optimum des primalen Problems ist, ist \mathbf{y}_0 wegen Folgerung 9.4 Optimum des dualen Problems. Insbesondere gilt $z_{\min} = \tilde{z}_{\max}$.

2.) Das Ziel besteht darin, diesen Teil des Beweises auf den ersten Teil zurückzuführen, indem gezeigt wird, dass das duale Problem des dualen Problems (9.2) gerade das primale Problem (9.1) ist. Dazu wird das duale Problem so umgeformt, dass es die Gestalt eines primalen Problems annimmt. Bildet man dann aus dieser Form das duale Problem, erhält man die Behauptung.

Es existiere ein endliches Maximum \tilde{z}_{\max} des dualen Problems (9.2). Wir setzen für einen beliebigen Vektor $\mathbf{y} \in \mathbb{R}^m$, der die Nebenbedingungen von (9.2) erfüllt, $\mathbf{y} = \mathbf{y}_1 - \mathbf{y}_2$, $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^m$, $\mathbf{y}_1, \mathbf{y}_2 \geq \mathbf{0}$. Aus den Nebenbedingungen von (9.2) erhält man

$$A^T (\mathbf{y}_1 - \mathbf{y}_2) + \mathbf{y}_3 = \mathbf{c} \in \mathbb{R}^n,$$

mit den Schlupfvariablenvektor $\mathbf{y}_3 \in \mathbb{R}^n$, $\mathbf{y}_3 \geq \mathbf{0}$. Daraus bilden wir folgendes zu (9.2) äquivalentes Problem, wobei das Vorzeichen der Zielfunktion geändert wird

$$\begin{aligned} \mathbf{b}^T (\mathbf{y}_2 - \mathbf{y}_1) &\rightarrow \min ! \\ -A^T \mathbf{y}_1 + A^T \mathbf{y}_2 - \mathbf{y}_3 &= -\mathbf{c} \\ \mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3 &\geq \mathbf{0}. \end{aligned}$$

Setzt man

$$\begin{aligned} \mathbf{w} &:= \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \end{pmatrix} \in \mathbb{R}^{2m+n}, \quad \mathbf{d} := \begin{pmatrix} -\mathbf{b} \\ \mathbf{b} \\ \mathbf{0}_n \end{pmatrix} \in \mathbb{R}^{2m+n}, \\ \mathcal{A} &:= \begin{pmatrix} -A^T & A^T & -I_n \end{pmatrix} \in \mathbb{R}^{n \times (2m+n)}, \end{aligned}$$

so kann man dieses System in der Form

$$\begin{aligned} \mathbf{d}^T \mathbf{w} &\rightarrow \min ! \\ \mathcal{A} \mathbf{w} &= -\mathbf{c} \\ \mathbf{w} &\geq \mathbf{0} \end{aligned} \tag{9.6}$$

schreiben.

Einerseits ist dieses Problem zum dualen Problem (9.2) äquivalent. Damit besitzt die Zielfunktion von (9.6) nach Voraussetzung ein endliches Optimum $z_{\min}^{(\mathbf{w})}$, für welches gilt $z_{\min}^{(\mathbf{w})} = \tilde{z}_{\max}$.

Andererseits besitzt das Problem (9.6) die gleiche Gestalt wie das primale Problem (9.1). Die duale Aufgabe zu (9.6) hat nun folgende Gestalt: Gesucht ist $\mathbf{x} \in \mathbb{R}^n$ mit

$$-\mathbf{c}^T \mathbf{x} \rightarrow \max !, \quad \mathcal{A}^T \mathbf{x} \leq \mathbf{d},$$

das heißt

$$\begin{aligned} \mathbf{c}^T \mathbf{x} &\rightarrow \min ! \\ -\mathcal{A} \mathbf{x} &\leq -\mathbf{b} \\ \mathcal{A} \mathbf{x} &\leq \mathbf{b} \\ -\mathbf{x} &\leq \mathbf{0}. \end{aligned}$$

Aus den ersten beiden Nebenbedingungen folgt $\mathcal{A} \mathbf{x} = \mathbf{b}$. Damit ist gezeigt, dass (9.1) das duale Problem zu (9.2) ist. Aus dem ersten Teil des Beweises wissen wir, dass die duale Aufgabe zu (9.6) ein endliches Maximum besitzt und dass dieses Maximum mit $z_{\min}^{(\mathbf{w})} = \tilde{z}_{\max}$ übereinstimmt. ■

Jetzt wird der Fall betrachtet, dass die Zielfunktion der primalen Aufgabe nach unten nicht beschränkt ist.

Satz 9.6 *Ist die Zielfunktion $z = \mathbf{c}^T \mathbf{x}$ der primalen Aufgabe (9.1) auf der Menge der zulässigen Lösungen nach unten unbeschränkt, dann besitzt die zugehörige duale Aufgabe (9.2) keine zulässige Lösung. Analog gilt, dass im Falle dass die Zielfunktion der dualen Aufgabe auf der Menge der zulässigen Lösungen nicht nach oben beschränkt ist, die primale Aufgabe keine zulässige Lösung besitzt.*

Beweis: Indirekter Beweis. Sei die Zielfunktion des primalen Problems nicht nach unten beschränkt und sei \mathbf{y} eine zulässige Lösung des dualen Problems, das heißt es gilt $\mathcal{A}^T \mathbf{y} \leq \mathbf{c}^T$. Aus Satz 9.3 folgt dann aber $z(\mathbf{x}) \geq \tilde{z}(\mathbf{y})$ für alle zulässigen Lösungen \mathbf{x} des primalen Problems und die Zielfunktion wäre nach unten beschränkt.

Die zweite Aussage folgt aus der ersten Aussage und daraus, dass das primale Problem (9.1) das duale Problem des dualen Problems (9.2) ist. ■

Folgerung 9.7 *Eine zulässige Lösung \mathbf{x}_0 des primalen Problems (9.1) ist genau dann optimal, wenn eine zulässige Lösung \mathbf{y}_0 des dualen Problems (9.2) existiert, mit $\mathbf{c}^T \mathbf{x}_0 = \mathbf{b}^T \mathbf{y}_0$. Eine analoge Aussage gilt, wenn man vom dualen Problem ausgeht.*

Satz 9.8 Komplementaritätssatz. *Es sei $\mathbf{x}_0 = (x_1^{(0)}, \dots, x_m^{(0)}, 0, \dots, 0)^T$ eine zulässige Basislösung des primalen Problems (9.1). Dann ist \mathbf{x}_0 genau dann optimal, wenn es eine zulässige Lösung \mathbf{y} des dualen Problems (9.2) mit folgenden Eigenschaften gibt:*

- 1) für alle Indizes $i \in \{1, \dots, m\}$ mit $x_i^{(0)} > 0$ gilt $\mathbf{a}_i^T \mathbf{y} = c_i$,
- 2) für alle Indizes $j \in \{1, \dots, n\}$ mit $\mathbf{a}_j^T \mathbf{y} < c_j$ gilt $x_j^{(0)} = 0$.

Beweis: i) Sei \mathbf{x}_0 optimal. Nach Folgerung 9.7 gibt es dann ein \mathbf{y}_0 mit $\mathbf{c}^T \mathbf{x}_0 = \mathbf{b}^T \mathbf{y}_0$. Einsetzen von $A\mathbf{x}_0 = \mathbf{b}$ ergibt

$$\mathbf{c}^T \mathbf{x}_0 = \mathbf{b}^T \mathbf{y}_0 = (A\mathbf{x}_0)^T \mathbf{y}_0 = \underbrace{\mathbf{x}_0^T A^T}_{\in \mathbb{R}} \mathbf{y}_0 = \mathbf{y}_0^T A\mathbf{x}_0 \iff (\mathbf{c}^T - \mathbf{y}_0^T A) \mathbf{x}_0 = 0.$$

oder in Summenschreibweise

$$\sum_{j=1}^n (c_j - \mathbf{y}_0^T \mathbf{a}_j) x_j^{(0)} = 0.$$

Da \mathbf{y}_0 eine zulässige Lösung des dualen Problems ist und $\mathbf{x}^{(0)}$ eine zulässige Lösung des primalen Problems, sind alle Faktoren nichtnegativ. Damit die Summe Null wird, müssen alle Summanden verschwinden und wenigstens jeweils einer der Faktoren Null sein. Ist $x_j^{(0)} > 0$, muss $\mathbf{a}_j^T \mathbf{y}_0 = c_j$ sein. Ist $\mathbf{a}_j^T \mathbf{y}_0 < c_j$, so muss $x_j^{(0)} = 0$ sein. Die Optimallösung \mathbf{y}_0 des dualen Problems erfüllt also die Bedingungen 1) und 2).

ii) Es gibt einen Vektor \mathbf{y} der die Bedingungen 1) und 2) erfüllt. Wir nehmen an, \mathbf{x}_0 sei nicht optimal. Dann gilt $\mathbf{c}^T \mathbf{x}_0 > \mathbf{b}^T \mathbf{y}$. Analog zum ersten Teil erhält man

$$\sum_{j=1}^n (c_j - \mathbf{y}^T \mathbf{a}_j) x_j^{(0)} > 0.$$

Aus den Bedingungen 1), 2) folgt jedoch, dass die Summe verschwindet. Damit ist die Annahme falsch und \mathbf{x}_0 ist optimal. ■

Bemerkung 9.9 Mit Hilfe der Dualität ist die Möglichkeit der Bestimmung von Schranken für eine zulässige (optimale) Lösung gegeben. Es gilt

$$z(\mathbf{x}) \geq z(\mathbf{x}_0) = \tilde{z}(\mathbf{y}_0) \geq \tilde{z}(\mathbf{y}),$$

wobei \mathbf{x} eine zulässige Lösung des primalen Problems (9.1), \mathbf{x}_0 eine Optimallösung von (9.1), \mathbf{y} eine zulässige Lösung des dualen Problems (9.2) und \mathbf{y}_0 eine Optimallösung des dualen Problems ist. □

Bemerkung 9.10 Symmetrisches duales Programm. Ein Spezialfall des dualen linearen Programms ist das symmetrische duale lineare Programm. Gegeben sei das lineare Programm

$$\begin{aligned} z = \mathbf{c}^T \mathbf{x} &\rightarrow \min ! \\ A\mathbf{x} &\geq \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}. \end{aligned} \tag{9.7}$$

Aus (9.7) wird ein lineares Programm

$$\begin{aligned} \tilde{z} = \mathbf{b}^T \mathbf{y} &\rightarrow \max ! \\ A^T \mathbf{y} &\leq \mathbf{c} \\ \mathbf{y} &\geq \mathbf{0}. \end{aligned} \tag{9.8}$$

konstruiert. □

Satz 9.11 Die linearen Programm (9.7) und (9.8) sind duale lineare Programme im Sinne von Definition 9.1.

Beweis: Aus (9.7) konstruieren wir das lineare Programm in Normalform

$$z = \mathbf{c}^T \mathbf{x} \rightarrow \min !, \quad A\mathbf{x} - \mathbf{v} = \mathbf{b}, \quad \mathbf{x}, \mathbf{v} \geq \mathbf{0}.$$

Aus Definition 9.1 ergibt sich das folgende duale lineare Programm

$$\tilde{z} = \mathbf{b}^T \mathbf{y} \rightarrow \max !, \quad A^T \mathbf{y} \leq \mathbf{c}, \quad -I_m \mathbf{y} \leq \mathbf{0}.$$

Die letzte Bedingung ist äquivalent zu $\mathbf{y} \geq \mathbf{0}$. ■

Man hat nun eine Nichtnegativitätsbedingung an die zulässigen Lösungen des dualen Programms (9.8).

Satz 9.12 Komplementaritätssatz. *Sind \mathbf{x}_0 eine zulässige Lösung von (9.7) und \mathbf{y}_0 eine zulässige Lösung von (9.8), so sind sie genau dann optimal, wenn die folgenden Relationen erfüllt sind:*

$$\mathbf{y}_0^T (A\mathbf{x}_0 - \mathbf{b}) = 0, \tag{9.9}$$

$$(\mathbf{y}_0^T A - \mathbf{c}^T) \mathbf{x}_0 = 0. \tag{9.10}$$

Beweis: 1) Seien \mathbf{x}_0 und \mathbf{y}_0 optimal. Dann folgt aus der Nebenbedingung von (9.8), aus der Nichtnegativität von \mathbf{x}_0 und aus Folgerung 9.7

$$\mathbf{y}_0^T (A\mathbf{x}_0 - \mathbf{b}) = \mathbf{y}_0^T A\mathbf{x}_0 - \mathbf{y}_0^T \mathbf{b} \leq \mathbf{c}^T \mathbf{x}_0 - \mathbf{y}_0^T \mathbf{b} = 0.$$

Andererseits gilt mit $\mathbf{y}_0 \geq \mathbf{0}$ und der Nebenbedingung von (9.7)

$$\mathbf{y}_0^T (A\mathbf{x}_0 - \mathbf{b}) \geq 0.$$

Aus beiden Ungleichungen zusammen folgt (9.9). Die Beziehung (9.10) beweist man analog.

2) Gelten jetzt (9.9) und (9.10). Daraus folgt

$$\mathbf{c}^T \mathbf{x}_0 = \mathbf{y}_0^T A\mathbf{x}_0 = \mathbf{y}_0^T \mathbf{b}.$$

Nach Folgerung 9.7 sind \mathbf{x}_0 und \mathbf{y}_0 optimal. ■

Beispiel für Anwendungen des Dualitätsprinzips werden in den Übungsaufgaben behandelt.

Kapitel 10

Die duale Simplexmethode

Bemerkung 10.1 Motivation. Bei der dualen Simplexmethode ist eine Startlösung oftmals leichter angebar als bei der Simplexmethode für das ursprüngliche lineare Programm, da man keine Nichtnegativitätsanforderungen zu erfüllen hat. Des Weiteren ist die duale Simplexmethode ein wichtiges Verfahren zur Lösung von ganzzahligen linearen Programmen, das heißt, von linearen Programmen, bei denen die Lösung ganzzahlig sein soll.

Seien

$$\begin{aligned} z = \mathbf{c}^T \mathbf{x} &\rightarrow \min ! \\ A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned} \tag{10.1}$$

das primale Programm und

$$\begin{aligned} \tilde{z} = \mathbf{b}^T \mathbf{y} &\rightarrow \max ! \\ A^T \mathbf{y} &\leq \mathbf{c} \end{aligned} \tag{10.2}$$

das zugehörige duale Programm. Wir setzen voraus, dass \tilde{z} endlich ist. \square

Definition 10.2 Ecklösung. Ohne Beschränkung der Allgemeinheit sei $A_B = (\mathbf{a}_1, \dots, \mathbf{a}_m)$ eine Basis. Ein Punkt $\mathbf{y} = (y_1, \dots, y_m)^T$ heißt Ecklösung von (10.2), wenn

$$\begin{aligned} \mathbf{a}_i^T \mathbf{y} &= c_i \quad \text{für } i = 1, \dots, m, \\ \mathbf{a}_i^T \mathbf{y} &< c_i \quad \text{für } i = m + 1, \dots, n \end{aligned} \tag{10.3}$$

gelten. \square

Bemerkung 10.3 Ecklösung bedeutet, dass die Nebenbedingungen, die durch die Basisvektoren von A_B gegeben sind, mit Gleichheit erfüllt sind und die Nebenbedingungen mit den Nichtbasisvektoren als echte Ungleichung.

Seien A_B^{-1} die Inverse von A_B mit der Darstellung

$$A_B^{-1} = \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_m \end{pmatrix}, \quad \text{dass heißt } \mathbf{b}_i \mathbf{a}_j = \delta_{ij}, \quad i, j = 1, \dots, m,$$

und $\mathbf{y} = y_1 \mathbf{a}_1 + \dots + y_m \mathbf{a}_m \in \mathbb{R}^m$ ein beliebiger Vektor. Dann folgt (man beachte, die \mathbf{b}_i sind Zeilenvektoren)

$$\mathbf{b}_i \mathbf{y} = y_1 \mathbf{b}_i \mathbf{a}_1 + \dots + y_m \mathbf{b}_i \mathbf{a}_m = y_i, \quad i = 1, \dots, m.$$

Daraus erhält man insbesondere die Darstellung

$$\mathbf{y} = (\mathbf{b}_1 \mathbf{y}) \mathbf{a}_1 + \dots + (\mathbf{b}_m \mathbf{y}) \mathbf{a}_m. \quad (10.4)$$

□

Die Ausartung in der dualen Simplexmethode, die im folgenden Satz ausgeschlossen ist, wird in seinem Beweis definiert, siehe auch Bemerkung 10.5.

Satz 10.4 Hauptsatz der dualen Simplexmethode. *Sei \tilde{z} nach oben beschränkt und sei Ausartung ausgeschlossen. Ist $\mathbf{y} \in \mathbb{R}^m$ eine Ecklösung und gilt $\mathbf{b}_i \mathbf{b} < 0$ für wenigstens ein $i = 1, \dots, m$, so existiert eine Ecklösung $\bar{\mathbf{y}}$ mit größerem Wert der Zielfunktion \tilde{z} .*

Beweis: Seien \mathbf{y} eine Ecklösung, $\mathbf{b}_i \mathbf{b} < 0$ und $\theta > 0$ beliebig. Es wird ein $\bar{\mathbf{y}}$ konstruiert, welches die Bedingungen des Satzes erfüllt.

Man bildet

$$\bar{\mathbf{y}} = \mathbf{y} - \theta \mathbf{b}_l^T.$$

Aus der Eckpunkteigenschaft (10.3) folgt

$$\mathbf{a}_i^T \bar{\mathbf{y}} = \mathbf{a}_i^T \mathbf{y} - \theta \underbrace{\mathbf{a}_i^T \mathbf{b}_l^T}_{=0} = \mathbf{a}_i^T \mathbf{y} = c_i, \quad i = 1, \dots, m, \quad i \neq l, \quad (10.5)$$

$$\mathbf{a}_i^T \bar{\mathbf{y}} = \mathbf{a}_i^T \mathbf{y} - \theta \underbrace{\mathbf{a}_i^T \mathbf{b}_l^T}_{=1} = \mathbf{a}_i^T \mathbf{y} - \theta < c_i. \quad (10.6)$$

Damit man eine zulässige Lösung hat, muss auch

$$\mathbf{a}_i^T \bar{\mathbf{y}} = \mathbf{a}_i^T \mathbf{y} - \theta \mathbf{a}_i^T \mathbf{b}_l^T \leq c_i, \quad i = m+1, \dots, n,$$

gelten. Für wenigstens einen Index i gilt $\mathbf{a}_i^T \mathbf{b}_l^T < 0$. Anderenfalls, falls also $\mathbf{a}_i^T \mathbf{b}_l^T \geq 0$ für $i = m+1, \dots, n$, sind die Nebenbedingungen für beliebig großes θ erfüllt. Damit wäre

$$\tilde{z} = \mathbf{b}^T \bar{\mathbf{y}} = \mathbf{b}^T \mathbf{y} - \theta \underbrace{\mathbf{b}^T \mathbf{b}_l^T}_{<0 \text{ n.V.}}$$

unbeschränkt im Widerspruch zur Voraussetzung.

Wir wählen

$$\theta = \min_{i=m+1, \dots, n; \mathbf{b}_i \mathbf{a}_i < 0} \left(\frac{\mathbf{a}_i^T \mathbf{y} - c_i}{\mathbf{b}_i \mathbf{a}_i} \right) > 0. \quad (10.7)$$

Wir nehmen an, dass θ von genau einem Index $i = k$ bestimmt wird. Sonst hat man Ausartung. Es gelten:

- 1.) $\bar{\mathbf{y}}$ erfüllt (10.5) und (10.6), das heißt, die Basisvektoren die in der Basis verbleiben erfüllen die Nebenbedingung mit Gleichheit und die neue Nichtbasisvariable mit Index l als echte Ungleichung.
- 2.) Für den Index k , der in die Basis aufgenommen werden soll, gilt

$$\mathbf{a}_k^T \bar{\mathbf{y}} = \mathbf{a}_k^T \mathbf{y} - \frac{\mathbf{a}_k^T \mathbf{y} - c_k}{\mathbf{b}_l \mathbf{a}_k} \mathbf{b}_l \mathbf{a}_k = c_k,$$

- 3.) Aus der Wahl von θ folgt

$$\mathbf{a}_i^T \bar{\mathbf{y}} = \mathbf{a}_i^T \mathbf{y} - \theta \mathbf{a}_i^T \mathbf{b}_l^T < c_i, \quad i = m+1, \dots, n, \quad i \neq k.$$

Falls $\mathbf{a}_i^T \mathbf{b}_l^T$ nichtnegativ ist, ist das Erfülltsein dieser Bedingung klar. Ansonsten wurde bei der Wahl von θ gerade der Index k ausgewählt, der die k -te Nebenbedingung $\mathbf{a}_k^T \mathbf{b}_l^T < 0$ für $\bar{\mathbf{y}}$ zu einer Gleichung werden lässt, ohne dass die anderen Nebenbedingungen mit $\mathbf{a}_i^T \mathbf{b}_l^T < 0$ verletzt werden.

Aus 1) – 3) folgt, dass $\bar{\mathbf{y}}$ eine Ecklösung ist. Ferner gilt

$$\tilde{z} = \mathbf{b}^T \bar{\mathbf{y}} = \mathbf{b}^T \mathbf{y} - \theta \mathbf{b}^T \mathbf{b}_l^T > \mathbf{b}^T \mathbf{y}.$$

■

Bemerkung 10.5 Ausartung im dualen Programm. Ist der Index k bei der Wahl von θ in (10.7) nicht eindeutig, so liegt Ausartung vor. \square

Bemerkung 10.6 Unbeschränktheit der Zielfunktion. Die Unbeschränktheit der Zielfunktion \tilde{z} ist an $\mathbf{a}_j^T \mathbf{b}_l^T \geq 0$ für $j = m + 1, \dots, n$, zu erkennen. \square

Satz 10.7 Optimalitätskriterium. Sei \mathbf{y} eine Ecklösung von (10.2) und gelte $\mathbf{b}_i \mathbf{b} \geq 0$ für alle $i = 1, \dots, m$. Dann ist \mathbf{y} die Optimallösung von (10.2) und die Größen $x_i = \mathbf{b}_i \mathbf{b}$ stellen die Basisvariablen der Optimallösung des primalen Problems (10.1) dar.

Beweis: Es gilt mit (10.3)

$$z = \sum_{i=1}^m c_i x_i = \sum_{i=1}^m c_i \mathbf{b}_i \mathbf{b} = \sum_{i=1}^m \mathbf{y}^T \mathbf{a}_i \mathbf{b}_i \mathbf{b} = \mathbf{y}^T \underbrace{\left(\sum_{i=1}^m \mathbf{a}_i \mathbf{b}_i \right)}_{=I_m} \mathbf{b} = \mathbf{b}^T \mathbf{y} = \tilde{z}.$$

Nach dem starken Dualitätssatz, Satz 9.5, folgt die Aussage des Satzes.

Bemerkung zur Summe: Sei $\sum_{i=1}^m \mathbf{a}_i \mathbf{b}_i = C$ mit einer unbekanntenen Matrix C . Multiplikation diese Gleichung von rechts mit \mathbf{a}_j , $j = 1, \dots, m$, ergibt

$$C \mathbf{a}_j = \sum_{i=1}^m \mathbf{a}_i \underbrace{\mathbf{b}_i \mathbf{a}_j}_{=\delta_{ij}} = \mathbf{a}_j$$

für alle \mathbf{a}_j . Da die $\{\mathbf{a}_j\}$ eine Basis des \mathbb{R}^m bilden, gilt $C = I_m$. \blacksquare

Bemerkung 10.8 Duale Simplextablelle. Die duale Simplextablelle hat die Gestalt

i	c_i	Lösung	$m+1$	\dots	k	\dots	n
			c_{m+1}	\dots	c_k	\dots	c_n
1	c_1	$\mathbf{b}_1 \mathbf{b}$	$\mathbf{b}_1 \mathbf{a}_{m+1}$	\dots	$\mathbf{b}_1 \mathbf{a}_k$	\dots	$\mathbf{b}_1 \mathbf{a}_n$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
l	c_l	$\mathbf{b}_l \mathbf{b}$	$\mathbf{b}_l \mathbf{a}_{m+1}$	\dots	$\mathbf{b}_l \mathbf{a}_k$	\dots	$\mathbf{b}_l \mathbf{a}_n$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
m	c_m	$\mathbf{b}_m \mathbf{b}$	$\mathbf{b}_m \mathbf{a}_{m+1}$	\dots	$\mathbf{b}_m \mathbf{a}_k$	\dots	$\mathbf{b}_m \mathbf{a}_n$
		\tilde{z}	$\mathbf{a}_{m+1}^T \mathbf{y} - c_{m+1}$	\dots	$\mathbf{a}_k^T \mathbf{y} - c_k$	\dots	$\mathbf{a}_n^T \mathbf{y} - c_n$

Wie bei der Simplexmethode, wird die Zeile l Hauptzeile und die Spalte k Hauptspalte genannt. Das Pivotelement ist $\mathbf{b}_l \mathbf{a}_k$. Aus den Nebenbedingungen des dualen linearen Programms (10.2) folgt, dass die Einträge in der letzten Zeile im Nichtbasisteil nichtpositiv sind.

- In der Schlusszeile stehen die Größen, die man zur Berechnung von θ in (10.7) benötigt.
- Die Spalte *Lösung* enthält eine Basislösung des primalen Problems. Sei $\tilde{\mathbf{x}} \in \mathbb{R}^m$ der Vektor mit den Basisvariablen des primalen Problems. Aus den Nebenbedingungen des primalen Problems folgt

$$A_B \tilde{\mathbf{x}} = \mathbf{b} \implies \tilde{\mathbf{x}} = A_B^{-1} \mathbf{b} = \begin{pmatrix} \mathbf{b}_1 \mathbf{b} \\ \vdots \\ \mathbf{b}_m \mathbf{b} \end{pmatrix}.$$

Diese Basislösung ist im allgemeinen nicht zulässig, da sie negative Komponenten besitzt. Gilt jedoch $\mathbf{b}_i \mathbf{b} \geq 0$ für alle $i = 1, \dots, m$, dann ist sie primale Optimallösung, siehe Satz 10.7.

- Die Nichtbasisvektoren lassen sich als Linearkombination der Basisvektoren darstellen $\mathbf{a}_j = A_B \mathbf{a}$ mit einem unbekanntem Koeffizientenvektor \mathbf{a} . Dieser lässt sich durch $\mathbf{a} = A_B^{-1} \mathbf{a}_j$ berechnen, was in Komponentenschreibweise zu

$$\mathbf{a}_j = \sum_{i=1}^m (\mathbf{b}_i \mathbf{a}_j) \mathbf{a}_i, \quad j = m+1, \dots, n$$

führt, siehe (10.4).

- Falls man das Optimum des dualen Problems mit der dualen Simplexmethode gefunden hat, ist die Tabelle der dualen Simplexmethode eine Optimaltabelle für die primale Aufgabe.

□

Bemerkung 10.9 Herleitung der Transformationsregeln. Sei $A_B = (\mathbf{a}_1, \dots, \mathbf{a}_m)$ und sei $\hat{A}_B = (\mathbf{a}_1, \dots, \mathbf{a}_{l-1}, \mathbf{a}_k, \mathbf{a}_{l+1}, \dots, \mathbf{a}_m)$. Dann ist

$$\begin{aligned} A_B^{-1} \hat{A}_B &= \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_m \end{pmatrix} (\mathbf{a}_1, \dots, \mathbf{a}_{l-1}, \mathbf{a}_k, \mathbf{a}_{l+1}, \dots, \mathbf{a}_m) \\ &= \begin{pmatrix} \mathbf{b}_1 \mathbf{a}_1 & \mathbf{b}_1 \mathbf{a}_2 & \cdots & \mathbf{b}_1 \mathbf{a}_k & \cdots & \mathbf{b}_1 \mathbf{a}_m \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{b}_l \mathbf{a}_1 & \mathbf{b}_l \mathbf{a}_2 & \cdots & \mathbf{b}_l \mathbf{a}_k & \cdots & \mathbf{b}_l \mathbf{a}_m \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{b}_m \mathbf{a}_1 & \mathbf{b}_m \mathbf{a}_2 & \cdots & \mathbf{b}_m \mathbf{a}_k & \cdots & \mathbf{b}_m \mathbf{a}_m \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & \cdots & 0 & \mathbf{b}_1 \mathbf{a}_k & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \mathbf{b}_2 \mathbf{a}_k & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \mathbf{b}_l \mathbf{a}_k & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \mathbf{b}_m \mathbf{a}_k & 0 & \cdots & 1 \end{pmatrix} =: I_m^{(k)}. \end{aligned}$$

Mit

$$\left(\hat{A}_B \right)^{-1} = \begin{pmatrix} \hat{\mathbf{b}}_1 \\ \vdots \\ \hat{\mathbf{b}}_m \end{pmatrix} \implies A_B^{-1} = I_m^{(k)} \left(\hat{A}_B \right)^{-1},$$

oder

$$\begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_l \\ \vdots \\ \mathbf{b}_m \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{b}}_1 + (\mathbf{b}_1 \mathbf{a}_k) \hat{\mathbf{b}}_k \\ \vdots \\ (\mathbf{b}_l \mathbf{a}_k) \hat{\mathbf{b}}_k \\ \vdots \\ \hat{\mathbf{b}}_m + (\mathbf{b}_m \mathbf{a}_k) \hat{\mathbf{b}}_k \end{pmatrix}.$$

Das Ziel besteht darin, die Vektoren \mathbf{b}_i durch die Vektoren $\hat{\mathbf{b}}_i$ zu ersetzen. Im Prinzip stehen nun die Transformationsregeln da. Für das Pivotelement gilt

$$\hat{\mathbf{b}}_k = \frac{\mathbf{b}_l}{\mathbf{b}_l \mathbf{a}_k} \implies \hat{\mathbf{b}}_k \mathbf{a}_l = \frac{\mathbf{b}_l \mathbf{a}_l}{\mathbf{b}_l \mathbf{a}_k} = \frac{1}{\mathbf{b}_l \mathbf{a}_k}.$$

Für die Hauptspalte erhält man daraus

$$\hat{\mathbf{b}}_i = \mathbf{b}_i - (\mathbf{b}_i \mathbf{a}_k) \hat{\mathbf{b}}_k \implies \hat{\mathbf{b}}_i \mathbf{a}_l = \mathbf{b}_i \mathbf{a}_l - (\mathbf{b}_i \mathbf{a}_k) \left(\hat{\mathbf{b}}_k \mathbf{a}_l \right) = 0 - \frac{\mathbf{b}_i \mathbf{a}_k}{\mathbf{b}_l \mathbf{a}_k} \quad i \neq l.$$

Für die Hauptzeile ergibt sich unmittelbar, wobei $\mathbf{a}_0 := \mathbf{b}$ ist,

$$\hat{\mathbf{b}}_k \mathbf{a}_j = \frac{\mathbf{b}_l \mathbf{a}_j}{\mathbf{b}_l \mathbf{a}_k}, \quad j = 0, m+1, \dots, n, j \neq k.$$

Damit ergibt sich auch die Rechteckregel

$$\hat{\mathbf{b}}_i \mathbf{a}_j = \mathbf{b}_i \mathbf{a}_j - \mathbf{b}_i \mathbf{a}_k \hat{\mathbf{b}}_k \mathbf{a}_j = \mathbf{b}_i \mathbf{a}_j - \frac{\mathbf{b}_l \mathbf{a}_j}{\mathbf{b}_l \mathbf{a}_k} \mathbf{b}_i \mathbf{a}_k.$$

Zusammenfassung: Falls in der Spalte *Lösung* wenigstens ein $\mathbf{b}_i \mathbf{b} < 0$ steht, zum Beispiel für $i = l$, so transformiert man wie folgt:

- setze $\mathbf{a}_0 = \mathbf{b}$,
- vertausche die Indizes l und k ,
- Pivotelement: $\hat{\mathbf{b}}_k \mathbf{a}_l = 1/(\mathbf{b}_l \mathbf{a}_k)$,
- Hauptspalte:

$$\hat{\mathbf{b}}_i \mathbf{a}_l = -\frac{\mathbf{b}_i \mathbf{a}_k}{\mathbf{b}_l \mathbf{a}_k}, \quad i = 1, \dots, m, i \neq l,$$

- Hauptzeile:

$$\hat{\mathbf{b}}_k \mathbf{a}_j = \frac{\mathbf{b}_l \mathbf{a}_j}{\mathbf{b}_l \mathbf{a}_k}, \quad j = 0, m+1, \dots, n, j \neq k,$$

- Rechteckregel:

$$\hat{\mathbf{b}}_i \mathbf{a}_j = \mathbf{b}_i \mathbf{a}_j - \frac{\mathbf{b}_l \mathbf{a}_j}{\mathbf{b}_l \mathbf{a}_k} \mathbf{b}_i \mathbf{a}_k, \quad i = 1, \dots, m, i \neq l, j = 0, m+1, \dots, n, j \neq k.$$

Die Rechteckregel wird auch auf die letzte Zeile angewandt *Übungsaufgabe*. Das sind dieselben Regeln wie im primalen Fall! \square

Beispiel 10.10 Wir betrachten noch einmal das Problem aus Beispiel 5.5:

$$\begin{aligned} z = -3x_1 - 2x_2 - 4x_3 - x_4 &\rightarrow \min ! \\ \begin{pmatrix} 2 & 2 & 3 & 0 & 1 & 0 & 0 \\ 1 & 3 & 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 5 & 2 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_7 \end{pmatrix} &= \begin{pmatrix} 700 \\ 400 \\ 500 \end{pmatrix} \\ \mathbf{x} &\geq \mathbf{0}. \end{aligned}$$

In Beispiel 5.5 hatten wir die Optimallösung $\mathbf{x} = (320, 0, 20, 40, 0, 0, 0)^T$ mit $z = -1080$ erhalten. Das duale Problem zum obigen linearen Programm lautet

$$\begin{aligned} \tilde{z} = 700y_1 + 400y_2 + 500y_3 &\rightarrow \max ! \\ \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 1 \\ 3 & 0 & 5 \\ 0 & 2 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} &\leq \begin{pmatrix} -3 \\ -2 \\ -4 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

Wir nehmen uns die erste, dritte und fünfte Nebenbedingung des dualen Problems her und betrachte diese Bedingungen als Gleichungen:

$$\begin{pmatrix} 2 & 1 & 1 \\ 3 & 0 & 5 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -3 \\ -4 \\ 0 \end{pmatrix} \implies \mathbf{y} = \begin{pmatrix} 0 \\ -11/5 \\ -4/5 \end{pmatrix}.$$

Durch Einsetzen in die anderen Nebenbedingungen verifiziert man, dass man damit eine Ecklösung des dualen Problems gefunden hat. Es ist

$$A_B = (\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5) = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 0 & 0 \\ 1 & 5 & 0 \end{pmatrix},$$

$$A_B^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1/5 & 1/5 \\ 1 & -7/5 & -3/5 \end{pmatrix} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_3 \\ \mathbf{b}_5 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 700 \\ 400 \\ 500 \end{pmatrix}.$$

Damit kann man alle Größen für die duale Simplextabelle bestimmen:

i	c_i	Lösung	2	4	6	7
1	-3	400	-2	-1	0	0
3	-4	20	3	2	1	0
5	0	-160	-2/5	0	-1/5	1/5
		-1280	-14/5	-4	-7/5	-3/5
			-27/5	-5	-11/5	-4/5

Die Lösung ist nicht optimal, da $-160 < 0$. Damit ist $l = 5$ die Hauptzeile. Zur Bestimmung der Hauptspalte berechnet man

$$\theta = \min_{j \in \{2,4,6,7\}, \mathbf{b}_l \mathbf{a}_j < 0} \left(\frac{\mathbf{a}_j^T \mathbf{y} - c_j}{\mathbf{b}_l \mathbf{a}_j} \right) = \min \left\{ \frac{27}{14}, \frac{5}{4}, \frac{11}{7}, \frac{4}{3} \right\} = \frac{5}{4}.$$

Damit ist die Hauptspalte $k = 4$. Mit den Transformationsregeln der Simplexmethode erhält man die neue duale Simplextabelle

i	c_i	Lösung	2	5	6	7
1	-3	320	-2	0	0	0
3	-4	20	8/5	1/2	3/10	-3/10
4	-1	40	-2/5	0	-1/5	1/5
		-1080	7/10	-1/4	7/20	3/20
			-19/10	-5/4	-9/20	-1/20

Damit ist das Optimum bestimmt. Die Optimallösung des primalen Problems findet man in der Spalte *Lösung* ebenso den zugehörigen Zielfunktionswert.

Die Lösung des dualen Problems \mathbf{y} kann man im allgemeinen nicht direkt aus der dualen Simplextabelle ablesen. In der letzten Zeile steht nämlich $\mathbf{a}_j^T \mathbf{y} - c_j$. Das direkte Ablesen geht nur, wenn $c_j = 0$ und die \mathbf{a}_j Einheitsvektoren sind. Das ist in diesem Beispiel gegeben, nämlich für die Indizes 5, 6, 7. Die Lösung des dualen Problems ist also $\mathbf{y} = (-5/4, -9/20, -1/20)^T$ mit dem Zielfunktionswert $\tilde{z} = -1080$. Im allgemeinen muss man noch ein lineares Gleichungssystem lösen, um die Lösung des dualen Problems zu berechnen. \square

Kapitel 11

Die duale Simplexmethode zur Lösung rein ganzzahliger linearer Programme

Wir betrachten folgendes Optimierungsproblem

$$z = \mathbf{c}^T \mathbf{x} \rightarrow \min !$$
$$A\mathbf{x} = \mathbf{b} \tag{11.1}$$

$$\mathbf{x} \geq \mathbf{0} \tag{11.2}$$

$$x_j \text{ ganz für } j = 1, \dots, n_1 \leq n, \tag{11.3}$$

$$a_{ij}, b_i \text{ ganz für } i = 1, \dots, m, j = 1, \dots, n. \tag{11.4}$$

Definition 11.1 Ganzzahliges lineares Programm. Das lineare Programm (11.1) – (11.4) heißt rein ganzzahliges lineares Programm falls $n_1 = n$. Ansonsten heißt es für $n_1 > 0$ gemischt ganzzahliges lineares Programm. \square

Wir werden nur rein ganzzahlige lineare Programme betrachten.

Bemerkung 11.2 Häufig enthalten ganzzahlige lineare Programme Bedingungen der folgenden Art

$$0 \leq x_j \leq 1, \quad x_j \text{ ganz,}$$

für gewisse Indizes j . \square

Beispiel 11.3 Wir betrachten

$$z = -8x_1 - 4x_2 \rightarrow \min !$$

$$-2x_1 + 3x_2 \leq 6$$

$$8x_1 + 3x_2 \leq 20$$

$$\mathbf{x} \geq \mathbf{0}$$

$$x_1, x_2 \text{ ganz.}$$

Das Problem ohne die Ganzzheitsforderung kann man graphisch lösen, siehe Abbildung 11.1.

Das stetige Optimum ist

$$\mathbf{x} = \left(\frac{7}{5}, \frac{44}{15} \right), \quad z = -\frac{344}{15}.$$

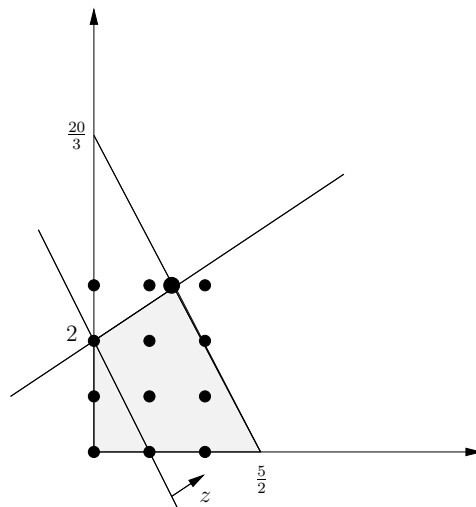


Abbildung 11.1: Illustration zu Beispiel 11.3.

Eine einfache Idee zur Bestimmung des ganzzahligen Optimums ist Runden. Man erhält damit $\mathbf{x} = (1, 3)^T$. Dieser Punkt ist jedoch nicht zulässig. Durch Abrunden folgt $\mathbf{x} = (1, 2)^T$. Man erhält mit diesem Punkt den Zielfunktionswert $z = -16$. Dieser ist jedoch nicht optimal, da man mit $\mathbf{x} = (2, 1)^T$ den Wert $z = -20$ erhält.

Runden ist also keine geeignete Lösungstechnik. \square

Bemerkung 11.4 Motivation für das Schnittprinzip. Mit dem normalen Simplexverfahren kann man nicht arbeiten, da das Optimum ein innerer Punkt des zulässigen Bereiches ist. Das könnte man durch die Bestimmung der konvexen Hülle aller zulässigen Punkte ändern. Dieses Vorgehen ist aber im allgemeinen viel zu aufwendig. Stattdessen versucht man in der Nähe des stetigen Optimums durch Abschneiden den gesuchten ganzzahligen optimalen Punkt zu einem Eckpunkt zu machen. Dieses Schnittprinzip soll jetzt auf eine Art (nach Gomory (1957)) realisiert werden. \square

Definition 11.5 Schnittbedingung. Die Nebenbedingung

$$\sum_{j=1}^n \beta_j x_j \leq \beta \tag{11.5}$$

heißt Schnittbedingung, wenn folgendes erfüllt ist:

- i) Es sei $\mathbf{x}^{(0)}$ ein Optimum mit den Nebenbedingungen (11.1) – (11.2), aber nicht (11.3). Dann erfüllt $\mathbf{x}^{(0)}$ die Bedingung (11.5) nicht, das heißt

$$\sum_{j=1}^n \beta_j x_j^{(0)} > \beta.$$

- ii) Jede Lösung, welche die Nebenbedingungen (11.1) – (11.3) erfüllt, erfüllt auch (11.5), das heißt

$$\{\mathbf{x} : \mathbf{x} \text{ erfüllt (11.1) – (11.3)}\} \subset \{\mathbf{x} : \mathbf{x} \text{ erfüllt (11.5)}\}.$$

\square

Bemerkung 11.6 Vorgehen. Mit Hilfe der Schnittbedingung soll der zulässige Bereich verkleinert werden (eine Ecke wird abgeschnitten) ohne dass damit ganzzahlige Lösungen abgeschnitten werden.

Jetzt soll (11.1) – (11.4) durch die Einführung von Schnittbedingungen gelöst werden. Es sei dazu zuerst (11.1) – (11.2) mit Hilfe der Simplexmethode gelöst. Das Optimum sei $\mathbf{x}^{(0)} = (x_1^{(0)}, \dots, x_m^{(0)}, 0, \dots, 0)^T$. Dabei sei wenigstens ein $x_i^{(0)}$, $i \in \{1, \dots, m\}$, nicht ganz.

Sei $A_B = (\mathbf{a}_1, \dots, \mathbf{a}_m)$ die Matrix der Basisvektoren. Die Auflösung von (11.1) nach den Basisvariablen liefert für jedes zulässige \mathbf{x} die Darstellung

$$x_i = \alpha_i + \alpha_{i,m+1}(-x_{m+1}) + \dots + \alpha_{i,n}(-x_n), \quad i = 1, \dots, m. \quad (11.6)$$

□

Lemma 11.7 Die Koeffizienten von (11.6) sind die Koeffizienten einer optimalen Simplextabelle.

Beweis: Wir betrachten die Nebenbedingung (11.1) und zerlegen $A = (A_B|A_N)$ sowie $\mathbf{x} = (\mathbf{x}_B|\mathbf{x}_N)^T$. Wir wissen, dass $A_N = A_B X$, wobei X die Einträge der Simplextabelle sind. Aus

$$A_B \mathbf{x}_B + A_N \mathbf{x}_N = \mathbf{b}$$

folgt

$$\mathbf{x}_B = A_B^{-1} \mathbf{b} + A_B^{-1} A_N (-\mathbf{x}_N) = \underbrace{A_B^{-1} \mathbf{b}}_{\alpha_i} + \underbrace{X(-\mathbf{x}_N)}_{\text{Rest von (11.6)}} .$$

■

Bemerkung 11.8 Weiteres Vorgehen. Sei jetzt für $i = p$ die Variable $x_p^{(0)}$ nicht ganz, $i \in \{1, \dots, m\}$. Falls mehrere $x_i^{(0)}$ nicht ganz sind, wähle man einen dieser Indizes. Welches der beste ist, ist im allgemeinen nicht zu beantworten.

Sei $a \in \mathbb{R}$. Dann bezeichnen wir

$$[a] = \text{INT}(a), \quad \{a\} = a - [a],$$

wobei $\text{INT}(a)$ der größte ganzzahlige Bestandteil von a ist. Es gilt $\{a\} \in [0, 1)$. Insbesondere gilt $\{x_p^{(0)}\} > 0$. Da für das Optimum die Komponenten mit den Indizes $m+1, \dots, n$ verschwinden und wegen (11.6) folgt damit

$$\alpha_p = x_p^{(0)} = \underbrace{[x_p^{(0)}]}_{\geq 0} + \underbrace{\{x_p^{(0)}\}}_{> 0} > 0.$$

□

Satz 11.9 Die Bedingung

$$s_1 = -\{\alpha_p\} - \{\alpha_{p,m+1}\}(-x_{m+1}) - \dots - \{\alpha_{p,n}\}(-x_n), \quad s_1 \geq 0 \quad (11.7)$$

stellt eine Schnittbedingung gemäß Definition 11.5 dar.

Beweis: Die Bedingungen von Definition 11.5 müssen geprüft werden. Wir fügen die Bedingung (11.7) zum System der Nebenbedingungen (11.1), (11.2) hinzu und setzen $\mathbf{x}^{(0)}$ ein. Aus (11.6) und wegen $x_{m+1}^{(0)} = \dots = x_n^{(0)} = 0$ folgt

$$s_1 = -\alpha_p < 0,$$

also ist $\mathbf{x}^{(0)}$ nicht zulässig.

Nun ist zu zeigen, dass mit (11.7) kein bezüglich (11.1) – (11.3) zulässiger Punkt weggeschnitten wird. Sei $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_n)^T$ ein Punkt, der (11.1) – (11.3) erfüllt, also insbesondere ganzzahlige Komponenten besitzt. Dann folgt aus (11.7)

$$s_1 = - \underbrace{(\alpha_p - [\alpha_p])}_{\in [0,1)} - \sum_{j=m+1}^n \underbrace{(\alpha_{p,j} - [\alpha_{p,j}])}_{\in (0,1)} \underbrace{(-\tilde{x}_j)}_{\leq 0}.$$

Damit ist $s_1 > -1$. Andererseits gilt

$$s_1 = \underbrace{[\alpha_p] + \sum_{j=m+1}^n [\alpha_{p,j}(-\tilde{x}_j)]}_{\in \mathbb{Z}} - \underbrace{\alpha_p - \sum_{j=m+1}^n \alpha_{p,j}(-\tilde{x}_j)}_{(11.6) = -\tilde{x}_p \in \mathbb{Z}}.$$

Damit ist $s_1 \in \mathbb{Z}$. Da $s_1 > -1$ folgt $s_1 \geq 0$. ■

Bemerkung 11.10 Zusammenfassung. Man hat mit der Schnittbedingung (11.7) das Optimum bezüglich der Nebenbedingungen (11.1), (11.2) abgeschnitten, ohne dabei auch ganzzahlige Lösungen wegzuschneiden. Folgendes lineare Programm ist jetzt zu lösen:

$$\begin{aligned} z = \mathbf{c}^T \mathbf{x} &\rightarrow \min ! \\ \mathbf{A} \mathbf{x} &= \mathbf{b} \\ \sum_{j=m+1}^n \{\alpha_{p,j}\} (-x_j) + s_1 &= -\{\alpha_p\} \\ \mathbf{x} &\geq \mathbf{0} \\ s_1 &\geq 0. \end{aligned} \tag{11.8}$$

Zur Lösung von (11.8) berechnet man zuerst die optimale Lösung des linearen Programms ohne Ganzzahligkeitsbedingung mit Hilfe der dualen Simplexmethode. Hat man diese, und ist sie nicht ganzzahlig, betrachtet man im nächsten Schritt die duale Simplextabelle mit

$$x_i = \alpha_i, \quad i = 1, \dots, m, \quad s_1 = -\alpha_p \tag{11.9}$$

(beachte: in der Simplextabelle des dualen Problems steht $x_i = \mathbf{b}_i \mathbf{b} = \alpha_i$). Der Vektor (11.9) ist eine dual zulässige Lösung, das heißt, die Nebenbedingungen des dualen Problems sind erfüllt. *Übungsaufgabe* Eventuell ist die Einführung weiterer Schnittbedingungen nötig. Das oben beschriebene Vorgehen wird im nächsten Beispiel demonstriert. □

Beispiel 11.11 Wir betrachten das lineare Programm

$$\begin{aligned} z = -x_1 - 2x_2 &\rightarrow \min ! \\ \begin{pmatrix} 2 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_5 \end{pmatrix} &= \begin{pmatrix} 10 \\ 5 \\ 4 \end{pmatrix} \\ \mathbf{x} &\geq \mathbf{0} \\ \mathbf{x} &\text{ ganz.} \end{aligned}$$

Die optimale duale Simplextabelle lautet

i	c_i	Lösung	3	4
1	-1	5/3	1/3	-1/3
2	-2	20/3	1/3	2/3
5	0	7/3	-1/3	1/3
		-15	-1	-1

Die Lösung ist optimal, aber nicht ganzzahlig. Heuristisch wählt man $\{\alpha_p\}$ möglichst groß, hier zum Beispiel $\{20/3\} = 2/3$. Es besteht allerdings die Gefahr, dass man an der falschen Stelle abschneidet. Nun verwendet man, dass die Tabelle der dualen Simplexmethode eine Optimaltabelle der primalen Aufgabe ist, Bemerkung 10.8. Damit kann man Lemma 11.7 für die Formulierung der Schnittbedingung nutzen, da die benötigten Koeffizienten $\alpha_{p,j}$ gerade im Nichtbasisteil der ausgewählten Zeile stehen:

$$s_1 = -\frac{2}{3} - \frac{1}{3}(-x_3) - \frac{2}{3}(-x_4) \geq 0.$$

Führt man die Variable s_1 als $m+1$ -ste Eckvariable in die duale Simplextabelle ein, dann hat man die neue Matrix der Basisvektoren

$$\tilde{A}_B = \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_m & \mathbf{0} \\ 0 & \cdots & 0 & 1 \end{pmatrix} \implies \tilde{A}_B^{-1} = \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_m \\ \mathbf{e}_{m+1}^T \end{pmatrix},$$

wobei \mathbf{e}_{m+1} der $m+1$ -ste Einheitsvektor ist. Somit erhält man in der Spalte Lösung für $s_1 : \mathbf{e}_{m+1}^T \mathbf{b} = -\alpha_p = -2/3$. In den Nichteckspalten der Zeile von s_1 erhält man die Zahlen $-\{\alpha_{p,j}\}$, siehe (11.8). Man hat die neue duale Simplextabelle

i	c_i	Lösung	3	4
			0	0
1	-1	5/3	1/3	-1/3
2	-2	20/3	1/3	2/3
5	0	7/3	-1/3	1/3
s_1	0	-2/3	-1/3	-2/3
		-15	-1	-1

Die Hauptzeile ist die Zeile von s_1 . Aus

$$\theta = \min_{j \in \{3,4\}} \left\{ \frac{-1}{-1/3}, \frac{-1}{-2/3} \right\} = \frac{3}{2}$$

folgt, dass $k = 4$ die Hauptspalte ist. Der Simplexschritt führt zu folgender Tabelle

i	c_i	Lösung	3	s_1
			0	0
1	-1	2	-1/2	-1/2
2	-2	6	0	1
5	0	2	-1/2	1/2
4	0	1	1/2	-3/2
		-14	-1/2	-3/2

Damit hat man die Optimallösung des ganzzahligen linearen Programms gefunden.

In der Praxis sind im allgemeinen mehr Schnittbedingungen nötig. Die Endlichkeit des Verfahrens ist nicht gesichert. \square

Kapitel 12

Innere–Punkt–Verfahren

Innere–Punkt–Verfahren verfolgen die Idee, bei der Lösung von linearen Programmen durch das Innere des konvexen Polyeders zum Optimum zu gelangen. Damit unterscheiden sie sich grundsätzlich von der Simplexmethode.

Bemerkung 12.1 Historie von Innere–Punkt–Verfahren.

- Dikin 1967: hat bereits die Idee von Karmarkar (1984) umgesetzt, Arbeit wurde aber nicht wahrgenommen.
- Fiacco, McCormick 1968: Innere–Punkt–Verfahren für nichtlineare Optimierungsprobleme, waren aber nicht wettbewerbsfähig im Vergleich zu anderen Verfahren.
- Khachian 1979: Ellipsoidmethode. Polynomiale Komplexität konnte bewiesen werden, allerdings war das Verfahren wenig praxistauglich wegen numerischen Instabilitäten.
- Karmarkar 1984: projektive Methode, erste praxistaugliche Innere–Punkt–Methode.

□

Bemerkung 12.2 Motivationen. Die Motivation zur Entwicklung von Alternativen zur Simplexmethode liegt in deren schlechter theoretischer Komplexität begründet. Man kann nicht ausschließen, dass die Anzahl der Iterationen exponentiell mit der Problemgröße wächst, siehe Abschnitt 8.2. Das hat im wesentlichen zwei Gründe:

- Die Simplexmethode kennt als Abstiegsrichtungen nur die Kanten auf dem Rand des zulässigen Bereiches. Deren Anzahl wächst exponentiell mit der Dimension des Polyeders.
- Die Simplexmethode sucht die Abstiegsrichtung lokal aus. Für diese Richtung spielen nur die im aktuellen Eckpunkt aktiven Nebenbedingungen eine Rolle, nicht aber die Gesamtheit der Nebenbedingungen.

Alternativen zur Simplexmethode müssen an diesen beiden Schwachstellen angreifen: sie sollten Abstiegsrichtungen durch das Innere des Polyeders zulassen und bei der Bestimmung dieser Richtungen Informationen von allen Nebenbedingungen einbeziehen. □

Bemerkung 12.3 Ellipsoidmethoden. Die Idee von Khachians Ellipsoidmethode lässt sich grob wie folgt beschreiben: Man startet mit einem zulässigen Punkt $\bar{\mathbf{x}}$ und sucht einen möglichst zentral gelegenen Punkt $\tilde{\mathbf{x}}$ in einem Restpolyeder, welches nur Punkte \mathbf{x} mit $\mathbf{c}^T \mathbf{x} \leq \mathbf{c}^T \bar{\mathbf{x}}$ enthält. Diesen Punkt findet man dadurch, dass man mittels einer Iteration das Restpolyeder möglichst gut in ein Ellipsoid einbettet und dessen Mittelpunkt $\tilde{\mathbf{x}}$ betrachtet. Ist $\tilde{\mathbf{x}}$ zulässig, setzt man $\bar{\mathbf{x}} = \tilde{\mathbf{x}}$ und startet die

Konstruktion von Neuem. Die Instabilitäten rührten daher, dass die Ellipsoide im Laufe der Iteration immer schmaler wurden.

Das Verfahren von Karmarkar geht andersherum vor: hier wird ein Ellipsoid in den zulässigen Bereich eingebettet. Es funktioniert grob wie folgt: Man startet mit einem inneren Punkt des zulässigen Bereiches $\bar{\mathbf{x}}$ und konstruiert ein im zulässigen Bereich eingeschriebenes, möglichst großes Ellipsoid. Dann nimmt man anstelle des zulässigen Bereiches nur das Ellipsoid und minimiert darüber die Zielfunktion. Es ergibt sich ein Randpunkt $\tilde{\mathbf{x}}$, an dem das Minimum angenommen wird. Man setzt $\bar{\mathbf{x}} = \tilde{\mathbf{x}}$ und startet die Konstruktion von Neuem.

Das Verfahren von Karmarkar arbeitet praktisch wesentlich besser als die Ellipsoidmethode von Khachian. Man kann auch beweisen, dass die Komplexität des Verfahrens von Karmarkar besser ist. Die Komplexität ist also insbesondere polynomial. Für sehr große Problem ist dieses Verfahren oft schneller als die Simplex-methode. \square

In diesem Kapitel soll jedoch nur eine einfache Innere-Punkt-Methode im Detail besprochen werden.

12.1 Das Newton-Verfahren

Die betrachtete Innere-Punkt-Methode beruht auf dem Newton-Verfahren. Hier werden nur kurz einige Fakten zu diesem Verfahren zusammengestellt.

Bemerkung 12.4 Herleitung. Sei $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $\mathbf{f} \in C^1(\mathbb{R}^n)$ gegeben. Gesucht ist eine Nullstelle $\bar{\mathbf{x}}$ dieser Funktion $\mathbf{f}(\bar{\mathbf{x}}) = \mathbf{0}$. Sei \mathbf{x} eine Näherung an $\bar{\mathbf{x}}$. Setzt man $\bar{\mathbf{x}} = \mathbf{x} + \Delta\mathbf{x}$, dann erhält man mit der Taylorentwicklung, abgebrochen nach dem linearen Term, die Approximation

$$\mathbf{0} = \mathbf{f}(\bar{\mathbf{x}}) = \mathbf{f}(\mathbf{x} + \Delta\mathbf{x}) \approx \mathbf{f}(\mathbf{x}) + D\mathbf{f}(\mathbf{x})\Delta\mathbf{x},$$

wobei $D\mathbf{f}(\mathbf{x})$ die Jacobi-Matrix von $\mathbf{f}(\mathbf{x})$ in \mathbf{x} ist. Unter der Annahme, dass die Jacobi-Matrix regulär ist, erhält man

$$\Delta\mathbf{x} \approx -(D\mathbf{f}(\mathbf{x}))^{-1} \mathbf{f}(\mathbf{x}) \implies \bar{\mathbf{x}} \approx \mathbf{x} - (D\mathbf{f}(\mathbf{x}))^{-1} \mathbf{f}(\mathbf{x}).$$

Diese Beziehung motiviert die Berechnung der nächsten Iterierten \mathbf{x}^+ des Newton-Verfahrens

$$\mathbf{x}^+ := \mathbf{x} - (D\mathbf{f}(\mathbf{x}))^{-1} \mathbf{f}(\mathbf{x}).$$

\square

Satz 12.5 Konvergenzverhalten des Newton-Verfahrens. Seien $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $\mathbf{f} \in C^3(\mathbb{R}^n)$, $\bar{\mathbf{x}}$ eine Nullstelle von $\mathbf{f}(\mathbf{x})$ und $|D\mathbf{f}(\bar{\mathbf{x}})| \neq 0$ (Determinante). Dann gibt es ein $\varepsilon > 0$, so dass das Newton-Verfahren für alle Startwerte $\mathbf{x}^{(0)}$ mit $\|\bar{\mathbf{x}} - \mathbf{x}^{(0)}\|_2 \leq \varepsilon$ quadratisch gegen $\bar{\mathbf{x}}$ konvergiert.

Bemerkung 12.6 Interpretation. Der Satz besagt zum einen, dass das Newton-Verfahren lokal konvergent ist, das heißt, es konvergiert, wenn man nahe genug an der Lösung beginnt. Quadratische Konvergenz bedeutet, dass es Konstanten $c > 0$ und $k_0 > 0$ gibt, so dass

$$\|\bar{\mathbf{x}} - \mathbf{x}^{(k+1)}\|_2 \leq c \|\bar{\mathbf{x}} - \mathbf{x}^{(k)}\|_2^2$$

für alle $k \geq k_0$. \square

12.2 Ein Kurz-Schritt-Algorithmus

Bemerkung 12.7 Überführung des Optimierungsproblems in ein äquivalentes nichtlineares Gleichungssystem. Wir betrachten das lineare Programm

$$\min_{\mathbf{x} \in \mathbb{R}^n} \{z = \mathbf{c}^T \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\} \quad (12.1)$$

mit dem zugehörigen dualen Programm

$$\max_{\mathbf{y} \in \mathbb{R}^m} \{\tilde{z} = \mathbf{b}^T \mathbf{y} : A^T \mathbf{y} + \mathbf{s} = \mathbf{c}, \mathbf{s} \geq \mathbf{0}\}. \quad (12.2)$$

Im dualen Programm haben wir hierbei die Schlupfvariablen \mathbf{s} eingeführt, um Gleichungsnebenbedingungen zu erhalten.

Innere-Punkte-Verfahren erzeugen eine Folgen von Punkten $\{(\mathbf{x}^{(k)}, \mathbf{y}^{(k)}, \mathbf{s}^{(k)})\}$, $k = 0, 1, \dots$, mit $\mathbf{x}^{(k)} \geq \mathbf{0}$, $\mathbf{s}^{(k)} \geq \mathbf{0}$, deren Grenzwerte Optimallösungen von (12.1) und (12.2) liefern. Die Verfahren nennt man zulässige-Innere-Punkte-Verfahren, wenn alle Punkte $(\mathbf{x}^{(k)}, \mathbf{y}^{(k)}, \mathbf{s}^{(k)})$ zulässige innere Punkte von (12.1) und (12.2) sind, das heißt es gilt

$$A\mathbf{x}^{(k)} = \mathbf{b}, \mathbf{x}^{(k)} > \mathbf{0}, \quad A^T \mathbf{y}^{(k)} + \mathbf{s}^{(k)} = \mathbf{c}, \mathbf{s}^{(k)} > \mathbf{0}.$$

Um zulässige-Innere-Punkte-Verfahren verwenden zu können, müssen wir natürlich voraussetzen, dass die obigen Mengen nicht leer sind. Neben diesen Verfahren gibt es auch unzulässige-Innere-Punkte-Verfahren.

Aus der Nichtnegativität von \mathbf{x} und \mathbf{s} folgt

$$0 \leq \mathbf{x}^T \mathbf{s} = \mathbf{x}^T (\mathbf{c} - A^T \mathbf{y}) = \mathbf{c}^T \mathbf{x} - (A\mathbf{x})^T \mathbf{y} = \mathbf{c}^T \mathbf{x} - \mathbf{b}^T \mathbf{y} = z - \tilde{z}.$$

Von Satz 9.5 (Starker Dualitätssatz) wissen wir, dass für die Optimallösungen von (12.1) und (12.2) gilt $z = \tilde{z}$. Also folgt im Optimum $\mathbf{x}^T \mathbf{s} = 0$ und wegen Nichtnegativität dieser beiden Vektoren sogar $x_i s_i = 0$ für alle $i = 1, \dots, n$. Zusammen mit den Nebenbedingungen von (12.1) und (12.2) sind die Optimallösungen von (12.1) und (12.2) Lösungen des nichtlinearen Systems

$$\Psi_0(\mathbf{x}, \mathbf{y}, \mathbf{s}) := \begin{pmatrix} A\mathbf{x} - \mathbf{b} \\ A^T \mathbf{y} + \mathbf{s} - \mathbf{c} \\ X\mathbf{s} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{x}, \mathbf{s} \geq \mathbf{0}, \quad (12.3)$$

wobei $X := \text{diag}(\mathbf{x}) \in \mathbb{R}^{n \times n}$ ist. Die Nichtlinearität ist in der letzten Gleichung, welche ausgeschrieben $x_i s_i = 0$, $i = 1, \dots, n$, bedeutet. Die Jacobi-Matrix von Ψ_0 ist gegeben durch

$$D\Psi_0(\mathbf{x}, \mathbf{y}, \mathbf{s}) = \begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I_n \\ S & 0 & X \end{pmatrix} \in \mathbb{R}^{(2n+m) \times (2n+m)} \quad (12.4)$$

mit $S := \text{diag}(\mathbf{s}) \in \mathbb{R}^{n \times n}$ und I_n der n -dimensionalen Einheitsmatrix. □

Lemma 12.8 Regularität der Jacobi-Matrix. *Unter den Voraussetzungen $\text{rg}(A) = m$ und $\mathbf{x} > \mathbf{0}$, $\mathbf{s} > \mathbf{0}$ ist die Jacobi-Matrix (12.4) regulär.*

Beweis: Indirekt. Sei $(\mathbf{u}^T, \mathbf{v}^T, \mathbf{w}^T)^T \neq \mathbf{0}$ ein Vektor mit

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I_n \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

Aus der ersten Gleichung folgt $\mathbf{A}\mathbf{u} = \mathbf{0}$. Diese Beziehung wird in die umgestellte zweite Gleichung eingesetzt

$$\mathbf{w} = -A^T \mathbf{v} \implies \mathbf{u}^T \mathbf{w} = -\mathbf{u}^T A^T \mathbf{v} = -(\mathbf{A}\mathbf{u})^T \mathbf{v} = 0.$$

Aus der dritten Gleichung folgt $\mathbf{u} = -S^{-1}X\mathbf{w}$. Die Invertierbarkeit von S folgt aus $\mathbf{s} > \mathbf{0}$. Mit der eben bewiesenen Beziehung und der Symmetrie der Diagonalmatrizen ergibt sich

$$0 = \mathbf{u}^T \mathbf{w} = -\mathbf{w}^T X S^{-1} \mathbf{w}.$$

Da $\mathbf{x}, \mathbf{s} > \mathbf{0}$, folgt daraus $\mathbf{w} = \mathbf{0}$. Aus der dritten Gleichung folgt damit $\mathbf{u} = \mathbf{0}$. Damit vereinfacht sich die zweite Gleichung zu $A^T \mathbf{v} = \mathbf{0}$. Wegen des vollen Spaltenrangs von A^T folgt daraus $\mathbf{v} = \mathbf{0}$. Damit ist der Widerspruch zur Annahme konstruiert. ■

Bemerkung 12.9 Das Verfahren. Auf Grund dieses Satzes liegt es nahe, die Lösung des Systems (12.3) mit dem Newton-Verfahren zu versuchen. Dabei gibt es allerdings einige Schwierigkeiten:

- Die Regularität der Jacobi-Matrix ist nur für innere Punkte ($\mathbf{x}, \mathbf{s} > \mathbf{0}$) bewiesen. Die gesuchte Lösung liegt aber nicht im Inneren, da aus der dritten Gleichung $x_i s_i = 0$, $i = 1, \dots, n$, folgt, dass mindestens eine der Variablen x_i oder s_i im Optimum verschwindet.
- Es ist nicht klar, ob die in einem Newton-Schritt berechnete neue Iterierte überhaupt ein zulässiger Punkt ist. Falls nicht, ist die Regularität der Jacobi-Matrix nicht gesichert. Außerdem kann es passieren, dass das Newton-Verfahren gar nicht konvergiert oder zu einer Lösung konvergiert, welche die Nichtnegativitätsbedingung nicht erfüllt.

Um doch noch ein Newton-ähnliches Verfahren für (12.3) nutzen zu können, modifiziert man dieses System. Seien $\mathbf{e}_n = (1, \dots, 1)^T \in \mathbb{R}^n$ und sei $\mu > 0$ ein Parameter. Anstelle von (12.3) betrachtet man nun

$$\Psi_\mu(\mathbf{x}, \mathbf{y}, \mathbf{s}) := \begin{pmatrix} \mathbf{A}\mathbf{x} - \mathbf{b} \\ A^T \mathbf{y} + \mathbf{s} - \mathbf{c} \\ X\mathbf{s} - \mu \mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{x}, \mathbf{s} > \mathbf{0}. \quad (12.5)$$

Die Jacobi-Matrix dieses Systems ist (12.4).

Das schwerere Problem, eine Nullstelle von (12.3) zu finden, ersetzt man durch das einfachere Problem (12.5). Für festes $\mu > 0$ kann die Lösung nur im Inneren des zulässigen Bereichs liegen, da $x_i s_i = \mu$ für alle i . Da die Jacobi-Matrix nach Lemma 12.8 dort regulär ist, gibt es eine Umgebung der Lösung $(\mathbf{x}(\mu)^T, \mathbf{y}(\mu)^T, \mathbf{s}(\mu)^T)^T$, in der das Newton-Verfahren ohne Berücksichtigung der Ungleichungen $\mathbf{x}, \mathbf{s} > \mathbf{0}$ quadratisch konvergiert. In diesem Sinne hat man das Problem (12.3) mit den Nichtnegativitätsbedingungen durch ein System ohne Nebenbedingungen ersetzt.

Da das Optimum von (12.1) auf dem Rand angenommen wird ist zu erwarten, dass die Lösung $(\mathbf{x}(\mu)^T, \mathbf{y}(\mu)^T, \mathbf{s}(\mu)^T)^T$ umso weiter vom Rand entfernt liegt, je größer μ ist. Das Vorgehen ist:

- Man beginnt mit einem großen Wert μ_0 , für den sich die Lösung leicht berechnen lässt. Für große μ_0 erwartet man ein großes Konvergenzgebiet.
- Dann wird μ sukzessive verkleinert und jedesmal wird die zugehörige Lösung bis auf eine gewisse Genauigkeit mit dem Newton-Verfahren approximiert.

Die Punktmenge

$$\{(\mathbf{x}(\mu)^T, \mathbf{y}(\mu)^T, \mathbf{s}(\mu)^T)^T : \mu > 0\}$$

wird zentraler Pfad genannt. Für seine Punkte existiert eine sogenannte Dualitätslücke:

$$\begin{aligned} 0 < \mu n &= \mathbf{x}(\mu)^T \mathbf{s}(\mu) = \mathbf{x}(\mu)^T (\mathbf{c} - A^T \mathbf{y}(\mu)) = \mathbf{c}^T \mathbf{x}(\mu) - (\mathbf{A}\mathbf{x}(\mu))^T \mathbf{y}(\mu) \\ &= \mathbf{c}^T \mathbf{x}(\mu) - \mathbf{b}^T \mathbf{y}(\mu) = z_\mu - \tilde{z}_\mu. \end{aligned} \quad (12.6)$$

□

Lemma 12.10 Lösbarkeit des modifizierten Systems. Sei die dual zulässige Menge $\{\mathbf{y} : A^T \mathbf{y} \leq \mathbf{c}\}$ beschränkt und sei das Innere dieser Menge nichtleer. Dann besitzt das Problem (12.5) für jedes $\mu > 0$ eine eindeutige Lösung.

Beweis: Literatur, zum Beispiel [JS04, S 75ff]. ■

Bemerkung 12.11 Analyse eines Newton-Schrittes. Sei $(\mathbf{x}^T, \mathbf{y}^T, \mathbf{s}^T)^T$ die gegenwärtige Iterierte zur Lösung von (12.5) für festes $\mu > 0$. Die Korrektur für die nächste Iterierte des Newton-Verfahrens berechnet sich aus der Lösung von

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I_n \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x} \\ \Delta \mathbf{y} \\ \Delta \mathbf{s} \end{pmatrix} = - \begin{pmatrix} A\mathbf{x} - \mathbf{b} \\ A^T \mathbf{y} + \mathbf{s} - \mathbf{c} \\ X\mathbf{s} - \mu \mathbf{e} \end{pmatrix}. \quad (12.7)$$

Nach der Lösung dieses Systems folgt, dass die neue Iterierte $((\mathbf{x} + \Delta \mathbf{x})^T, (\mathbf{y} + \Delta \mathbf{y})^T, (\mathbf{s} + \Delta \mathbf{s})^T)^T$ die ersten zwei Gleichungen von (12.5) erfüllt, da diese linear sind. *Übungsaufgabe* Wir können annehmen, dass die gegenwärtige Iterierte bereits durch einen Newton-Schritt berechnet wurde, sie damit ebenfalls diese Eigenschaft besitzt und sich (12.7) vereinfacht zu

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I_n \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x} \\ \Delta \mathbf{y} \\ \Delta \mathbf{s} \end{pmatrix} = - \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ X\mathbf{s} - \mu \mathbf{e} \end{pmatrix} =: \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ -\mathbf{r} \end{pmatrix}. \quad (12.8)$$

Der Vektor \mathbf{r} wird Residuum genannt. Das Residuum der neuen Newton-Iterierten ist, unter Verwendung der Beziehungen von (12.8),

$$\begin{aligned} \tilde{\mathbf{r}} &= (X + \Delta X)(\mathbf{s} + \Delta \mathbf{s}) - \mu \mathbf{e} = X\mathbf{s} + \underbrace{X\Delta \mathbf{s} + S\Delta \mathbf{x}}_{=-\mathbf{r}} + \Delta X\Delta \mathbf{s} - \mu \mathbf{e} \\ &= X\mathbf{s} - \mathbf{r} + \Delta X\Delta \mathbf{s} - \mu \mathbf{e} = \Delta X\Delta \mathbf{s}. \end{aligned} \quad (12.9)$$

Um den Newton-Schritt analysieren zu können, werden einige Bezeichnungen eingeführt

$$D := \sqrt{XS^{-1}}, \quad \mathbf{q} := DX^{-1}\mathbf{r}, \quad R := \text{diag}(\mathbf{r}).$$

Direktes Nachrechnen oder Einsetzen zeigt, dass man die Lösung von (12.8) wie folgt aufschreiben kann: *Übungsaufgabe*

$$\begin{aligned} \Delta \mathbf{y} &= (AD^2A^T)^{-1}AD\mathbf{q}, \\ \Delta \mathbf{x} &= D^2A^T\Delta \mathbf{y} - D\mathbf{q}, \\ \Delta \mathbf{s} &= -D^{-1}\mathbf{q} - D^{-2}\Delta \mathbf{x}. \end{aligned}$$

Die Matrix

$$DA^T(AD^2A^T)^{-1}AD$$

definiert die Orthogonalprojektion auf den Bildraum von DA^T . Diese Projektion wird Π_R genannt. Demzufolge ist die Projektion in den Nullraum von DA^T gegeben durch $\Pi_N := I - \Pi_R$, wobei I die identische Abbildung ist. Nun kann man zeigen (*Übungsaufgabe*), dass

$$\Delta \mathbf{x} = -D\Pi_N\mathbf{q}, \quad \Delta \mathbf{s} = -D^{-1}\Pi_R\mathbf{q}$$

gelten. Das sind die beiden Ausdrücke, die wir in (12.9) analysieren müssen.

Wir nehmen nun an, dass man für gegebene $\mathbf{x} > \mathbf{0}$, $\mathbf{s} > \mathbf{0}$ und \mathbf{y} ein $\beta \in [0, 1/2]$ finden kann, für welches

$$\|\mathbf{r}\|_2 \leq \beta\mu \quad (12.10)$$

gilt. Das heißt, der gegebene Punkt ist in einer gewissen Umgebung des zentralen Pfades. Für $\beta = 0$ muss er direkt auf dem zentralen Pfad sein.

Wegen $\mathbf{r} = X\mathbf{s} - \mu\mathbf{e}$ gilt

$$DX^{-1} = \sqrt{XS^{-1}X^{-2}} = \sqrt{X^{-1}S^{-1}} = \sqrt{(R + \mu I)^{-1}} = \left(\sqrt{R + \mu I}\right)^{-1}.$$

Bei dieser Rechnung und auch bei folgenden Rechnungen wird stark ausgenutzt, dass hier die Matrizen D , R , S und X Diagonalmatrizen sind. Für Nichtdiagonalmatrizen funktionieren die Rechnungen nicht mehr. Es folgt

$$\|DX^{-1}\|_2^2 = \|(R + \mu I)^{-1}\|_2 = \max_{1 \leq i \leq n} \frac{1}{|r_i + \mu|} \leq \frac{1}{\mu(1 - \beta)}. \quad (12.11)$$

Hierbei wurde die Definition der Spektralnorm einer Matrix ausgenutzt, sowie dass wir es mit einer Diagonalmatrix zu tun haben. Die letzte Ungleichung folgt aus der Dreiecksungleichung und (12.10)

$$|r_i + \mu| \geq \mu - |r_i| \geq \mu - \beta\mu = \mu(1 - \beta).$$

Aus der Verträglichkeit der Euklidischen Vektornorm und der Spektralnorm von Matrizen sowie (12.10) folgt damit

$$\tilde{\beta} := \|\mathbf{q}\|_2 = \|DX^{-1}\mathbf{r}\|_2 \leq \|DX^{-1}\|_2 \|\mathbf{r}\|_2 \leq \beta \sqrt{\frac{\mu}{1 - \beta}}.$$

Sei θ der Winkel zwischen \mathbf{q} und $\Pi_N\mathbf{q}$. Dann folgt mit der Orthogonalität der Projektionen

$$\cos(\theta) = \frac{(\mathbf{q}, \Pi_N\mathbf{q})}{\|\mathbf{q}\|_2 \|\Pi_N\mathbf{q}\|_2} = \frac{(\Pi_R\mathbf{q} + \Pi_N\mathbf{q}, \Pi_N\mathbf{q})}{\|\mathbf{q}\|_2 \|\Pi_N\mathbf{q}\|_2} = \frac{\|\Pi_N\mathbf{q}\|_2^2}{\|\mathbf{q}\|_2 \|\Pi_N\mathbf{q}\|_2} = \frac{\|\Pi_N\mathbf{q}\|_2}{\|\mathbf{q}\|_2}.$$

Es folgen

$$\begin{aligned} \|D^{-1}\Delta\mathbf{x}\|_2 &= \|\Pi_N\mathbf{q}\|_2 = \cos(\theta) \|\mathbf{q}\|_2 = \tilde{\beta} \cos(\theta), \\ \|D\Delta\mathbf{s}\|_2 &= \|\Pi_R\mathbf{q}\|_2 = \|(I - \Pi_N)\mathbf{q}\|_2 = \sin(\theta) \|\mathbf{q}\|_2 = \tilde{\beta} \sin(\theta). \end{aligned} \quad (12.12)$$

Diese Abschätzungen kann man nun in (12.9) einsetzen. Mit der Cauchy-Schwarz-Ungleichung und $\beta \leq 1/2$ erhält man

$$\begin{aligned} \|\tilde{\mathbf{r}}\|_2 &= \|\Delta X \Delta\mathbf{s}\|_2 \leq \|D^{-1}\Delta X\|_2 \|D\Delta\mathbf{s}\|_2 \leq \tilde{\beta}^2 \cos(\theta) \sin(\theta) \\ &= \frac{\tilde{\beta}^2}{2} \sin(2\theta) \leq \frac{\tilde{\beta}^2}{2} \leq \beta^2 \frac{\mu}{2(1 - \beta)} \leq \beta^2 \mu. \end{aligned} \quad (12.13)$$

Wir haben damit erhalten, dass unter der Voraussetzung (12.10) der relative Fehler $\|X\mathbf{s} - \mu\mathbf{e}\|_2 / \mu = \|\mathbf{r}\|_2 / \mu$ in jedem Schritt des Newton-Verfahrens quadriert wird.

Aus (12.12) folgt insbesondere $\|D^{-1}\Delta\mathbf{x}\|_2 \leq \tilde{\beta}$ und daraus, sowie mit (12.11), der Definition von $\tilde{\beta}$ und $\beta \leq 1/2$,

$$\begin{aligned} \|X^{-1}\Delta\mathbf{x}\|_2 &= \|X^{-1}DD^{-1}\Delta\mathbf{x}\|_2 \leq \|X^{-1}D\|_2 \|D^{-1}\Delta\mathbf{x}\|_2 \leq \frac{\tilde{\beta}}{\sqrt{\mu(1 - \beta)}} \\ &= \frac{\beta}{1 - \beta} \leq 1. \end{aligned}$$

Somit ist $(\Delta x_i)/x_i \leq 1$ woraus

$$x_i - |\Delta x_i| \geq x_i - x_i = 0, \quad i = 1, \dots, n$$

folgt. Eine analoge Abschätzung kann man für \mathbf{s} durchführen. Die Nichtnegativität der neuen Iterierten ist gesichert.

Es gilt sogar die Positivität. *Übungsaufgabe ab hier.* Mit den Abschätzungen (12.13) und (12.9) folgt

$$\|\tilde{\mathbf{r}}\|_2 < \mu \implies \|\Delta X \Delta \mathbf{s}\|_2 < \mu \implies |\Delta \mathbf{x}_i \Delta \mathbf{s}_i| < \mu$$

für alle $i = 1, \dots, n$. Aus (12.9) erhält man ebenso

$$\|\tilde{\mathbf{r}}\|_2 = \|(X + \Delta X)(\mathbf{s} + \Delta \mathbf{s}) - \mu \mathbf{e}\|_2 < \mu,$$

was man abgekürzt und quadriert wie folgt schreiben kann

$$\sum_{i=1}^n (a_i - \mu)^2 < \mu^2,$$

mit $a_i = (X + \Delta X)(\mathbf{s} + \Delta \mathbf{s})$. Wir wissen bereits, dass beide Faktoren von a_i nichtnegativ sind. Sei für einen Index, zum Beispiel $i = j$, einer der Faktoren Null. Dann folgt

$$\mu^2 + \sum_{i=1, i \neq j}^n (a_i - \mu)^2 < \mu^2.$$

Diese Aussage kann nicht gelten, da alle Summanden der Summe nichtnegativ sind. Also sind alle Faktoren der a_i , $i = 1, \dots, n$, positiv. Das heißt, es gelten $\mathbf{x} + \Delta \mathbf{x} > \mathbf{0}$, $\mathbf{s} + \Delta \mathbf{s} > \mathbf{0}$ und der Newton-Schritt liefert einen inneren Punkt. *Übungsaufgabe bis hier.* \square

Algorithmus 12.12 Kurz-Schritt-Algorithmus. Es seien $\mathbf{x}^{(0)} > \mathbf{0}$, $\mathbf{y}^{(0)}$, $\mathbf{s}^{(0)} > \mathbf{0}$ und $\mu^{(0)}$ so gegeben, dass

$$\begin{aligned} A\mathbf{x}^{(0)} &= \mathbf{b}, \\ A^T\mathbf{y}^{(0)} + \mathbf{s}^{(0)} &= \mathbf{c}, \\ X^{(0)}\mathbf{s}^{(0)} - \mu^{(0)}\mathbf{e} &= \mathbf{r}^{(0)}, \\ \frac{\|\mathbf{r}^{(0)}\|_2}{\mu^{(0)}} &\leq \frac{1}{2} \end{aligned} \tag{12.14}$$

gelten. Es sei des weiteren eine gewünschte Genauigkeit $\varepsilon > 0$ vorgegeben. Setze $k = 0$.

1. Führe einen Newton-Schritt aus, das heißt löse (12.8). Setze $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta \mathbf{x}$, $\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} + \Delta \mathbf{y}$ und $\mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \Delta \mathbf{s}$.
2. Verkleinere den Parameter mittels der Vorschrift

$$\mu^{(k+1)} = \mu^{(k)} \left(1 - \frac{1}{6\sqrt{n}} \right).$$

3. Setze $k := k + 1$.
4. Falls $\mu^{(k)} \leq \varepsilon/n$ dann beende das Verfahren, ansonsten gehen zu Schritt 1. \square

Bemerkung 12.13 Die Bedingung (12.14) bedeutet, dass die Anfangsiterierte des Verfahrens sich relativ nahe am zentralen Pfad befinden muss. Die Bezeichnung „Kurz-Schritt-Algorithmus“ hat sich eingebürgert, weil der Parameter μ in jedem Schritt nur etwa um den Faktor $1 - 1/\sqrt{n}$ verkleinert wird, der für große n nahe bei Eins liegt. \square

Wir wollen nun die oben erarbeitete abstrakte Theorie auf den Kurz-Schritt-Algorithmus anwenden.

Lemma 12.14 *Der Kurz-Schritt-Algorithmus 12.12 erfüllt die Bedingung (12.10) mit $\beta = 1/2$.*

Beweis: Der Beweis erfolgt induktiv. Wegen der Forderung (12.14) an den Startpunkt, ist der Induktionsanfang gegeben. Gelte nun $\|\mathbf{r}^{(k)}\|_2 \leq \mu^{(k)}/2$ für ein $k \geq 0$. Aus (12.8), (12.9) (erste Gleichung) und der Definition des Parameters im Algorithmus 12.12 folgt

$$\mathbf{r}^{(k+1)} = X^{(k+1)} \mathbf{s}^{(k+1)} - \mu^{(k+1)} \mathbf{e} = \tilde{\mathbf{r}}^{(k)} + \mu^{(k)} \mathbf{e} - \mu^{(k+1)} \mathbf{e} = \tilde{\mathbf{r}}^{(k)} + \frac{\mu^{(k)}}{6\sqrt{n}} \mathbf{e}.$$

Aus (12.13) folgt $\|\tilde{\mathbf{r}}^{(k)}\|_2 \leq \mu^{(k)}/4$. Mit der Dreiecksungleichung, $\|\mathbf{e}\|_2 = \sqrt{n}$ und dem größtmöglichen Verhältnis von $\mu^{(k)}$ und $\mu^{(k+1)}$ (für $n = 1$) erhält man

$$\begin{aligned} \|\mathbf{r}^{(k+1)}\|_2 &\leq \|\tilde{\mathbf{r}}^{(k)}\|_2 + \frac{\mu^{(k)}}{6\sqrt{n}} \|\mathbf{e}\|_2 \leq \mu^{(k)} \left(\frac{1}{4} + \frac{1}{6} \right) = \frac{5}{12} \mu^{(k)} \leq \frac{5}{12} \frac{6}{5} \mu^{(k+1)} \\ &= \frac{1}{2} \mu^{(k+1)}. \end{aligned}$$

Das ist die Aussage des Lemmas. Die Induktionsvoraussetzung ist bei der Nutzung von (12.13) implizit verwendet wurden. ■

In einer Übungsaufgabe wurde gezeigt, dass die Iterierten bei der Ausführung des Newton-Schritts strikt zulässig bleiben ($\mathbf{x}, \mathbf{s} > \mathbf{0}$). Damit ist der Kurz-Schritt-Algorithmus 12.12 wohldefiniert.

Satz 12.15 Polynomiale Komplexität des Kurz-Schritt-Algorithmus. *Der Kurz-Schritt-Algorithmus 12.12 hält nach spätestens*

$$6\sqrt{n} \ln \left(\frac{n\mu^{(0)}}{\varepsilon} \right) \quad (12.15)$$

Iterationen mit Näherungslösungen $\mathbf{x} > \mathbf{0}$, \mathbf{y} , $\mathbf{s} > \mathbf{0}$ von (12.1) und (12.2), deren Dualitätslücke

$$\mathbf{c}^T \mathbf{x} - \mathbf{b}^T \mathbf{y} \leq 2\varepsilon$$

erfüllt.

Beweis: Man muss den Abbruchindex und den Abbruchfehler untersuchen. Die Abbruchbedingung in Schritt 4 ist erfüllt, wenn

$$\ln \mu^{(k)} \leq \ln \frac{\varepsilon}{n}.$$

Nach Konstruktion gilt

$$\mu^{(k)} = \left(1 - \frac{1}{6\sqrt{n}} \right)^k \mu^{(0)}.$$

Einsetzen liefert

$$k \ln \left(1 - \frac{1}{6\sqrt{n}} \right) + \ln \left(\mu^{(0)} \right) \leq \ln \frac{\varepsilon}{n}$$

oder

$$-k \ln \left(1 - \frac{1}{6\sqrt{n}} \right) \geq \ln \frac{n\mu^{(0)}}{\varepsilon} \iff k \ln \left(1 + \frac{1}{6\sqrt{n}-1} \right) \geq \ln \frac{n\mu^{(0)}}{\varepsilon}.$$

Wegen $\ln x \leq x - 1$ folgt daraus

$$k \left(1 - \frac{1}{6\sqrt{n}-1} - 1 \right) \geq \ln \frac{n\mu^{(0)}}{\varepsilon} \iff k \geq (6\sqrt{n}-1) \ln \frac{n\mu^{(0)}}{\varepsilon}.$$

Daraus folgt die erste Aussage des Satzes.

Die Abschätzung der Dualitätslücke erfolgt unter Nutzung von (12.6), der dritten Gleichung von (12.8), der Cauchy–Schwarz–Ungleichung, (12.10) und der Abbruchbedingung des Verfahrens

$$\begin{aligned} \mathbf{c}^T \mathbf{x} - \mathbf{b}^T \mathbf{y} &= \mathbf{x}^T \mathbf{s} = \mathbf{e}^T (X\mathbf{s}) = \mathbf{e}^T (\mu^{(k)} \mathbf{e} + \mathbf{r}) \leq n\mu^{(k)} + \|\mathbf{e}\|_2 \|\mathbf{r}\|_2 \\ &= n\mu^{(k)} + \sqrt{n}\beta\mu^{(k)} \leq 2n\mu^{(k)} \leq 2\varepsilon. \end{aligned}$$

■

Bemerkung 12.16 Man beachte, dass weder das Abbruchkriterium des Kurz–Schritt–Verfahrens noch die Dualitätslücke etwas darüber aussagen, wie groß der Abstand zwischen der mit dem Kurz–Schritt–Verfahren berechneten numerischen Lösung zur Lösung des linearen Programms (12.1) ist. □

Bemerkung 12.17 Zur Anzahl der Iterationen des Kurz–Schritt–Verfahrens. In vielen Anwendungen lässt sich ein Startpunkt zum Algorithmus 12.12 mit $n\mu^{(0)} \leq 10^{10}$ angeben. Die gewünschte Genauigkeit liegt im allgemeinen bei $\varepsilon = 10^{-8}$ bis $\varepsilon = 10^{-15}$. Für den letzten Wert ist der Faktor $6 \ln \left(\frac{n\mu^{(0)}}{\varepsilon} \right)$ in (12.15) ungefähr 345. Die Anzahl der Iterationen hängt im wesentlichen von \sqrt{n} ab. □

Bemerkung 12.18 Innere–Punkt–Verfahren in der Praxis. In der Praxis nutzt man vor allem unzulässige Innere–Punkt–Verfahren, das heißt, mit Iterierten, die außerhalb des zulässigen Bereichs liegen können. Die analytische Untersuchung dieser Verfahren ist jedoch komplizierter. □