Lineare Optimierung

bei Prof. Walter Alt

Semester: SS 2006 und WS 2009

Vorwort

Dieses Dokument wurde als Skript für die auf der Titelseite genannte Vorlesung erstellt und wird jetzt im Rahmen des Projekts „Vorlesungsskripte der Fakultät für Mathematik und Informatik“ weiter betreut. Das Dokument wurde nach bestem Wissen und Gewissen angefertigt. Dennoch garantiert weder der auf der Titelseite genannte Dozent, die Personen, die an dem Dokument mitgewirkt haben, noch die Mitglieder des Projekts für dessen Fehlerfreiheit. Für etwaige Fehler und dessen Folgen wird von keiner der genannten Personen eine Haftung übernommen. Es steht jeder Person frei, dieses Dokument zu lesen, zu verändern oder auf anderen Medien verfügbar zu machen, solange ein Verweis auf die Internetadresse des Projekts `http://uni-skripte.lug-jena.de/` *enthalten ist.*

Diese Ausgabe trägt die Versionsnummer 2680 und ist vom 20. Februar 2010. Eine (mögliche) aktuellere Ausgabe ist auf der Webseite des Projekts verfügbar.

Jeder ist dazu aufgerufen, Verbesserungen, Erweiterungen und Fehlerkorrekturen für das Skript einzureichen bzw. zu melden oder diese selbst einzupflegen – einfach eine E-Mail an die Mailingliste <`uni-skripte@lug-jena.de`> *senden. Weitere Informationen sind unter der oben genannten Internetadresse verfügbar.*

Hiermit möchten wir allen Personen, die an diesem Skript mitgewirkt haben, vielmals danken:

- *Jens Kubieziel* <`jens@kubieziel.de`> *(2006)*
- *Jörg Bader* <`joerg.bader@wurzel.org`> *(2009/2010)*

Inhaltsverzeichnis

Auflistung der Theoreme

Sätze

Definitionen

1. Einführung

1.1. Grundbegriffe und Beispiele

Die Zielstellung in der linearen Optimierung ist die Betrachtung einer Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ mit dem Ziel der Berechnung eines Minimums von f.

Beispiel 1.1
(1) $f\colon \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$ hat genau ein Minimum. (2) $f\colon \mathbb{R} \to \mathbb{R}$ mit $f(x) = x$ ist nicht nach unten beschränkt und hat daher kein Minimum. (3) $f\colon \mathbb{R} \to \mathbb{R}$ mit $f(x) = \sin x$ ist 2π-periodisch und hat unendlich viele Minima.

Beispiel 1.2
Die Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$ mit $f(x_1, x_2) = x_1 \sin(x_1 + x_1 x_2 \sin(x_2))$.

Die Norm auf $\mathbb{R}^n$ ist im allgemeinen die euklidische Norm, d. h. für $x \in \mathbb{R}^n$, $x = (x_1, \dots, x_n)^T$ ist $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$. Für $\tilde{x} \in \mathbb{R}$ und $r > 0$ ist die **offene Kugel** um $\tilde{x}$ mit Radius r definiert durch $B(\tilde{x}, r) = \{x \in \mathbb{R}^n \colon \|x - \tilde{x}\| < r\}$.

Allgemeine Formulierung des Optimierungsproblems Sei $D \subset \mathbb{R}^n$ nichtleer und offen, $f\colon D \to \mathbb{R}, \mathcal{F} \subset D$:

$$\min_{x \in \mathcal{F}} f(x) \tag{1.1}$$

Die Funktion f heißt **Ziel-** oder **Kostenfunktion**.

$\mathcal{F} = D$ unrestringiertes Optimierungsproblem (Optimierungsproblem ohne Nebenbedingungen)

$\mathcal{F} \subset D$ restringiertes Optimierungsproblem (Optimierungsproblem mit Nebenbedingungen), $\mathcal{F}$ heißt **zulässige Menge** und wird meist durch (Un-)gleichungen beschrieben.

Beispiel 1.3
$n = 1, D = \mathbb{R}, \min f(x) = x^3$, NB: $x \geq 1$.
In diesem Fall ist $\mathcal{F} = \{x \in \mathbb{R} \colon x \geq 1\}$

Definition 1.4 (Lokale Minima)
Ein Punkt $\tilde{x} \in \mathcal{F}$ (zulässiger Punkt) heißt **lokaler Minimalpunkt** von f auf $\mathcal{F}$ oder **lokale Lösung** von Gleichung 1.1, falls es ein $r > 0$ mit $f(x) \geq f(\tilde{x})$ für alle $x \in \mathcal{F} \cap B(\tilde{x}, r)$ gibt.

Ein Punkt $\tilde{x} \in \mathcal{F}$ heißt **strikter lokaler Minimalpunkt** von f auf $\mathcal{F}$ oder **strikte lokale Lösung** von Gleichung 1.1, falls es ein $r > 0$ mit $f(x) > f(\tilde{x})$ für alle $x \in \mathcal{F} \cap B(\tilde{x}, r)$ mit $x \neq \tilde{x}$ gibt.

Die zugehörigen Funktionswerte heißen (strikter) **lokaler Minimalwert** oder (striktes) **lokales Minimum**.

Definition 1.5 (Globales Minimum)
Ein Punkt $\tilde{x} \in \mathcal{F}$ heißt **globaler Minimalpunkt** von f auf $\mathcal{F}$ oder globale Lösung von Gleichung 1.1, falls gilt $\forall x \in \mathcal{F} : f(x) \geq f(\tilde{x})$

Bemerkung 1.6
Die Optimierungsaufgabe $\max g(x)$ unter den Nebenbedingungen $x \in \mathcal{F}$ ist äquivalent zur Minimierung. Das Problem wird als $\min f(x) := -g(x)$ definiert.

Anstatt lokales Minimum sagt man auch **relatives Minimum**. Anstatt striktes Minimum sagt man auch **strenges Minimum**.

Eine lineare Optimierungsaufgabe besteht aus einer linearen Funktion f und $\mathcal{F}$ wird aus linearen (Un-)gleichungen beschrieben.

Beispiel 1.7 (Produktionsplanung)
Eine Firma produziert vier Lacke L_1, L_2, L_3, L_4. Dabei ist der Gewinn pro Kilogramm Lack:

- 1,50 € für L_1
- 1,00 € für L_2
- 2,00 € für L_3
- 1,40 € für L_4

Nebenbedingungen für die wöchentliche Produktion:

- von L_1 und L_2 können zusammen maximal 1300 kg produziert werden
- von L_1, L_3, L_4 können zusammen maximal 2000 kg produziert werden
- Mindestproduktion von L_4 sind 800 kg
- von L_3 sind nicht mehr als 500 kg zu verkaufen

Ziel ist die Maximierung des Gewinns. Wir bezeichnen mit x_i die Menge, die von Lack L_i produziert wird. Der Gewinn ist $1,5x_1 + x_2 + 2x_3 + 1,4x_4$, d. h. es ist $f(x_1, x_2, x_3, x_4) = -1,5x_1 - x_2 - 2x_3 - 1,4x_4$ unter den Nebenbedingungen:

$$\begin{aligned} x_1 + x_2 &\leq 1300 \qquad & x_4 &\geq 800 \Leftrightarrow -x_4 \leq -800 \\ x_1 + x_3 + x_4 &\leq 2000 \qquad & x_3 &\leq 500 \end{aligned}$$

zu minimieren. Weiterhin ergibt sich aus praktischen Überlegungen $x_i \geq 0$. Wir formulieren das Problem in Matrixschreibweise: Sei $c = -(1, 5, 1, 2, 1, 4)^T$. Dann hat die Zielfunktion die Form $c^T x$. Mit

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} 1300 \\ 2000 \\ -800 \\ 500 \end{pmatrix}$$

können wir die Nebenbedingungen in der Form $Ax \leq b$ schreiben. Dann können wir das Optimierungsproblem in der Form $\min c^T x$ unter den Nebenbedingungen $Ax \leq b, x \geq 0$ schreiben.

Beispiel 1.8
Eine Tierfarm kauft drei verschiedene Kornsorten, um daraus Futtermittel zu mischen.

	S_1	S_2	S_3	Mindestbedarf
Nährstoff A	2	3	7	1250
Nährstoff B	1	1	0	250
Nährstoff C	5	3	0	900
Nährstoff D	0,6	0,25	1	232,5
Kosten	41	35	96	

Gesucht ist eine Mischung, die den Nährstoffbedarf deckt und dabei möglichst billig herzustellen ist.

Wir bezeichnen mit $x_i, i = 1, 2, 3$ diejenige Menge, die von der Kornsorte s_i eingekauft wird. Dabei müssen wir die Mengen so bestimmen, dass die Gesamtkosten der Futtermischung minimiert wird, d. h. die Funktion $41x_2 + 35x_2 + 96x_3$ soll minimiert werden. Folgende Nebenbedingungen gelten:

- Sinnvollerweise ist $x_i \geq 0, i = 1, 2, 3$
- Um den Nährstoffbedarf zu decken, müssen wir fordern:

$$2x_1 + 3x_2 + 7x_3 \geq 1250 \qquad x_1 + x_2 \geq 250$$
$$5x_1 + 3x_2 \geq 900 \qquad 0,6x_1 + 0,25x_2 + x_3 \geq 232,5$$

- Wir definieren weiter:

$$c = \begin{pmatrix} 41 \\ 35 \\ 96 \end{pmatrix} \qquad A = \begin{pmatrix} 2 & 3 & 7 \\ 1 & 1 & 0 \\ 5 & 3 & 0 \\ 0,6 & 0,25 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 1250 \\ 250 \\ 900 \\ 232,5 \end{pmatrix}$$

- Somit können wir das Problem in der Form $\min c^T x$ unter den Nebenbedingungen $Ax \geq b, x \geq 0$ schreiben.

Allgemein zur Produktplanung Es sollen bestimmte Güter aus Rohstoffen hergestellt werden. Durch die Zielfunktion soll der Gewinn maximiert bzw. die Kosten minimiert werden und als Nebenbedingung sollen keine negativen Mengen produziert bzw. benutzt werden.

1.2. Numerische Lösung von Optimierungsproblemen

Allgemein sei das Problem

$$\min f(x) \text{ mit } x \in \mathcal{F}$$

mit der Zielfunktion $f\colon \mathbb{R}^n \to \mathbb{R}$ und der zulässigen Menge $\mathcal{F} \subset \mathbb{R}^n$ zu lösen.

Numerische Verfahren sind Iterationsverfahren. Sie berechnen ausgehend von einem Startpunkt $x^{(0)}$ eine Folge $\{x^{(k)}\} \subset \mathbb{R}^n, k = 1, 2, 3, \ldots$ mit dem Ziel:

- nach endlich vielen Schritten m ist $x^{(m)}$ eine Lösung oder, falls dies nicht möglich ist,
- deren Grenzwert ist eine Lösung.

Diese numerische Verfahren berechnen nur eine Näherungslösung. Da die Zielfunktion zu minimieren ist, versucht man, die Folge $x^{(k)}$ so zu berechnen, dass $f(x^{(k+1)}) < f(x^{(k)})$ ist. Dies nennt man **Abstiegsverfahren**. Prinzipiell geht man dabei wie folgt vor:

- Man berechnet eine **Abstiegsrichtung** $d^{(k)}$, d. h. eine Richtung $d^{(k)}$ mit $f(x^{(k)} + td^{(k)}) < f(x^{(k)})$ für hinreichend kleines $t > 0$.
- Man berechnet eine geeignete Schrittweite t_k.

Ist $\mathcal{F} \neq \mathbb{R}^n$, dann fordert man in der Regel

- $x^{(0)} \in \mathcal{F}$
- $x^{(k)} \in \mathcal{F}, k = 1, 2, \ldots$; Dazu bestimmt man die **zulässige Abstiegsrichtung** $d^{(k)}$ mit der Eigenschaft, dass $x^{(k)} + td^{(k)} \in \mathcal{F}$ für hinreichend kleines $t > 0$.

Es gibt verschiedene Verfahren, die diese Vorgehensweise für bestimmte Klassen von Optimierungsproblemen realisieren:

- Lineare Optimierungsprobleme: Simplexverfahren, Innere-Punkte-Verfahren
- Differenzierbare Optimierungsprobleme: Gradientenverfahren

Es gibt auch allgemeine Verfahren, die keine speziellen Voraussetzungen an die Zielfunktion und die Nebenbedingungen machen.

Beispiel 1.9 (Mutations-Selektions-Verfahren)
1. Wähle $x^{(0)} \in \mathcal{F}$. Setze $k := 0$
2. Mutation: Berechne einen neuen Punkt $v^{(k)}$ durch zufällige Änderung $x^{(k)}$.

3. Selektion: Ist $v^{(k)} \in \mathcal{F}$ und $f(v^{(k)}) < f(x^{(k)})$, dann $x^{(k+1)} := v^{(k)}$. Andernfalls $x^{(k+1)} := x^{(k)}$.
4. Setze $k := k + 1$ und gehe zu Punkt 2.

Dieses Verfahren hat kein Abbruchkriterium. Eine häufig benutzte Mutation ist die Berechnung von $v_i^{(k)} = x_i^{(k)} + t_k(d_i^{(k)} - 0,5)$. Dabei ist $d_i^{(k)} \in [0,1]$ zufällig erzeugt und t_k muss durch Probieren bestimmt werden.

Beispiel 1.10
Sei folgendes Optimierungsproblem gestellt:

$$\begin{aligned} \min -(2000x_1 + 1200x_2 + 1000x_3 + 1500x_4) \\ x_1 + x_2 + x_3 + x_4 &\leq 16 \\ 1200x_2 &\geq 4000 \\ 1200x_2 - 1500x_4 &\leq 0 \\ 2000x_1 &\leq 7000 \\ xi &\geq 0 \qquad i = 1, \dots, 4 \end{aligned}$$

Das Simplexverfahren berechnet die Lösung in zwei Schritten. Zuerst wird $x^{(0)} \in \mathcal{F}$ ermittelt. Dazu benötigt das Verfahren vier Iterationen. Danach erfolgt die eigentliche Berechnung in zwei Iterationen und es stoppt mit $x = \begin{pmatrix} 3,5 \\ 3,33 \\ 0 \\ 9,16 \end{pmatrix}$, $f(x) = -24750$.

Das Mutations-Selektions-Verfahren wird mit einer Schrittweite $t_0 = 0,8$ gestartet. Diese wird immer nach 5000 Schritten halbiert und nach 100000 Iterationen wird abgebrochen. Das ermittelte Ergebnis ist $x = \begin{pmatrix} 3,5 \\ 3,333 \\ 1,6557 \cdot 10^{-7} \\ 9,1666 \end{pmatrix}$, $f(x) = -24749,999985$.

2. Grundlagen der linearen Optimierung

Ein lineares Optimierungsproblem wird im englischen auch als „linear program“ bezeichnet. Daher rührt die oft benutzte Abkürzung *LP* her. Das Problem besteht aus einer linearen Zielfunktion $f(x) = c^T x$ und Nebenbedingungen, die in Form von Gleichungen und Ungleichungen definiert werden.

Es gibt verschiedene „Standardformen“ von linearen Optimierungsproblemen.

2.1. Lineare Optimierungsprobleme

2.1.1. Standardformen von linearen Optimierungsproblemen

Definition 2.1 (Lineares Optimierungsproblem in allgemeiner Form)
Ein lineares Optimierungsproblem in allgemeiner Form (LPA) ist die Aufgabe:

$$\min c^T x \text{ mit } l_x \leq x \leq u_x, l_A \leq A^{(1)}x \leq u_A, A^{(2)}x = b$$

wobei gilt: $x, c, l_x, u_x \in \mathbb{R}^n$, $A^{(1)}$ ist eine $m_1 \times n$-Matrix, $l_A, u_A \in \mathbb{R}^{m_1}$, $A^{(2)}$ ist eine $m_2 \times n$-Matrix, $b \in \mathbb{R}^{m_2}$.

Beispiel 2.2
Eine Supermarktkette stellt ein Billiggetränk auf Weinbasis her. Grundlage ist ein Landwein mit Kosten von 1 € pro Liter. Weitere Zutaten sind Zuckerlösung (1,20 € pro Liter) und Konservierungsmittel (1,80 € pro Liter). Das Ziel ist die Herstellung einer möglichst billigen Mischung. Dabei müssen folgende Restriktionen beachtet werden.

- mindestens ein Drittel Zuckerlösung

- mindestens halb soviel Wein wie Zuckerlösung

- Anteil des Konservierungsmittels mindestens halb so groß wie der Anteil der Zuckerlösung und höchstens so groß wie der Anteil der Zuckerlösung

- Anteil des Konservierungsmittels mindestens die Hälfte des Weinanteils

Wir legen folgende Optimierungsvariablen fest:

x_1 Anteil der Zuckerlösung

x_2 Anteil des Konservierungsmittels

x_3 Anteil des Weins

Die Kosten $1,2x_1 + 1,8x_2 + x_3$ sind unter folgenden Restriktionen zu minimieren.

$$
\begin{array}{ccc}
x_1, x_2, x_3 \geq 0 & x_1 + x_2 + x_3 = 1 & x_1 \geq \frac{1}{3} \\
\frac{1}{2}x_1 \leq x_2 \leq x_1 & 2x_3 \geq x_1 & x_2 \geq \frac{1}{2}x_3
\end{array}
$$

Bemerkung 2.3

- Ungleichungen zwischen Vektoren sind komponentenweise zu verstehen.
- Sinnvollerweise gilt: $l_x \leq u_x, l_A \leq u_A$. Gilt für einen Index $i \in \{1, \ldots, n\}$: $l_{x_i} = u_{x_i}$, d. h. wir fordern $l_{x_i} = x_i = u_{x_i}$, dann ist x_i fest.
- Das lineare Funktional $\mathcal{F} = \{x \in \mathbb{R}^n : l_x \leq x \leq u_x, l_A \leq A^{(1)} \leq u_A, A^{(2)}x = b\}$ heißt **zulässige Menge** (feasible set).
- Ein Punkt $\tilde{x} \in \mathcal{F}$ (zulässiger Punkt) heißt **optimal**, falls $\forall x \in \mathcal{F} : f(x) \geq f(\tilde{x})$ gilt, d. h. $\tilde{x}$ ist globales Minimum von f auf $\mathcal{F}$.

Mögliche Transformationen

- Sei $a \in \mathbb{R}^n, b \in \mathbb{R}$. Eine Ungleichung vom Typ $a^T x \leq -b$ ist äquivalent zu $-a^T x \geq b$.
- Eine Gleichung $a^T x = b$ ist äquivalent zu $-a^T x = -b$ und äquivalent zu $a^T x \leq b$ und $a^T x \geq b$.
- Gibt es für x_i keine Vorzeichenbedingung, dann kann man zwei neue Variablen x_i^+, x_i^- mit $x_i = x_i^+ - x_i^-$ mit $x_i^+, x_i^- \geq 0$ einführen.

Beispiel 2.4
Wir betrachten das Problem, $\min x_1 - x_2$ unter der Nebenbedingungen $x_1 \geq 0, x_2 \leq 5$. Die Lösung ist mit $x_1 = 0, x_2 = 5$ eindeutig bestimmt. Um ein Problem mit Vorzeichenbedingung für *alle* Variablen zu erhalten, definieren wir x_2^+, x_2^- und erhalten das Problem $\min x_1 - x_2^+ + x_2^-$ mit den Nebenbedingungen $x_1, x_2^+, x_2^- \geq 0, x_2^+ - x_2^- \leq 5$.

Definition 2.5 (Lineares Optimierungsproblem in kanonischer Form)
Ein lineares Optimierungsproblem in kanonischer Form (LPK) ist die Aufgabe:

$$\min c^T x \text{ mit } Ax \geq b, x \geq 0$$

mit $c \in \mathbb{R}^n, A$ ist eine $m \times n$-Matrix, $b \in \mathbb{R}^m$. In diesem Fall ist die zulässige Menge $\mathcal{F} = \{x \in \mathbb{R}^n : Ax \geq b, x \geq 0\}$.

Beispiel 2.6
siehe Satz 1.7 (Produktionsplanung)

$$\max 1,5x_1 + x_2 + 2x_3 + 1,4x_4$$

$$x_1 + x_2 + \leq 1300 \qquad x_1 + x_3 + x_4 \leq 2000 \qquad x_4 \leq 800 \qquad x_3 \leq 500$$

$$c = -\begin{pmatrix} 1,5 \\ 1 \\ 2 \\ 1,4 \end{pmatrix} \qquad A = \begin{pmatrix} -1 & -1 & 0 & 0 \\ -1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} -1300 \\ -2000 \\ 800 \\ -500 \end{pmatrix}$$

Definition 2.7 (Lineares Optimierungsproblem in Normalform)
Ein lineares Optimierungsproblem in Normalform oder Standardform ist die Aufgabe

$$\min c^T x \text{ unter den Nebenbedingungen} Ax = b, x \geq 0$$

mit $c \in \mathbb{R}^n, b \in \mathbb{R}^m$ und A ist eine $m \times n$-Matrix. Die zulässige Menge ist

$$\mathcal{F} = \{ x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \}$$

Bemerkung 2.8

- Wir können o. B. d. A. $b \geq 0$ setzen.
- Bei n Variablen und m Gleichungsrestriktionen folgt, dass ein lineares Gleichungssystem vorliegt.
- In der Regel fordert man $m < n$.

Wie kann man nun ein Problem in kanonischer Form in Standardform umwandeln? Die Lösung ist hier die Einführung von **Schlupfvariablen**.

Beispiel 2.9
Sei folgende Aufgabe gegeben:
Maximiere $x_1 + x_2$ unter den Nebenbedingungen $x_1 + 3x_2 \leq 13, 3x_1 + x_2 \leq 15, -x_1 + x_2 \leq 3, x_i \geq 0$.
Wir führen die **Schlupfvariablen** x_3, x_4, x_5 ein und ersetzen die Ungleichung durch:

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 13 & \qquad 3x_1 + x_2 + x_4 &= 15 \\ -x_1 + x_2 + x_5 &= 3 & x_i &\geq 0 \end{aligned}$$

Allgemein lässt sich der Sachverhalt so formulieren: Sei ein lineares Problem in kanonischer Form gegeben. Um das Problem in Standardform zu transformieren, führen wir die Schlupfvariablen

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} x_{n+1} \\ \vdots \\ x_{n+m} \end{pmatrix} \in \mathbb{R}^m$$

ein und erhalten somit das Problem

$$\min_{x\in\mathbb{R}^n, y\in\mathbb{R}^m} c^T x = (c^T, 0_m^T)\begin{pmatrix} x \\ y \end{pmatrix}$$

unter den Nebenbedingungen $x, y \geq 0$, d. h.

$$Ax - y = b = (A - I_m)\begin{pmatrix} x \\ y \end{pmatrix}$$

Beispiel 2.10 (Frannie verkauft Holz)
Frannie verkauft jedes Jahr drei Festmeter[1] Holz. Zwei Kunden machen folgendes Angebot. Kunde1 bietet 90 € für einen halben Festmeter und Kunde2 bietet 150 € für einen Festmeter. Seien also x_1 die Anzahl der halben Festmeter an Kunde1 und x_2 die Anzahl der Festmeter an Kunde2. Die Einnahmen sollen maximiert werden. Damit ergibt sich folgendes Optimierungsproblem:

$$\min -90x_1 - 150x_2$$

unter den Nebenbedingungen:

$$0,5x_1 + x_2 \leq 3 \quad x_i \geq 0$$

Wir betrachten die Geraden $x_2 = -0,5x_1 + 3$ und $-90x_1 - 150x_2 = r$

2.1.2. Grafische Lösung von linearen Optimierungsproblemen

Grafik einfügen

Hierzu muss man sich auf $n = 2$ beschränken und es sind nur Ungleichungen als Nebenbedingungen zugelassen

Zur Bestimmung der Lösungsmenge werden die Randgeraden eingezeichnet und die richtige Lösungshälfte bestimmt.

2.1.3. Lösbarkeit von linearen Optimierungsproblemen

Hier kann man drei Fälle unterscheiden:

1. Lösung ist eindeutig bestimmt
2. Es gibt unendlich viele Lösungen.
3. Es gibt keine Lösung.

[1] http://de.wikipedia.org/wiki/Festmeter

2.1.4. Software zur Lösung

Zur Lösung von linearen Optimierungsproblemen kann man u. a. MATLAB oder auch Scilab benutzen.

MATLAB ist eine kommerzielle Lösung der Firma The MathWorks, Inc. und hat die Webseite http://www.mathworks.de/.

Scilab ist eine kostenlose Entwicklung aus Frankreich. Sie wurde am INRIA entwickelt und ist über die Webseite http://www.scilab.org/ verfügbar.

Beispiel 2.11

Beispiel einfügen

Beispiel 2.12

Beispiel einfügen

Beispiel 2.13

Beispiel einfügen

Beispiel 2.14

Beispiel einfügen

2.2. Konvexe Mengen

Definition 2.15 (Konvexe Menge)
Ein Menge $C \subset \mathbb{R}^n$ heißt **konvex**, falls mit $x, y \in C$ und $\alpha \in (0, 1)$ auch der Punkt $(1-\alpha)x + \alpha y$ Element von C ist.

Bemerkung 2.16
Seien $x, y \in C$. Dann heißt $[x, y] := \{ (1-\alpha)x + \alpha y \mid \alpha \in [0, 1] \}$ **Verbindungsstrecke** von x und y. Konvexität heißt $x, y \in C \Rightarrow [x, y] \in C$.

Beispiel 2.17
Sei $n = 1$. Dann sind die konvexen Mengen Intervalle. Für den Fall, dass n beliebig ist, gilt:

Sei $x \in \mathbb{R}^n$ und $r \in \mathbb{R}$. Wir betrachten $B(x, r) = \{ y \in \mathbb{R}^n \mid \|y - x\| < r \}$ und behaupten, dass $B(x, r)$ konvex ist. Also ist zu zeigen, dass $v = (1-\alpha)y + \alpha x \in B(x, r)$ bzw. $\|v - x\| < r$. Denn sei $y, z \in B(x, r)$ und $\alpha \in (0, 1)$. Dann folgt:

$$\begin{aligned}
\|v - x\| &= \|(1-\alpha)y + \alpha z - x\| = \|(1-\alpha)y + \alpha z - (1-\alpha)x - \alpha x\| \\
&= \|(1-\alpha)(y - x) + \alpha(z - x)\| \leq \|(1-\alpha)(y - x)\| + \|\alpha(z - x)\| \\
&\leq (1-\alpha) \underbrace{\|y - x\|}_{<r} + \alpha \underbrace{\|z - x\|}_{<r} \\
&< (1 - \alpha + \alpha)r = r
\end{aligned}$$

Beispiel 2.18
Sei $(s, r) \in \mathbb{R}^n \times \mathbb{R}$, $s \neq 0$, wobei 0 der Nullvektor ist. Die Hyperebene $H_{s,r} = \{ x \in \mathbb{R}^n \mid \langle s, x\rangle = r \}$ ist konvex. Denn sei $y, z \in H_{s,r}$ und $\alpha \in [0, 1]$ beliebig. Also ist zu zeigen, $v = (1-\alpha)y + \alpha z \in H_{s,r}$, d. h. $\langle s, v\rangle = r = \langle s, (1-\alpha)y + \alpha z\rangle = (1-\alpha) \underbrace{\langle s, y\rangle}_{=r} + \alpha \underbrace{\langle s, z\rangle}_{=r} = r$. Gleiches gilt für den offenen Halbraum $\{ x \in \mathbb{R}^n \mid \langle s, x\rangle < r \}$ und den abgeschlossenen Halbraum $\{ x \in \mathbb{R}^n \mid \langle s, x\rangle \leq r \}$.

Lemma 2.19
Ist $(C_j)_{j\in J}$ eine beliebige Familie konvexer Mengen, dann ist auch der Durchschnitt dieser Mengen $C := \bigcap_{j\in J} C_j$ konvex.

Denn seien $x, y \in C$ und $\alpha \in [0, 1]$ beliebig. Dann folgt für alle $j \in J$, dass $x, y \in C_j$ und weiter ist $(1-\alpha)x + \alpha y \in C_j$. Damit gilt aber $(1-\alpha)x + \alpha y \in C$.

Beispiel 2.20
Sei $(a^{(i)}, b^{(i)}) \in \mathbb{R}^n \times \mathbb{R}$, $i = 1, \dots, s-1, s, s+1, \dots, m$. Weiter sei

$$K = \{ x \in \mathbb{R}^n \mid \langle a^{(i)}, x\rangle = b^{(i)}, i = 1, \dots, s, \langle a^{(i)}, x\rangle \leq b^{(i)}, i = s+1, \dots, m \}$$

die Lösungsmenge eines Systems linearer Gleichungen und Ungleichungen. Man nennt K auch **Polyeder**. Es wird behauptet, dass K konvex ist.

Sei $C_i = \{ x \in \mathbb{R}^n \mid \langle a^{(i)}, x\rangle = b_i \}$ mit $i = 1, \dots, s$ bzw. $C_i = \{ x \in \mathbb{R}^n \mid \langle a^{(i)}, x\rangle \leq b_i \}$ mit $i = s+1, \dots, n$. Dann ist $K = \bigcap_{i=1}^m C_i$ konvex nach Satz 2.19 und obigem Beispiel.

Ein lineares Optimierungsproblem in allgemeiner, kanonischer oder Standardform ist vom Typ $\min c^T x$ unter den Nebenbedingungen $x \in K$, wobei K ein Polyeder ist.

Satz 2.21 (Lineare Optimierung auf konvexen Mengen)
Ist $c \in \mathbb{R}^n$ und $K \subset \mathbb{R}^n$ nichtleer und konvex, dann ist jede lokale Lösung des Problems $\min c^T x$ unter den Nebenbedingungen $x \in K$ auch globale Lösung. Weiterhin ist die Lösungsmenge konvex.

Beweis: 1
Sei $\tilde{x} \in K$ eine beliebige lokale Lösung. Dann existiert ein $r > 0$ mit $f(x) \geq f(\tilde{x})$ für alle $x \in K \cap B(\tilde{x}, r)$. Sei weiter $y \in K, y \neq \tilde{x}$ beliebig. Wir zeigen: $f(y) \geq f(\tilde{x})$. Dazu nutzen wir die Konvexität aus. Da K konvex ist, gilt $\tilde{x} + \alpha(y - \tilde{x}) = (1-\alpha)\tilde{x} + \alpha y \in K$. Weiter gilt für hinreichend kleines $\alpha > 0$ (speziell für $\alpha < \frac{r}{\|y-\tilde{x}\|}$), dass $\tilde{x} + \alpha(y-\tilde{x}) \in B(\tilde{x}, r)$ ist. Für $0 < \alpha < \min\{1, \frac{r}{\|y-\tilde{x}\|}\}$ gilt also: $\tilde{x} + \alpha(y-\tilde{x}) \in K \cap B(\tilde{x}, r)$. Somit folgt: $f(\tilde{x} + \alpha(y-\tilde{x})) \geq f(\tilde{x})$. Dies ist gleichbedeutend mit $c^T(\tilde{x} + \alpha(y-\tilde{x})) \geq c^T\tilde{x} \Rightarrow c^T\tilde{x} + \alpha c^T y - \alpha c^T\tilde{x} \geq c^T\tilde{x} \Rightarrow \alpha c^T y \geq \alpha c^T \tilde{x} \Rightarrow c^T y \geq c^T\tilde{x} \Rightarrow f(y) \geq f(\tilde{x})$.

Seien $\tilde{x}, \tilde{y}$ beliebige Lösungen. Dann sind diese auch globale Lösungen, d. h. $f(\tilde{x}) = f(\tilde{y})$. Sei $\alpha \in [0, 1]$. Nun ist zu zeigen, dass auch $(1-\alpha)\tilde{x} + \alpha\tilde{y}$ Lösung ist. Es gilt: $f((1-\alpha)\tilde{x} + \alpha\tilde{y}) = c^T((1-\alpha)\tilde{x} + \alpha\tilde{y}) = (1-\alpha)c^T\tilde{x} + \alpha c^T\tilde{y} = (1-\alpha)f(\tilde{x}) + \alpha f(\tilde{y}) = (1-\alpha)f(\tilde{x}) + \alpha f(\tilde{x}) = f(\tilde{x})$. ∎

Bemerkung 2.22
Die Aussage von Satz 2.21 gilt allgemein für konvexe Zielfunktionen. Dabei heißt eine Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ konvex auf K, wenn gilt: $f((1-\alpha)x + \alpha y) \leq (1-\alpha)f(x) + \alpha f(y)$ für alle $x, y \in K$ und $\alpha \in [0, 1]$. Im Spezialfall $f(x) = c^T x$ gilt sogar $f((1-\alpha)x + \alpha y) = (1-\alpha)x + \alpha y$.

Definition 2.23 (Konvexkombination)
Eine **Konvexkombination** von Punkten $x^{(1)}, \dots, x^{(k)} \in \mathbb{R}^n$ ist ein Punkt der Form $\sum_{i=1}^k \alpha_i x^{(i)}$ mit $\alpha_i \geq 0, i = 1, \dots, k$ und $\sum_{i=1}^k \alpha_i = 1$.

Bemerkung 2.24
Für $k = 2$ gilt: $\alpha_1 x^{(1)} + \alpha_2 x^{(2)}$ ist Konvexkombination mit $\alpha_1, \alpha_2 \geq 0, \alpha_1 + \alpha_2 = 1 \Leftrightarrow \alpha_1 = 1 - \alpha_2$.

Satz 2.25
Eine Menge $C \subset \mathbb{R}^n$ ist dann konvex, wenn sie alle Konvexkombinationen von Punkten in C enthält.

Beweis: 2
Dies stellt eine Beweisidee dar:

In der Gegenrichtung zeigt man: Wenn C alle Konvexkombinationen von Punkten in C enthält, dann enthält sie alle Konvexkombinationen von zwei Punkten in C. Dann folgt, dass C konvex ist.

In der Hinrichtung nimmt man an, dass C konvex ist. Weiter sei $x^{(1)}, \dots, x^k \in C$ beliebig und $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ mit $\alpha_i \geq 0, i = 1, \dots, k, \sum_{i=1}^k \alpha_i = 1$. Dann ist zu zeigen, dass $\sum_{i=1}^k \alpha_i x^{(i)} \in C$. Dieser Beweis wird per Induktion geführt. ∎

Definition 2.26 (Konvexe Hülle)
Sei $A \in \mathbb{R}^n$ beliebig. Die **konvexe Hülle** von A ist die Menge

$$\operatorname{co} A := \bigcap_{A \subset C} C$$

Dies ist die kleinste konvexe Menge, die A umfasst.

Lemma 2.27
Für eine beliebige Menge $A \subset \mathbb{R}^n$ ist $\operatorname{co} A = \{ x \in \mathbb{R} \mid x \text{ ist Konvexkombination von Punkten in } A \} =: B$.

Beweis: 3
Für die Richtung $\operatorname{co} A \subset B$ gilt, dass B konvex ist und das $A \subset B$. Somit folgt, $\operatorname{co} A \subset B$.

Für $B \subset \operatorname{co} A$ gilt: Sei C eine beliebige konvexe Menge mit $C \supset A \Rightarrow C$ enthält alle Konvexkombinationen von Punkten in A, d. h. $C \supset B \Rightarrow B \subset \bigcap_{C \supset A} C \subset \operatorname{co} A$. ∎

2.3. Hauptsatz der linearen Optimierung

Ziel dieses Abschnittes ist, zu zeigen, wenn ein lineares Optimierungsproblem eine Lösung hat, existiert eine Ecke, die Lösung ist.

Definition 2.28 (Geometrische Definition einer Ecke)
Sei $\mathcal{F}$ die zulässige Menge eines linearen Optimierungsproblems. Dann heißt ein Punkt $x \in \mathcal{F}$ **Ecke** von $\mathcal{F}$, falls x nicht als Konvexkombination $x = (1 - \alpha)y + \alpha z$ mit $y, z \in \mathcal{F}, y \neq z, \alpha \in (0, 1)$ darstellbar ist.

Bemerkung 2.29
Ist $\mathcal{F}$ die zulässige Menge eines linearen Optimierungsproblems mit Vorzeichenbedingung $x \geq 0_n$ und ist der Nullpunkt $0_n \in \mathcal{F}$, dann ist 0_n eine Ecke von $\mathcal{F}$. Denn für $y, z \in \mathcal{F}$ mit $y \neq z, \alpha \in (0,1)$ ist $0_n = (1-\alpha)y + \alpha z \Rightarrow y = z = 0_n$.

Beispiel 2.30
Sei $\mathcal{F} \subset \mathbb{R}^2$ die zulässige Menge eines linearen Optimierungsproblems in kanonischer Form. Es ist definiert durch: $2x_1 + 3x_2 \leq 6, x_1 \leq 2, x_i \geq 0$.

Beispiel einfügen

Beispiel 2.31

Ab jetzt betrachten wir nur noch lineare Optimierungsprobleme in Standardform, d. h. mit der zulässigen Menge $\mathcal{F} = \{ x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \}$. Dabei ist A eine $m \times n$-Matrix. Die zulässigen Punkte $x \in \mathcal{F}$ sind insbesondere Lösung des linearen Gleichungssystems $Ax = b$.

Es gibt folgende Spezialfälle:

$n = m$ Hat A den vollen Rang, dann hat das System $Ax = b$ genau eine Lösung: $x = A^{-1}b$. Gilt für x auch noch $x \geq 0_n \Rightarrow \mathcal{F} = \{x\}$. Andernfalls ist $\mathcal{F} = \emptyset$.

$m > n$ In diesem Falle gibt es mehr Gleichungen als Variablen (überbestimmtes System). Dies ist für die Optimierung uninteressant.

$m < n$ Dieses System hat weniger Gleichungen als Variablen. Die Menge $\{ x \in \mathbb{R}^n \mid Ax = b \}$ ist dann ein affiner Unterraum. Spezialfall: Der Rang der Matrix A ist gleich m.

Satz 2.32 (Charakterisierung einer Ecke)
Sei $\mathcal{F} = \{ x \in \mathbb{R}^n \mid Ax = b, x \geq 0_n \}, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \leq n$ die zulässige Menge eines linearen Optimierungsproblems in Standardform. Dann gilt: $x \in \mathcal{F}$ ist genau dann Ecke von $\mathcal{F}$, wenn die zu positiven Komponenten von x gehörenden Spaltenvektoren linear unabhängig sind, d. h. wenn die Vektoren $A_{.i}$ linear unabhängig sind.

Beweis: 4
Ist $x = 0_n$, dann ist x Ecke von $\mathcal{F}$ nach Satz 2.29. Sei jetzt $I(x) \neq \emptyset$ (x hat mindestens eine Komponente, die größer als 0 ist).

„$\Rightarrow$“ Sei x Ecke von $\mathcal{F}$. Dann ist zu zeigen, dass die Vektoren $A_{.i}$ mit $i \in I(x)$ linear unabhängig sind. Angenommen die Vektoren $A_{.i}$ seien linear abhängig. Dann zeigen wir: $\exists y, z \in \mathcal{F}$ mit $y \neq z$ und $x = 0,5y + 0,5z$. Da die Vektoren $A_{.i}$ linear abhängig ist, existieren Zahlen $\alpha_i \neq 0$ mit $0 = \sum_{i \in I(x)} \alpha_i A_{.i}$ und $\gamma := \max_{i \in I(x)} |\alpha_i| > 0$. Sei $\beta := \min_{i \in I(x)} x \Rightarrow \beta > 0$. Wir definieren zwei Vektoren:

$$y_i = \begin{cases} x_i - \frac{\beta}{\gamma}\alpha_i & i \in I(x) \\ 0 & \text{sonst} \end{cases} \qquad z_i = \begin{cases} x_i + \frac{\beta}{\gamma}\alpha_i & i \in I(x) \\ 0 & \text{sonst} \end{cases}$$

Dann ist $x = 0,5y + 0,5z$ und weiter $y \neq z$. Nun ist noch zu zeigen, dass $y, z \in \mathcal{F}$, d. h. $y, z \geq 0, Ay = b = Az$. Aus der Definition von β und γ erhalten wir:

$$\left|\frac{\beta}{\gamma}\alpha_i\right| = \beta \underbrace{\frac{|\alpha_i|}{\gamma}}_{\leq 1} \leq \beta \Rightarrow y, z \geq 0_n$$

Weiter ist $x_i = 0$ für $i \notin I(x)$. Damit folgt, $Ay = \sum_{i=1}^{n} y_i A_{.i} = \sum_{i \in I(x)} y_i A_{.i} = \sum_{i \in I(x)} \left(x_i - \frac{\beta}{\gamma}\alpha_i\right) A_{.i} = \sum_{i \in I(x)} x_i A_{.i} - \frac{\beta}{\gamma} \underbrace{\sum_{i \in I(x)} \alpha_i A_{.i}}_{0_n} = \sum_{i=1}^{n} x_i A_{.i} = Ax$ ↯

„$\Leftarrow$“ Seien die $A_{.i}$ linear unabhängig. Dann ist zu zeigen, x ist Ecke von $\mathcal{F}$. Sei $x = (1-\lambda)x^{(1)} + \lambda x^{(2)}$ mit $x^{(1)}, x^{(2)} \in \mathcal{F}, \lambda \in (0,1)$. Dann ist zu zeigen, dass $x^{(1)} = x^{(2)}$. Sei nun $j \notin I(x) \Rightarrow 0 = x_j = (1-\lambda)x_j^{(1)} + \lambda x_j^{(2)} \Rightarrow x_j^{(1)} = x_j^{(2)} = 0$. Es gilt, $Ax^{(1)} = b, Ax^{(2)} = b \Rightarrow 0 = b - b = Ax^{(1)} - Ax^{(2)} = A(x^{(1)} - x^{(2)})$ und weiter $0 = \sum_{i=1}^{n}(x_i^{(1)} - x_i^{(2)})A_{.i} = \sum_{i \in I(x)}(x_i^{(1)} - x_i^{(2)})A_{.i}$. Da die Vektoren $A_{.i}$ linear unabhängig sind, folgt $x_i^{(1)} - x_i^{(2)} = 0 \Rightarrow x_i^{(1)} = x_i^{(2)}$. Somit ist gezeigt, dass $x^{(1)} = x^{(2)}$. ∎

Bemerkung 2.33
Bei einer Auswahl von Spaltenvektoren von A können höchstens m linear unabhängig sein, d. h. $I(x)$ hat höchtens m Elemente und es folgt, dass eine Ecke höchtens m positive Komponenten haben kann.

Beispiel 2.34
siehe Satz 2.31. Der Ursprung liegt im Satz 2.10.

Die Ecken der zulässigen Mengen sind anschaulich die Punkte:

$$x^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} \qquad x^{(2)} = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} \qquad x^{(3)} = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$$

Es gilt: $A = (1/2, 1, 1), b = 3, m = 1, n = 3$.

Beispiel 2.35 (siehe Satz 2.30)
Die zulässige Menge in kanonischer Form wird definiert durch:

$$A = \begin{pmatrix} -2 & -3 \\ -1 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} -6 \\ -2 \end{pmatrix}$$

Um Satz 2.32 anzuwenden, müssen wir das Problem in Standardform bringen:

$$\tilde{A} = \begin{pmatrix} -2 & -3 & -1 & 0 \\ -1 & 0 & 0 & -1 \end{pmatrix}$$

Kandidaten für die Ecken sind:

$$x^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 6 \\ 2 \end{pmatrix} \qquad x^{(2)} = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 0 \end{pmatrix} \qquad x^{(3)} = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 2 \end{pmatrix} \qquad x^{(4)} = \begin{pmatrix} 2 \\ 2/3 \\ 0 \\ 0 \end{pmatrix}$$

Sei $x \in \mathcal{F}_k$ und $s = Ax - b \geq 0 \Rightarrow \begin{pmatrix} x \\ s \end{pmatrix} \in \mathcal{F}$ zulässige Punkte für die Standardform.

Bemerkung 2.36
Sei $\mathcal{F}_k = \{\, x = (x_1, \dots, x_n)^T \in \mathbb{R}^n \mid x \geq 0, Ax \geq b \,\}$ die zulässige Menge eines linearen Optimierungsproblems in kanonischer Form. Transformiert man das Problem in Standardform, erhält man die zulässige Menge:

$$\mathcal{F} = \{\, \tilde{x} = (x_1, \dots, x_n, \underbrace{x_{n+1}, \dots, x_{n+m}}_{=s})^T \in \mathbb{R}^{n+m} \mid \tilde{x} \geq 0, (A \;\; -I_m)\tilde{x} = b \,\}$$

Dann gilt:

- $x \in \mathcal{F}_k$ und $s = Ax - b \Leftrightarrow \begin{pmatrix} x \\ s \end{pmatrix} \in \mathcal{F}$
- $x \in \mathcal{F}_k$ ist Ecke von $\mathcal{F}_k \Leftrightarrow \begin{pmatrix} x \\ s \end{pmatrix}$ mit $s = Ax - b$ ist Ecke von $\mathcal{F}$.

Satz 2.37
Die zulässige Menge $\mathcal{F} = \{\, x \in \mathbb{R}^n \mid x \geq 0, Ax = b \,\}$ eines linearen Optimierungsproblems in Standardform sei nicht leer. Dann existiert mindestens eine und höchstens endlich viele Ecken von $\mathcal{F}$.

Beweis einfügen

BEWEIS: 5

Lemma 2.38
Sei $\mathcal{F} = \{\, x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \,\} \neq \emptyset$ zulässige Menge eines linearen Optimierungsproblems in Standardform. Dann ist $\mathcal{F}$ genau dann beschränkt, wenn es kein $d \in \mathbb{R}^n \setminus \{0_n\}$ gibt mit $Ad = 0, d \geq 0$.

Beweis einfügen

BEWEIS: 6

Satz 2.39
Seien $\mathcal{F} = \{\, x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \,\} \neq \emptyset$ und $v^{(1)}, \dots, v^{(k)}$ Ecken von $\mathcal{F}$. Jeder Punkt $x \in \mathcal{F}$ hat eine Darstellung:

$$x = \sum_{i=1}^{k} \alpha_i v_k^{(i)} + d$$

mit $\alpha_i \geq 0, i = 1, \ldots, k, \sum_{i=1}^k \alpha_i = 1, d \geq 0, Ad = 0$. Ist $\mathcal{F}$ beschränkt, gilt:

$$\mathcal{F} = \operatorname{co}\{v^{(1)}, \ldots, v^{(k)}\} \tag{2.1}$$

Beweis: 7
Sei p_0 die Minimalzahl der positiven Komponenten von Vektoren $x \in \mathcal{F}$. Der Beweis wird nach dem Prinzip der vollständigen Induktion geführt:

Induktionsanfang $p = p_0$

Induktionsvoraussetzung Es gelte Gleichung 2.1 für $x \in \mathcal{F}$ und höchstens $n-1$ positiven Komponenten.

Induktionsbeweis Sei $x \in \mathcal{F}$ ein Vektor mit genau p positiven Komponenten. Wir legen den Index mit $I(x) := \{ 1 \leq i \leq n \mid x_i \geq 0 \}$ fest. Sei nun x keine Ecke von $\mathcal{F}$. Nach Satz 2.32 folgt, dass die $A_{.i}$ linear abhängig sind, d. h. $\exists w \in \mathbb{R}^n \setminus \{0\}$ mit $w_i = 0$ für $i \notin I(x)$ und $\sum_{i=1}^n w_i A_{.i} = 0$.

1. Fall Es gibt mindestens ein $w_j < 0$ und $w_k > 0$. Dann sind $x^{(1)}, x^{(2)} \in \mathcal{F}$ und haben höchstens $p-1$ positive Komponenten. Nach der Induktionsvoraussetzung haben die $x^{(1)}, x^{(2)}$ eine Darstellung der Form:

$$x^{(i)} = \sum_{j=1}^k \alpha_j^{(i)} v^{(j)} + d^{(i)} \qquad i = 1, 2$$

mit $\alpha_j^{(i)} \geq 0, \sum_{j=1}^k \alpha_j^{(i)} = 1, d^{(i)} \geq 0, Ad^{(i)} = 0$. Weiter gilt:

$$\begin{aligned}
x &= \underbrace{\frac{\lambda_2}{\lambda_1 + \lambda_2}}_{=(1-\mu)} x^{(1)} + \underbrace{\frac{\lambda_1}{\lambda_1 + \lambda_2}}_{=\mu \in (0,1)} = x^{(2)}(1-\mu)x^{(1)} + \mu x^{(2)} \\
&= (1-\mu) \sum_{j=1}^k \alpha_j^{(i)} v^{(j)} + d^{(i)} + \mu \sum_{j=1}^k \alpha_j^{(i)} v^{(j)} + d^{(i)} \\
&= \left(\sum_{j=1}^k \underbrace{(1-\mu)\alpha_j^{(i)} + \mu \alpha_j^{(i)}}_{=:\alpha_j} \right) v^{(j)} + (1-\mu) d^{(1)} + \mu d^{(2)} \\
&\Rightarrow \alpha_j \geq 0, \sum_{j=1}^k \alpha_j = 1
\end{aligned}$$

Für $d := (1-\mu)d^{(1)} + \mu d^{(2)} \geq 0$ und $Ad = (1-\mu)Ad^{(1)} + \mu Ad^{(2)}$.

2. Fall $w \geq 0$
Dann definiert man $x^{(1)}$ und λ_1 wie oben. Dann ist $x^{(1)} \in \mathcal{F}$ und hat höchstens $p-1$ positive Komponenten. Mit der Anwendung der Induktionsvoraussetzung auf $x^{(1)}$ folgt $x = x^{(1)} + \lambda_1 w = \sum_{j=1}^k \alpha_j^{(1)} v^{(j)} + d^{(1)} + \lambda_1 w$. Für $d := d^{(1)} + \lambda_1 w$ gilt $d \geq 0$ und $Ad^{(1)} + \lambda_1 w = 0$.

3. Fall $w \leq 0$
Man definiert $x^{(2)}$ und λ_2 wie oben. Die Beweisführung erfolgt wie im zweiten Fall.

Sei $\mathcal{F}$ nun beschränkt. Dann folgt nach Satz 2.38, dass es kein $d \in \mathbb{R}^n \setminus \{0\}$ mit $d \geq 0$ und $Ad = 0$ gibt. Damit hat jedes $x \in \mathcal{F}$ eine Darstellung wie in Gleichung 2.1 mit $d = 0_n$, d. h. für alle $x \in \mathcal{F}$ ist Konvexkombination der Ecken $v^{(1)}, \ldots, v^k \in \mathcal{F} \Rightarrow \mathcal{F} \subseteq \mathrm{co}\{v^{(1)}, \ldots, v^{(k)}\}$. Da $\mathcal{F}$ konvex und $v^{(1)}, \ldots, v^{(k)} \in \mathcal{F} \Rightarrow \mathcal{F} \supseteq \mathrm{co}\{v^{(1)}, \ldots, v^{(k)}\}$. ■

Satz 2.40 (Hauptsatz der linearen Optimierung)
Sei $A \in \mathbb{R}^{m \times n}, m \leq n, b \in \mathbb{R}^m$. Wir betrachten das lineare Optimierungsproblem in Standardform:

$$\min c^T x \text{ unter den Nebenbedingungen} x \in \mathcal{F} = \{\, x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \,\}$$

Dann gilt:

(a) Entweder ist eine der endlich vielen Ecken des linearen Optimierungsproblems in Standardform Lösung oder die Zielfunktion ist auf $\mathcal{F}$ nicht nach unten beschränkt.

(b) Ist $\mathcal{F}$ beschränkt, dann hat das lineare Optimierungsproblem in Standardform mindestens eine Lösung. Ein $x \in \mathcal{F}$ ist genau dann Lösung, wenn gilt: x ist Konvexkombination von Ecken, die ebenfalls Lösung sind.

Beweis: 8
Es existiere ein $d \in \mathbb{R}^n$ mit $d \geq 0, Ad = 0$ und $c^T d < 0$. Dann ist die Zielfunktion auf $\mathcal{F}$ nicht nach unten beschränkt. Denn sei $x \in \mathcal{F}$ beliebig. Dann folgt für alle $t \geq 0$: $x + td \in \mathcal{F} \Rightarrow c^T(x + td) = c^T x + tc^T d \xrightarrow{t \to \infty} -\infty$.

Betrachten jetzt ein beliebiges $d \in \mathbb{R}^n$ mit $d \geq 0, Ad = 0$: $c^T d \geq 0$. Es ist zu zeigen, dass mindestens eine der Ecken von $\mathcal{F}$ Lösung ist. Seien $v^{(1)}, \ldots, v^{(k)}$ Ecken. Nach Satz 2.39 hat jedes $x \in \mathcal{F}$ die Darstellung:

$$\begin{aligned} x &= \sum_{i=1}^{k} \alpha_i v^{(i)} + d \\ \Rightarrow c^T x &= c^T \left(\sum_{i=1}^{k} \alpha_i v^{(i)} + d \right) = \sum_{i=1}^{k} \alpha_i c^T v^{(i)} + \underbrace{c^T d}_{\geq 0} \end{aligned}$$

Sei $\mu := \min_{i=1,\ldots,k}\{c^T v^{(i)}\}$. Dann gilt $c^T x \geq \sum_{i=1}^{k} \alpha_i c^T v^{(i)} \geq \sum_{i=1}^{k} \alpha_i \mu = \mu \sum_{i=1}^{k} \alpha_i$. Somit ist jede Ecke $v^{(j)}$ von $\mathcal{F}$ mit $c^T v^{(j)} = \mu$ Lösung und der erste Teil des Satzes ist gezeigt.

Für den zweiten Teil des Satzes sei nun $\mathcal{F}$ beschränkt. Da $\mathcal{F}$ weiterhin auch abgeschlossen ist, folgt, dass $\mathcal{F}$ kompakt ist. Die Zielfunktion $f(x) = c^T x$ ist stetig. Nach dem Satz von Weierstraß[2] nimmt f auf $\mathcal{F}$ ihr Minimum an.

[2] Eine reelle und stetige Funktion auf einer kompakten Teilmenge des $\mathbb{R}^n$ ist beschränkt und nimmt die untere und obere Grenze an.

Sei x beliebige Lösung des linearen Optimierungsproblems in Standardform. Dann ist zu zeigen, dass das x eine Konvexkombination ist. Nach Satz 2.39 gilt, $x = \sum_{i=1}^k \alpha_i v^{(i)}$ mit $\alpha_i \geq 0, \sum_{i=1}^k \alpha_i = 1, i = 1, \dots, k$. Wir zeigen, $\alpha_i = 0$ für alle nicht optimalen Ecken. Wir wissen $\mu = c^T x = \sum_{i=1}^k \alpha_i c^T v^{(i)} = \min_{i=1,\dots,k}\{c^T v^{(i)}\}$, denn eine optimale Lösung existiert. Sei jetzt $j \in \{1, \dots, k\}$ für den $v^{(j)}$ nicht optimal ist. Dann ist $c^T v^{(j)} = \mu + \varepsilon$ mit $\varepsilon > 0 \Rightarrow \mu = \sum_{i=1}^k \alpha_i \underbrace{c^T v^{(i)}}_{\geq \mu} \geq \sum_{i=1}^k \alpha_i \mu + \alpha_j \varepsilon = \mu + \alpha_j \varepsilon \Rightarrow \alpha_j = 0$. Es ist nun noch zu zeigen: wenn x eine Konvexkombination optimaler Ecken ist, so ist auch x optimal. Sei $x \leq \sum_{j=1}^p \alpha_{ij} v^{(ij)}, p \leq k$ und optimale Ecken $v^{(ij)}, j = 1, \dots, p$ und $\alpha_{ij} \geq 0, j = 1, \dots, p, \sum \alpha_{ij} = 1$. Es gilt: $c^T v^{(ij)} = \mu \Rightarrow c^T x = \sum_{j=1}^p \alpha_{ij} \underbrace{c^T v^{(ij)}}_{=\mu} = \mu$. ∎

Wäre es hier besser i_j statt ij zu schreiben?

2.4. Basislösungen

Sei $J_B \subset \{1, \dots, n\}, J_N = \{1, \dots, n\} \setminus J_B$. Somit ist $J_B \cup J_N = \{1, \dots, n\}$ und $J_B \cap J_N = \emptyset$. Für einen Vektor $x \in \mathbb{R}^n$ bezeichnen wir mit:

- x_B den Vektor mit den Komponenten $x_i, i \in J_B$
- x_N den Vektor mit den Komponenten $x_i, i \in J_N$

Entsprechend bezeichnen wir für eine $m \times n$-Matrix A mit:

- A_B die Matrix mit den Spaltenvektoren $A_{.i}, i \in J_B$
- A_N die Matrix mit den Spaltenvektoren $A_{.i}, i \in J_N$

Sei $m \leq n$ und $A \in \mathbb{R}^{m \times n}$ mit vollem Zeilenrang $m \in \mathbb{R}^m$. Weiter sei $J_B \subset \{1, \dots, n\}$ mit genau m Elementen, so dass die Matrix A_B invertierbar ist. Wie oben sei J_N. Wir definieren B, N durch:

$$B := \{ A_{.i} \mid i \in J_B \} \qquad N := \{ A_{.i} \mid i \in J_N \}$$

Beispiel 2.41
Sei:

$$(A|b) = \begin{pmatrix} 0 & 1 & 5 & 1 & | & 2 \\ 1 & 0 & -1 & 0 & | & 2 \\ 1 & 0 & 2 & 1 & | & 3 \end{pmatrix} \qquad m = 3 \qquad n = 4$$

Wir wählen $J_B = \{1,2,4\} \Rightarrow J_N = \{3\}$. Die Matrix $A_B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ ist invertierbar.

$$A_N = \begin{pmatrix} 5 \\ -1 \\ 2 \end{pmatrix} \qquad x_B = \begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix} \qquad x_N = (x_3)$$

Sei $x \in \mathbb{R}^n$ eine Lösung des linearen Systems $Ax = b$. Dann gilt:

$$b = Ax = (A_B|A_N)\begin{pmatrix} x_B \\ x_N \end{pmatrix} = A_B x_B + A_N x_N$$

$$x_B = A_B^{-1}(b - A_N x_N) \tag{2.2}$$

Das bedeutet, dass man das System nach x_B auflösen kann und es folgt, dass die Lösungsmenge

$$\{\, x \in \mathbb{R}^n \mid x_N \in \mathbb{R}^{n-m}, x_B = A_B^{-1}(b - A_N x_N)\,\}$$

ein Unterraum des $\mathbb{R}^n$ ist. Die Dimension des Lösungsraumes ist $n - m$. In jedem $x_N \in \mathbb{R}^{n-m}$ existiert genau eine Lösung, wobei x_B durch die Gleichung 2.2 bestimmt ist.

Für die Optimierung ist die Lösung x zu $x_N = 0_{n-m}$ von besonderem Interesse, d. h. $x_B = A_B^{-1} b \Leftrightarrow A_B x_B = b$.

Definition 2.42 (Basislösung, Basisvariablen, Nichtbasisvariablen)
Sei $m \le n, A \in \mathbb{R}^{m\times n}$, wobei der Rang der Matrix A gleich m ist. Weiter sei $J_B \subset \{1, \ldots, n\}$ eine Indexmenge, so dass A_B invertierbar ist. Der eindeutig bestimmte Vektor x mit $A_B x_B = b, x_N = 0_{n-m}$ heißt **Basislösung** des Systems $Ax = b$ zur Basis B. Die Variablen x_B heißen **Basisvariablen**, x_N heißen **Nichtbasisvariablen**. Die Basislösung x zur Basis B heißt **zulässige Basislösung**, falls $x \ge 0$.

Bemerkung 2.43
Wegen $x_N = 0$ hat eine Basislösung höchstens m von Null verschiedene Komponenten. Für die Zulässigkeit eines Basislösung reicht es zu fordern, dass $x_B \ge 0_m$ ist.

Beispiel 2.44
Es sei ein lineares Optimierungsproblem in kanonischer Form mit den Nebenbedingungen $Ax \ge b$ gegeben. Wir benutzen die Schlupfvariablen $x_{n+1}, \ldots, x_{n+m}$ zur Transformation auf die Standardform:

$$\underbrace{(A \;\; -I_m)}_{=:\tilde{A}} \begin{pmatrix} x_1 \\ \vdots \\ x_{n+m} \end{pmatrix} = b$$

Der Rang der Matrix $\tilde{A}$ ist m. Hier bietet es sich an, $J_B = \{n+1, \dots, n+m\}$, $J_N = \{1, \dots, n\}$ zu wählen, d. h. $A_B = -I_m$, $B = (-e_1, \dots, -e_m)$, wobei e_i der i-te Einheitsvektor ist.
zulässige Basislösung:

$$x_N = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0_n \qquad\qquad x_B = A_B^{-1} b = -I_m^{-1} b = -I_m b = -b$$

Das bedeutet, $x = (0, \dots, 0, -b_1, \dots, -b_m)^T$ und x ist zulässig, wenn $x_B = -b \geq 0_m$, d. h. $b \leq 0_m$.

Beispiel 2.45 (siehe Satz 2.41)

$$(A|b) = \begin{pmatrix} 0 & 1 & 5 & 1 & | & 2 \\ 1 & 0 & -1 & 0 & | & 2 \\ 1 & 0 & 2 & 1 & | & 3 \end{pmatrix} \qquad m = 3 \qquad n = 4$$

Wir wählen $J_B = \{1, 2, 4\}$, $J_N = \{3\}$. Dann ist $A_B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. Die eindeutig bestimmte Lösung zur Basis B ist gegeben durch $x_N = (x_3) = 0_{4-3} = 0_1$ und $x_B = A_B^{-1} b = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \Leftrightarrow A_B x_B = b$

Satz 2.46 (Allgemeine Charakterisierung einer Ecke)
Sei $\mathcal{F} = \{ x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \}$ zulässige Menge eines linearen Optimierungsproblems in Standardform, $A \in \mathbb{R}^{m \times n}, m \leq n, \operatorname{rg}(A) = m, b \in \mathbb{R}^m$. Ein Punkt $x \in \mathcal{F}$ ist genau dann Ecke von $\mathcal{F}$, wenn x zulässige Basislösung ist.

Beweis: 9

„$\Rightarrow$" Sei x Ecke von $\mathcal{F}$ und $I(x) = \{ 1 \leq i \leq n \mid x_i > 0 \}$. Nach der Voraussetzung ist $b = Ax = \sum_{i=1}^n x_i A_{.i} = \sum_{i \in I(x)} x_i A_{.i}$. Nach Satz 2.32 sind die Spaltenvektoren $A_{.i}$ linear unabhängig. Es sei $\tilde{B} := \{ A_{.i} \mid i \in I(x) \}$ und p die Anzahl der Elemente von $\tilde{B}$. Falls $p = m$, dann ist $\tilde{B}$ Basis von A und x ist zugehörige Basislösung. Da $x \in \mathcal{F}$, ist x zulässig und es folgt, dass x zulässige Basislösung ist. Ist nun $p < m$, dann kann man $\tilde{B}$ wegen $\operatorname{rg}(A) = m$ zu einer Basis von A ergänzen und x ist wie oben zulässige Basislösung.

„$\Leftarrow$" Sei x zulässige Basislösung zu einer Basis B von A. Dann ist $x_N = 0_{n-m} \Rightarrow I(x) \subset J_B$. Die Spaltenvektoren $A_{.i}, i \in J_B$ sind linear unabhängig. Damit sind auch die Spaltenvektoren $A_{.i}, i \in I(x)$ linear unabhängig und nach Satz 2.32 ist x eine Ecke. ∎

Korollar 2.47
Die Voraussetzungen seien wie oben. Dann existiert mindestens eine Ecke von $\mathcal{F}$ und höchstens $\binom{n}{m}$ Ecken.

Beweis: 10
Die Existenz einer Ecke ergibt sich aus Satz 2.37. Es existieren $\binom{n}{m}$ Möglichkeiten um Spaltenvektoren auszuwählen. Damit gibt es höchstens $\binom{n}{m}$ Basislösungen. ∎

Beispiel 2.48

$$A = (1/2, 1, 1) \qquad b = (3) \qquad m = 1 \qquad n = 3$$

Es gibt höchstens $\binom{n}{m} = 3$ Basislösungen. Für $B = B_1 = \{A_{.1}\}$ gilt $A_B^{-1}b = 2 \cdot 3 = 6 \Rightarrow x^{(1)} = (6, 0, 0)^T$. Für $B = B_2 = \{A_{.2}\}$ gilt $A_B^{-1}b = 3 \Rightarrow x^{(2)} = (0, 3, 0)^T$. Für $B = B_3 = \{A_{.3}\}$ gilt $A_B^{-1}b = 3 \Rightarrow x^{(3)} = (0, 0, 3)^T$. Es gilt $x^i \geq 0, i = 1, 2, 3 \Rightarrow$ alle Basislösungen sind zulässig und Ecken.

Beispiel 2.49

$$A = \begin{pmatrix} -2 & -3 & -1 & 0 \\ -1 & 0 & 0 & -1 \end{pmatrix} \qquad b = \begin{pmatrix} -6 \\ -2 \end{pmatrix} \qquad n = 4 \qquad m = 2$$

Damit gibt es höchstens sechs Basislösungen:

$$B_1 = \{A_{.1}, A_{.2}\} \quad : A_B^{-1}b = \begin{pmatrix} 2 \\ 2/3 \end{pmatrix} \quad x^{(1)} = (2, 2/3, 0, 0)^T \tag{2.3}$$

$$B_2 = \{A_{.1}, A_{.3}\} \quad : A_B^{-1}b = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad x^{(2)} = (2, 0, 2, 0)^T \tag{2.4}$$

$$B_3 = \{A_{.1}, A_{.4}\} \quad : A_B^{-1}b = \begin{pmatrix} 3 \\ -1 \end{pmatrix} \quad x^{(3)} = (3, 0, 0, -1)^T \tag{2.5}$$

$$B_4 = \{A_{.2}, A_{.3}\} \quad : A_B^{-1}b \quad \text{keine Basislösung} \tag{2.6}$$

$$B_5 = \{A_{.2}, A_{.4}\} \quad : A_B^{-1}b = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad x^{(5)} = (0, 2, 0, 2)^T \tag{2.7}$$

$$B_6 = \{A_{.3}, A_{.4}\} \quad : A_B^{-1}b = \begin{pmatrix} 6 \\ 2 \end{pmatrix} \quad x^{(6)} = (0, 0, 6, 2)^T \tag{2.8}$$

Außer den Gleichungen 2.5 (negative Komponente) und 2.6 (Matrix nicht invertierbar) sind alle anderen Gleichungen Ecken.

Bei der Lösung eines linearen Optimierungsproblems können folgende Situationen auftreten:

1. $\mathcal{F} = \emptyset$: Dann existiert keine zulässige Basislösung.
2. $\mathcal{F} \neq \emptyset$ und die Zielfunktion ist auf $\mathcal{F}$ nach unten beschränkt: Dann existiert mindestens eine Ecke, die Lösung ist.
3. $\mathcal{F} \neq \emptyset$ und die Zielfunktion ist nicht nach unten beschränkt.

Für den zweiten Fall kann man folgende „naive“ Strategie anwenden:

- Bestimme alle möglichen Basislösungen
- Teste auf Zulässigkeit und erhalte so alle Ecken
- Bestimme die Ecke mit dem minimalen Funktionswert

3. Das Simplexverfahren

Das Simplexverfahren wurde 1947 von George Dantzig entwickelt. Der amerikanische Mathematiker lebte 1914 bis 2005 und war Mitbegründer der Mathematical Programming Society.

Der Ausgangspunkt für das Simplexverfahren ist ein lineares Optimierungsproblem in Standardform $\min c^T x$ unter den Nebenbedingungen $Ax = b, x \geq 0$ mit $A \in \mathbb{R}^{m \times n}, c \in \mathbb{R}^n, b \in \mathbb{R}^m$. Als Voraussetzung soll A den vollen Zeilenrang besitzen und es ergibt sich die folgende Vorgehensweise:

1. Man bestimme die Ausgangsecke (Phase 1).
2. Iteration: Ist die aktuelle Ecke optimal, dann stoppe.
3. Falls die Ecke nicht optimal ist, berechne eine bessere Ecke. Dies wird solange wiederholt, bis die optimale Ecke gefunden wurde (Phase 2), wobei die Ecken zulässige Basislösungen sind.

3.1. Grundlagen des Simplexverfahren

Ist x Ecke zur Basis B, dann gilt: $x_N = 0_{n-m}, x_B = A_B^{-1}(b - A_N x_N)$. In der Phase zwei ergibt sich folgende Vorgehensweise:

- Test, ob Ecke x optimal ist.
- Falls x nicht optimal, Test, ob Zielfunktion auf zulässiger Menge nach unten beschränkt ist. Falls nein, existiert keine Lösung und das Programm kann abbrechen.
- Berechnung einer neuen Ecke mit dem Ziel, dass der Zielfunktionswert in der neuen Ecke kleiner ist.

3.1.1. Reduktion der Variablen

Wir betrachen das System $Ax = b$ mit B Basis von A. Dann können wir das System nach x_B auflösen:

$$x_B = A_B^{-1}(b - A_N x_N) \qquad x_N \in \mathbb{R}^{n-m} \tag{3.1}$$

Somit kann das Ausgangsproblem mit den n Variablen auf ein Problem in den Variablen $x_N \in \mathbb{R}^{n-m}$ reduziert werden.

Das Einsetzen in die Zielfunktion ergibt:

$$\begin{aligned} c^T x &= c_B^T + x_B c_N^T x_N = c_B^T A_B^{-1}(b - A_N x_N) + c_N^T x_N \\ &= \underbrace{c_B^T A_B^{-1} b}_{=: z_B} + \underbrace{(c_N^T - c_B^T A_B^{-1} A_N)}_{=: s_N^T} x_N \end{aligned}$$

Dabei ist z_B konstant und s_N heißt Vektor der reduzierten Kosten. Mit:

$$s_N^T = c_N^T - c_B^T A_B^{-1} A_N \in \mathbb{R}^{n-m} \tag{3.2}$$

gilt

$$c^T x = z_B + s_N^T x_N \tag{3.3}$$

Die Vorzeichenbedingung $x \geq 0$ ist äquivalent zu $x_B \geq 0 \wedge x_N \geq 0$. Es gilt $x_B \geq 0 \Leftrightarrow A_B^{-1}(b - A_N x_N) \geq 0 \Leftrightarrow A_B^{-1} A_N x_N \leq A_B^{-1} b$. Damit erhalten wir das zum linearen Problem äquivalente, reduzierte Problem mit $x_N = v$

$$\min s_N^T v + z_B \tag{3.4}$$

$$\text{unter den Nebenbedingungen} v \in \mathcal{F}_B = \{ v \in \mathbb{R}^{n-m} \mid v \geq 0, A_B^{-1} A_N v \leq A_B^{-1} b \} \tag{3.5}$$

Die Konstante z_B wird in der Regel weggelassen.

Lemma 3.1
Die Probleme seien wie oben definiert. Dann gilt:

(i) Ist $v \in \mathcal{F}_B$ und ist $x \in \mathbb{R}^n$ definiert durch $x_N = v, x_B = A_B^{-1}(b - A_N v)$. Dann ist $x \in \mathcal{F}$. Ist nun $x \in \mathcal{F}$, dann ist $x_N \in \mathcal{F}_B$.

   Somit hat man eine eineindeutige Zuordnung: $\mathcal{F}_B \ni v \Leftrightarrow x \in \mathcal{F}$ mit $x_N = v, x_B = A_B^{-1}(b - A_N x_N)$.

   Beweis: 11
   klar ■

(ii) Ist $v^* \in \mathcal{F}_B$ Lösung des reduzierten Problems und ist x^* definiert durch $x_N^* = v^*, x_B^* = A_B^{-1}(b - A_N v^*)$, dann ist x^* Lösung des linearen Optimierungsproblems. Ist umgekehrt $x^* \in \mathcal{F}$ Lösung des linearen Problems und $v^* = x_N^*$, dann ist $v^* \in \mathcal{F}_B$ Lösung des reduzierten Problems.

   Beweis: 12
   Sei $v^* \in \mathcal{F}_B$ Lösung von Gleichung 3.4, x^* definiert durch $x_N^* = v^*, x_B^* = A_B^{-1}(b - A_N v^*)$. Nach dem ersten Punkt ist $x^* \in \mathcal{F}$. Sei $x \in \mathcal{F}$ beliebig. Dann gilt nach Gleichung 3.3: $c^T x = z_B + s_N^T x_N \geq z_B + s_N^T v^* = c^T x^*$, d. h. x^* ist Lösung des linearen Problems.

Sei umgekehrt x^* Lösung des linearen Problems und $v^* = x^*_N$. Nach dem ersten Punkt ist $v^* \in \mathcal{F}_B$. Sei $v \in \mathcal{F}_B$ und x definiert durch $x_N = v, x_B = A_B^{-1}(b - A_N x_N)$. Dann ergibt sich $z_B + s_N^T v = z_B + s_N^T x_N = c^T x \geq c^T x^* = z_B + s_N^T x^*_N = z_B + s_N^T v^*$, d. h. $s_N^T v \geq s_N^T v^*$ für alle $v \in \mathcal{F}_B$. ∎

(iii) Ist x^* zulässige Basislösung des linearen Problems zur Basis B, dann ist x^* genau dann Lösung, wenn $x^*_N = 0_{n-m}$ Lösung von Gleichung 3.4 ist.

Beweis: 13
Sei x^* Basislösung zur Basis B. Dann ist $x^*_N = 0_{n-m}$. Nach Punkt (ii) ist x^* genau dann Lösung des linearen Problems, wenn $v^* = x^*_N = 0_{n-m}$ Lösung von Gleichung 3.4 ist. ∎

Damit folgt der Optimalitätstest: Sei $x^* \in \mathcal{F}$ zulässige Basislösung zur Basis B. Man stellt (theoretisch) die reduzierte Gleichung 3.4 auf und prüft, ob $v^* = 0_{n-m}$ Lösung von Gleichung 3.4 ist. Falls ja, ist x^* Lösung des linearen Problems.

3.1.2. Ein Optimalitätskriterium

Gesucht ist eine hinreichende Bedingung dafür, dass $v^* = 0_{n-m}$ Lösung von Gleichung 3.4 ist.

Es gilt $s_N^T v^* = 0$, d. h. v^* ist Lösung, wenn gilt $s_N^T v \geq s_N^T v^* = 0$. Für $v \in \mathcal{F}_B$ gilt $v \geq 0$ und es folgt, wenn

$$s_N \geq 0 \tag{3.6}$$

erfüllt ist, gilt $s_N^T v \geq 0 = s_N^T v^*$

Lemma 3.2
Ist unter den obigen Voraussetzungen $s_N \geq 0$, dann ist die aktuelle Ecke x^* optimal. Wenn $s_N > 0$, dann ist x^* eindeutig bestimmt.

Beweis: 14
Der erste Teil ist klar. Für den zweiten Teil sei $x \in \mathcal{F}$. Dann ist $x_N \geq 0, x_B = A_B^{-1}(b - A_N x_N) \geq 0$. Ist $x \neq x^* \Rightarrow x_N \neq x^*_N = v^* = 0_{n-m}$, d. h. für mindestens eine Komponente $j \in J_N$ gilt $x_j > 0 \Rightarrow s_N^T x_N > 0 \Rightarrow c^T x = z_B + \underbrace{s_N^T x_N}_{>0} > z_B + \underbrace{s_N^T v^*}_{=0} = c^T x^*$, d. h. $c^T x > c^T x^*$ für alle $x \in \mathcal{F}, x \neq x^*$. Somit ist x eindeutig bestimmte Lösung. ∎

Für den Optimalitätstest ergibt sich nun: Sei x^* die aktuelle Ecke eine zulässige Basislösung zur Basis B. Berechne s_N nach Gleichung 3.2, prüfe, ob $s_N \geq 0$. Falls ja, ist x^* optimal.

Beispiel 3.3
Sei folgendes Problem gegeben:

$$\begin{aligned} \min 2x_1 + 6x_2 - 7x_3 + 2x_4 + 4x_5 & \\ 4x_1 - 3x_2 + 8x_3 - x_4 \qquad &= 12 \\ - x_2 + 12x_3 - 3x_4 + 4x_5 &= 20 \\ x_i &\geq 0 \end{aligned}$$

Damit folgt:

$$A = \begin{pmatrix} 4 & -3 & 8 & -1 & 0 \\ 0 & -1 & 12 & -3 & 4 \end{pmatrix} \qquad b = \begin{pmatrix} 12 \\ 20 \end{pmatrix} \qquad c = (2, 6, -7, 2, 4)^T$$

Wir wählen $J_B = \{1, 5\}, J_N = \{2, 3, 4\}, B = \{A_{.1}, A_{.5}\}$.

$$A_B = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \qquad A_B^{-1} = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix}$$

$$c_B = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \qquad A_B^{-1} b = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

$$A_B^{-1} A_N = \begin{pmatrix} -3/4 & 2 & -1/4 \\ -1/4 & 3 & -3/4 \end{pmatrix} \qquad s_N^T = c_N^T - c_B^T A_B^{-1} A_N = (17/2, -23, 11/2)^T$$

Somit ist die hinreichende Optimalitätsbedingung nicht erfüllt.

Nunmehr stellt sich die Frage, ob die Bedingung $s_N^T \geq 0$ auch notwendig ist. Dies ist mit ja zu beantworten, falls die Ecke x^* nicht entartet ist.

Definition 3.4 (Nicht entartete Basislösung)
Sei $A \in \mathbb{R}^{m \times n}$ mit $\text{rg}(A) = m \leq n, b \in \mathbb{R}^m, \mathcal{F} = \{ x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \}$, B Basis von A. Eine zulässige Basislösung $x \in \mathcal{F}$ heißt **nicht entartet**, wenn $x_B > 0$.

Bemerkung 3.5
Zu jeder Basis B existiert genau eine Basislösung x mit $x_N = 0_{n-m}, x_B = A_B^{-1}(b - A_N x_N)$, d. h. zu jeder Basis existiert höchstens eine Ecke.

Ist $x \in \mathcal{F}$ Ecke, dann kann x zu verschiedenen Basen die zugehörige Ecke sein.

Beispiel 3.6

$$A = (I_m, I_m) \qquad b = (0, 1, \ldots, 1)^T \in \mathbb{R}^m \qquad x = \begin{pmatrix} b \\ 0_m \end{pmatrix}$$

Dann ist x zulässige Basislösung zur Basis mit den Indizes $J_B = \{1, \ldots, m\}$, aber auch zu $J_B = \{2, \ldots, m + 1\}$

Lemma 3.7
Sei $x \in \mathcal{F}$ eine nicht entartete Ecke. Ist $x^* \in \mathcal{F}$ Lösung des linearen Problems, dann ist $s_N \geq 0$.

Beweis: 15
Nach Satz 3.1 gilt $v^* = x_N^* = 0_{n-m}$ ist Lösung des reduzierten Problems. Angenommen, es existiert ein $j \in \{1, \ldots, n-m\}$ mit $s_{N,j} < 0$. Wir definieren die Richtung $d \in \mathbb{R}^{n-m}$ durch $d_j = 1, d_i = 0$ für $i \in \{1, \ldots, n-m\} \setminus \{j\}$ (j-ter Einheitsvektor von $\mathbb{R}^{n-m}$) und zeigen, d ist eine zulässige Abstiegsrichtung im Punkt $v^* = 0_{n-m}$. Sei $v(t) := v^* + td = td \geq 0$ für $t \geq 0$ sowie $\widehat{A} := A_B^{-1} A_N$ und $\widehat{A_{.j}}$ der zu obigen Index j gehörige Spaltenvektor. Da x_B^* nicht entartet ist, gilt, $0 < x_B^* = A_B^{-1} b$. Für hinreichend kleines $t > 0$ gilt:

$$t\widehat{A_{.j}} < A_B^{-1} b \Rightarrow \exists \sigma > 0 \colon A_B^{-1} A_N v(t) = t\widehat{A_{.j}} \leq A_B^{-1} b \qquad \forall t \in [0, \sigma]$$

Damit ist $v(t) \in \mathcal{F}_B$ (zulässig für reduziertes Problem). Speziell gilt: $v(\sigma) \in \mathcal{F}_B$ und $s_N^T v(\sigma) = \sigma s_N^T d = \sigma s_{N,j} < 0$. Dies ist jedoch eine Widerspruch zur Optimalität von v^*. ■

3.1.3. Ein Unbeschränktheitskriterium

Sei wieder x^* Ecke und damit zulässige Basislösung zur Basis B von A. Das Optimalitätskriterium $s_N \geq 0$ sei nicht erfüllt, d. h. es existiert ein $j \in \{1, \ldots, n-m\} \colon s_{N,j} < 0$. Wir definieren wie oben $d \in \mathbb{R}^{n-m}$ durch $d_j = 1, d_i = 0$ für $i \in \{1, \ldots, n-m\} \setminus \{j\}$. Dann gilt wieder $v(t) := v^* + td = td \geq 0$ für alle $t \geq 0$. Weiter sei $\widehat{A} := A_B^{-1} A_N$. Für den j-ten Spaltenvektor gelte $\widehat{A_{.j}} \leq 0$. Dann folgt für $t \geq 0$:

$$A_B^{-1} A_N v(t) = t\widehat{A_{.j}} \leq 0 \leq x_B^* = A_B^{-1} b$$

d. h. $v(t) \in \mathcal{F}_B$ und $\mathcal{F}_B$ ist unbeschränkt. Nach Satz 3.1 ist $x(t) \in \mathcal{F}$ mit $x_N(t) = v(t), x_B(t) = A_B^{-1}(b - A_N v(t))$. Damit ist auch $\mathcal{F}$ nicht beschränkt.

Zusammen mit $s_{N,j} < 0$ folgt, $c^T x(t) = z_B + s_N^T v(t) = z_B + t s_{N,j} \xrightarrow{t \to +\infty} -\infty$, d. h. die Zielfunktion des linearen Problems ist auf $\mathcal{F}$ nicht nach unten beschränkt. Also existiert keine Lösung.

Lemma 3.8
Existiert ein Index $j \in \{1, \ldots, n-m\}$ mit

$$s_{N,j} < 0 \text{ und } \widehat{A_{.j}} \leq 0 \tag{3.7}$$

dann ist $\mathcal{F}$ nicht nach unten beschränkt und die Zielfunktion ist auf $\mathcal{F}$ nicht nach unten beschränkt.

Beispiel 3.9 (Fortsetzung von Satz 3.3)
Im obigen Beispiel hatten wir den Vektor $s_N^T = (17/2, -23, 11/2)^T$ berechnet. Für $j = 2$ gilt: $\widehat{A_{.j}} = A_B^{-1} A_N = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix} \begin{pmatrix} 8 \\ 12 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} > 0$, d. h. die Gleichung 3.7 ist nicht erfüllt.

3.1.4. Basiswechsel

Sei $x^* \in \mathcal{F}$ Ecke mit Basis B von A. Das Optimalitätskriterium und das Unbeschränktheitskriterium seien nicht erfüllt. Das Ziel ist nun, eine neue Basis B^+ zu konstruieren, so dass für die Ecke x^+ gilt:

$$c^T x^+ \leq c^T x^*$$

Dazu definieren wir d wie oben. Dann gilt $v(t) := v^* + td \in \mathcal{F}_B$ mit $t \geq 0$ nicht zu groß. Bestimme das maximale σ so, dass $v(\sigma) \in \mathcal{F}_B$. Sei $x(\sigma)$ definiert durch $x_N(\sigma) = v(\sigma), x_B(\sigma) = A_B^{-1}(b - A_N v(\sigma))$. Dann kann man $x^+ = x(\sigma)$ wählen.

Bezeichnungen: $J_B = \{b(1), \ldots, b(m)\}, J_N = \{n(1), \ldots, n(m)\}$

Dann ist $B = \{A_{.,b(1)}, \ldots, A_{.,b(m)}\}, N = \{A_{.,n(1)}, \ldots, A_{.,n(m)}\}, x_{B,i} = x_{b(i)}, i = 1, \ldots, m, x_N = x_{n(i)}, i = 1, \ldots, n - m, \overline{A} := A_B^{-1}, \widehat{A} = A_B^{-1} A_N$.

Ist die Gleichung 3.6 nicht erfült, dann existiert ein

$$s_{N,j} < 0 \qquad j \in \{1, \ldots, n - m\} \tag{3.8}$$

Wir definieren:

$$d_j = 1 \qquad d_i = 0 \qquad i \in \{1, \ldots, n - m\} \setminus \{j\} \tag{3.9}$$

Damit ist d eine zulässige Abstiegsrichtung für das reduzierte Problem, falls x^* nicht entartet ist. Weiter definieren wir:

$$v(t) := \underbrace{v^*}_{0_{n-m}} + td = td \qquad \forall t \geq 0$$

Ist das Unbeschränktheitskriterium erfüllt, dann gilt $v(t) \in \mathcal{F}_B$. Sei also das Unbeschränktheitskriterium nicht erfüllt. Wir wollen das $\max \sigma \geq 0$ mit $v(t) \in \mathcal{F}_B$ für alle $t \in [0, \sigma]$ bestimmen. Wegen $v(t) \geq 0$ ist die Vorzeichenbedingung immer erfüllt. Wir müssen nur noch die Restriktion

$$\underbrace{A_B^{-1} A_N}_{=\widehat{A}} v(t) \leq A_B^{-1} b \tag{3.10}$$

einhalten. Diese Bedingung ist äquivalent zu:

$$t\widehat{A_{.j}} \leq x_B^* \tag{3.11}$$

Sei $\nu \in \{1, \ldots, m\}$ ein Index mit $\widehat{A_{\nu,j}} \leq 0 \Rightarrow t\widehat{A_{\nu,j}} \leq 0 \leq x_{B,\nu}^*$, d. h. entscheidend sind nur die Indizes $\nu \in \{1, \ldots, m\}$ mit $\widehat{A_{\nu,j}} > 0$. Für einen solchen Index gilt:

$$t\widehat{A_{\nu,j}} \leq x_{B,\nu}^* \Leftrightarrow t \leq \frac{x_{B,\nu}}{\widehat{A_{\nu,j}}}$$

Damit folgt:

$$\sigma = \min\left\{ \frac{x^*_{B,\nu}}{\widehat{A_{\nu,j}}} \;\middle|\; \nu \in \{1,\dots,m\} \text{ mit } \widehat{A_{\nu,j}} > 0 \right\} \tag{3.12}$$

Der Punkt $v(\sigma)$ ist **Randpunkt** von $\mathcal{F}_B$. Zunächst gilt, $x^+ \in \mathcal{F}$ wegen des Satz 3.1. Wir müssen noch eine Basis B^+ von A finden, so dass x^+ Basislösung von B^+ ist.

Sei $j \in \{1,\dots,n-m\}$ der Index mit $s_N < 0$ nach Gleichung 3.8 und weiter $l \in \{1,\dots,m\}$ ein Index mit $\widehat{A_{l,j}} > 0$ und

$$\sigma = \frac{x^*_{B,l}}{\widehat{A_{l,j}}} \tag{3.13}$$

Die neue Basis B^+ wird dann definiert durch:

$$\begin{aligned} J_{B^+} &:= (J_B \setminus \{b(l)\}) \cup \{n(j)\} & \qquad B^+ &= \{ A_{.i} \mid i \in J_{B^+} \} \\ J_{N^+} &:= (J_N \setminus \{n(j)\}) \cup \{b(l)\} & N^+ &= \{ A_{.i} \mid i \in J_{N^+} \} \end{aligned}$$

Damit x^+ Basislösung zu B^+ ist, muss gelten, $x^+_j = 0$, d. h. wir müssen zeigen, $x^+_{N,\mu} = 0, \forall \mu \in \{1,\dots,n-m\} \setminus \{j\}$ und $x^+_{B,l} = 0, \forall j \in J_{N^+}$. Wegen $v(\sigma) = \sigma d = x^+_N$ ist $x^+_{N,\mu} = 0$ und nach der Wahl von l gilt: $\sigma\widehat{A_{l,j}} = x^*_{B,l} \Rightarrow x^+_{B,l} = (\underbrace{A_B^{-1} b}_{=x^*_{B,l}} - \underbrace{A_B^{-1} A_N v(\sigma)}_{=\sigma\widehat{A_{l,j}}}) = x^*_{B,l} - \sigma\widehat{A_{l,j}} = 0$. Damit ist $x_{N^+} = 0_{n-m}$ und es ist noch zu zeigen, dass $A_{B^+} x^+_{B^+} = b$.

Wir wissen, dass $x \in \mathcal{F}$. Es gilt, $b = Ax^+ = \sum_{i=1}^n A_{.i} x^+_i = \sum_{i \in J_{B^+}} x^+_i A_{.i} = A_{B^+} x^+_{B^+}$. Somit bleibt noch zu zeigen, dass B^+ eine Basis von A ist.

Lemma 3.10
Sei $B = \{v^{(1)},\dots,v^{(m)}\} \subseteq \mathbb{R}^m$ eine Basis des $\mathbb{R}^m$, $l \in \{1,\dots,m\}$ ein Index, $w \in \mathbb{R}^m$ ein Vektor sowie V eine $m \times m$-Matrix mit den Spaltenvektoren $v^{(1)},\dots,v^{(m)}$. Bildet man B^+ aus B durch Basistausch des Vektors $v^{(l)}$ gegen w, also $B^+ = \{v^{(1)},\dots,v^{(l-1)}, w, v^{(l+1)},\dots,v^{(m)}\}$, dann gilt, dass B^+ genau dann Basis des $\mathbb{R}^m$ ist, wenn $\widehat{w_l} \neq 0$ mit $\widehat{w_l} := V^{-1}w$.

Beweis: 16
Sei V^+ die Matrix mit den Spaltenvektoren $v^{(1)},\dots,v^{(l-1)}, w, v^{(l+1)},\dots,v^{(m)}$. Wir zeigen, dass V^+ genau dann invertierbar ist, wenn $\widehat{w_l} \neq 0$ gilt. Wir wissen, V^+ ist genau dann invertierbar, wenn $V^{-1}V^+$ invertierbar ist und es gilt, $V^{-1}V^+ = (e^{(1)},\dots,e^{(l-1)}, \widehat{w_l}, e^{(l+1)},\dots,e^{(m)})$. ∎

Wir wenden nun das Lemma an und setzen $v^{(i)} = A_{.b(i)}$ sowie $w = A_{.n(j)}$. Dann ist $\widehat{w} = A_B^{-1} A_{.n(j)} = \widehat{A_{.j}}$. Nach Satz 3.10 ist B^+ genau dann eine Basis von A, wenn $\widehat{w} \neq 0 = \widehat{A_{l,j}}$. Nach der obigen Wahl von l ist $\widehat{A_{l,j}} > 0$ und es folgt, dass B^+ Basis ist.

Bemerkung 3.11
Die obigen Betrachtungen zeigen, $x^*_{n(j)} = 0$, da $x^*_N = 0_{n-m}$, $x^+_{n(j)} = \sigma$. Für σ gilt: $\sigma > 0$, falls die Ecke x^* nicht entartet ist und andernfalls ist $\sigma = 0$, d. h. in diesem Fall ist es möglich, dass nur der Basiswechsel stattfindet. Die aktuelle Ecke ändert sich dann nicht.

Wie berechnet sich nun der Wert der Zielfunktion in der neuen Ecke x^+? Nach der Gleichung 3.3 gilt:

$$c^T x^+ = z_B + s_N^T x_N^+ = z_B + s_N^T v(\sigma) = \underbrace{z_B}_{=c^T x^*} + \sigma s_{N,j} \tag{3.14}$$

Somit folgt, $c^T x^+ \leq z_B = c^T x^*$ und, falls x^* nicht entartet ist: $c^T x^+ < z_B = c^T x^*$

Satz 3.12
Sei x^* Ecke von $\mathcal{F}$ zur Basis B von A. Ist x^* nicht optimal und ist das Unbeschränktheitskriterium nicht erfüllt, dann ist die zur neuen Basis gehörige Basislösung x^+ zulässig und damit Ecke von $\mathcal{F}$ mit $c^T x^+ \leq c^T x^*$.

3.2. Durchführung des Simplexverfahrens

3.2.1. Verfahrenskonzept

Für die Phase 2 des Simplexverfahrens lässt sich folgendes Verfahren beschreiben: Gegeben seien

- eine Basis $B = B^{(0)}$ von A,
- die Matrizen $\overline{A} = A_B^{-1}$,
- $\widehat{A_B^{-1}} = A_B^{-1} A_N$,
- der Testvektor $x_B^{(0)}$
- und die Indexmengen J_B, J_N

Berechne

- den Testvektor $s_N^T = c_N^T - c_B^T A_B^{-1} A_N$ und
- den Zielfunktionswert $z_B = c_B^T x_B^{(0)}$
- Setze $k := 0$.

Danach werden folgende Iterationsschritte ausgeführt:

1. Optimalitätskriterium: Gilt $s_N \geq 0$, dann ist die aktuelle Ecke x^* optimal. Das Programm endet hier.

2. Unbeschränktheitskriterium: Wähle einen Index $j \in \{1, \dots, n-m\}$ mit $s_{N,j} < 0$. Gilt nun $\widehat{A.j} < 0$, ist die Zielfunktion nicht nach unten beschränkt und das Programm endet.
3. Setze $B^{(k+1)} := B := B^+$. Berechne mit der neuen Basis $\overline{A} = A_B^{-1}, \widehat{A_B^{-1}} A_N$, den Vektor s_N und z_B.
4. Setze $k := k+1$ und gehe zum ersten Punkt.

3.2.2. Konvergenz des Simplexverfahrens

Satz 3.13 (Konvergenz des Simplexverfahrens)
Sind alle Ecken $x^{(k)}$ zu den berechneten Basen $B^{(k)}$ nicht entartet, dann stoppt das Verfahren nach endlich vielen Schritten mit einer optimalen Basis oder das Unbeschränktheitskriterium ist erfüllt.

Beweis: 17
Sei $x^{(k)}$ die in der k-ten Iteration berechnete Ecke zur Basis $B^{(k)}$. Dabei sei weder das Optimalitätskriterium noch das Unbeschränktheitskriterium erfüllt. Nach dem Satz 3.12 folgt: Das Verfahren berechnet eine neue Basis $B^+ = B^{(k+1)}$, so dass für die zugehörige Ecke $x^{(k+1)}$ gilt:

$$c^T x^{(k+1)} < c^T x^{(k)}$$

d. h. $x^{(k+1)}$ ist nicht entartet. Da es nur endlich viele Ecken gibt, kann der Fall nur endlich oft auftreten.
■

3.2.3. Tableaudarstellung

	J_N		J_B
J_B	$\widehat{A} = A_B^{-1} A_N$	$x_B^* = A_B^{-1} b$	A_B^{-1}
	s_N^T	z_B	$c_B^T A_B^{-1}$

Beispiel 3.14

	2	3	4		1	5
1	$-3/4$	2	$-1/4$	3	$1/4$	0
5	$-1/4$	3	$-3/4$	5	0	$1/4$
	$17/2$	-23	$11/2$	26	$1/2$	1

Im Beispiel gilt $x_B^* > 0$, d. h. die aktuelle Ecke ist nicht entartet. Allerdings ist wegen $s_{N,2} < 0$ das Optimalitätskriterium nicht erfüllt. Also müssen wir im vorliegenden Fall $j = 2$ wählen. Für den Basistausch muss der Index l so gewählt werden, dass gilt:

$$\begin{aligned} \frac{x_{B,l}}{\widehat{A}_{l,j}} = \sigma &= \min\left\{ \frac{x^*}{\widehat{A}_{\nu,j}} \;\middle|\; \nu \in \{1, \dots, m\} \text{ mit } \widehat{A}_{\nu,j} > 0 \right\} \\ &= \min\left\{ \frac{x_{B,1}^*}{\widehat{A}_{1,j}}, \frac{x_{B,2}^*}{\widehat{A}_{2,j}} \right\} = \frac{3}{2} \Rightarrow l = 1 \end{aligned}$$

Beispiel 3.15
Führt man den Basistausch mit $j = 2, l = 1$ durch, so erhält man:

	2	1	4		3	5
3	$-3/8$	$1/2$	$-1/8$	$3/2$	$1/8$	0
5	$7/8$	$-3/2$	$-3/8$	$1/2$	$-3/8$	$1/4$
	$-1/8$	$23/2$	$21/8$	$-17/2$	$-19/8$	1

Auch hier gilt, $x^{(1)}$ ist nicht entartet, da $x_B^* > 0$. Das Optimalitätskriterium ist nicht erfüllt, da $s_{N,1} < 0$ und auch das Unbeschränktheitskriterium ist nicht erfüllt, da nicht die gesamte Spalte kleiner oder gleich Null ist.

Aus einem erneuten Basistausch mit $j = 1, l = 2$ resultiert das folgende Tableau:

	5	1	4		3	2
3	$3/7$	$-1/7$	$-2/7$	$12/7$	$-1/28$	$3/28$
2	$8/7$	$-12/7$	$-3/7$	$4/7$	$-3/7$	$2/7$
	$1/7$	$79/7$	$18/7$	$-60/7$	$-65/28$	$27/28$

Das Optimalitätskriterium ist erfüllt. Somit ist die aktuelle Ecke optimal.

$$x_B^* = \begin{pmatrix} x_2^* \\ x_3^* \end{pmatrix} = \begin{pmatrix} 4/7 \\ 12/7 \end{pmatrix} \qquad x_N^* = \begin{pmatrix} x_1^* \\ x_4^* \\ x_5^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Wir führen folgende, neue Bezeichnungen ein:

$$\widehat{A} = A_B^{-1} A_N \qquad \overline{A} = A_B^{-1} \qquad \overline{x} = x_B^*$$
$$\widehat{s} = s_N \qquad \overline{z} = z_B \qquad \overline{y} = A_B^{-T} c_B$$

Mit den neuen Bezeichnungen wählen wir zum Basistausch einen Index $\{1, \ldots, n-m\}$ mit $\widehat{s}_j < 0$ und einen Index $l \in \{1, \ldots, m\}$ mit $\frac{\overline{x}_l}{\widehat{A}_{l,j}} = \min\left\{ \frac{\overline{x}_\nu}{\widehat{A}_{\nu,j}} \;\middle|\; \nu \in \{1, \ldots, m\} \text{ mit } \widehat{A}_{\nu,j} > 0 \right\}$.

3.3. Der Austauschschritt

Die Spalten des Simplextableaus sind von der Form $A_B^{-1} v$. Zum Update des Tableaus müssen wir die entsprechenden Vektoren $A_{B^+}^{-1} v$ berechnen. Betrachten wir die allgemeine Situation von Satz 3.10:

Sei $B = \{v^{(1)}, \ldots, v^{(m)}\}$ eine Basis des $\mathbb{R}^m$. Weiter sei $l \in \{1, \ldots, m\}$ ein gegebener Index, $w \in \mathbb{R}^m$ und $V = [v^{(1)}, \ldots, v^{(m)}]$. Die Koordinaten $\widehat{w} = V^{-1} w$ von w bezüglich Basis B seien bekannt. Weiter sei $\widehat{w}_l \neq 0$. Tauscht man $v^{(l)}$ gegen w, d. h. definiert man

$$B^+ = \{v^{(1)}, \ldots, v^{(l-1)}, w, v^{(l+1)}, \ldots, v^{(m)}\}$$

nach Satz 3.10 ist B^+ wieder eine Basis.

Sei $V^+ = [v^{(1)}, \dots, v^{(l-1)}, w, v^{(l+1)}, \dots, v^{(m)}]$. Für beliebige $v \in \mathbb{R}^m$ sei $\widehat{v} = V^{-1}v$ und $v^+ = (V^+)^{-1}v$. Berechnen $w^+ = (V^+)^{-1}v^{(l)}$. Wir kennen $\widehat{w} = V^{-1}w \Leftrightarrow V\widehat{w} = w \Leftrightarrow w = \sum_{i=1}^m \widehat{w}_i v^{(i)} \Rightarrow \widehat{w}_l v^{(l)} = w - \sum_{i=1, i\neq l}^m \widehat{w}_i v^{(i)}$. Da $\widehat{w}_l \neq 0$ folgt, $v^{(l)} = \frac{w}{\widehat{w}_l} - \sum_{i=1,i\neq l}^m \frac{\widehat{w}_i}{\widehat{w}_l} v^{(i)}$. Dann folgt:

$$v^{(l)} = V^+ w^+ \text{ mit } w^+ = \begin{cases} \frac{1}{\widehat{w}_l} & i = l \\ -\frac{\widehat{w}_i}{\widehat{w}_l} & i \in \{1, \dots, m\} \setminus \{l\} \end{cases} \tag{3.15}$$

Nachdem man w^+ entsprechend Gleichung 3.15 berechnet hat, kann man für beliebigen Vektor $v \in \mathbb{R}^n$ jetzt aus $\widehat{v} = V^{-1}v$ einfach $v^+ = (V^+)^{-1}v$ berechnen. Es gilt:

$$\widehat{v} = V^{-1}v \Leftrightarrow v = V\widehat{v} \Leftrightarrow v = \sum_{i=1}^m \widehat{V}_i v^{(i)}$$

Für $v^{(l)}$ setzen wir die Darstellung $v^{(l)} = V^+ w^+$ ein:

$$\begin{aligned} v &= \sum_{i=1, i\neq l}^m \widehat{v}_i v^{(i)} + \widehat{v}_l v^{(l)} = \sum_{i=1,i\neq l}^m \widehat{v}_i v^{(i)} + \widehat{v}_l V^+ w^+ \\ &= \sum_{i=1,i\neq l}^m \widehat{v}_i v^{(i)} + \widehat{v}_l \sum_{i=1,i\neq l}^m w_i^+ v^{(i)} + \widehat{v}_l w_l^+ w \\ &= \sum_{i=1,i\neq l}^m (\widehat{v}_i + \widehat{v}_l w_i^+) v^{(i)} + \widehat{v}_l w_l^+ w \end{aligned}$$

zwei Tabellen einfügen

Allgemeine Situation beim Basistausch:
aktuelle Basis: $B = \{v^{(1)}, \dots, v^{(l-1)}, v^{(l)}, v^{(l+1)}, \dots, v^{(m)}\}$
neue Basis: $B^+ = \{v^{(1)}, \dots, v^{(l-1)}, w, v^{(l+1)}, \dots, v^{(m)}\}$ mit $w \in \mathbb{R}^m, \widehat{w}_l \neq 0$, wobei $\widehat{w} = V^{-1}w$

Seien $w^+ := (V^+)^{-1}v^{(l)}$ Koordinaten von $v^{(l)}$ (des aus der aktuellen Basis gestrichenen Vektors) bezüglich der neuen Basis B^+

$$w_i^+ = \begin{cases} \frac{1}{\widehat{w}_l} & i = l \\ -\frac{\widehat{w}_i}{\widehat{w}_l} & i \neq l \end{cases} \tag{3.16}$$

Sei allgemein $v \in \mathbb{R}^m$ ein beliebiger Vektor, $\widehat{v} := A_B^{-1}v$ (Koordinaten von v bezüglich der aktuellen Basis B), $v^+ := A_{B^+}^{-1}v$ (Koordinaten von v bezüglich der neuen Basis B^+).

$$v_i^+ = \begin{cases} \widehat{v}_l \widehat{w}_l & i = l \\ \widehat{v}_i + \widehat{v}_l \widehat{w}_i & i \neq l \end{cases} \tag{3.17}$$

Anwendung auf das Simplextableau:

- Umrechnung der Pivotspalte nach Gleichung 3.16
- Umrechnung der restlichen Spalten von $\widehat{A}$ nach Gleichung 3.17
- Umrechnung von $\overline{x}$ und $\overline{A}$ nach Gleichung 3.17
- Umrechnung von $\widehat{s}, \overline{z}, \overline{y}$

3.3.1. Neuberechnung des Simplextableaus

Indextausch: $b^+(l) = n(j), n^+(j) = b(l)$, restliche Indizes bleiben unverändert

Neuberechnung der Pivotspalte: Der Vektor der aktuellen Pivotspalte ist $\widehat{w} = \widehat{A}_{.j}$ und $\widehat{w}$ wird durch den Vektor $w^+ = \widehat{A}^+_{.j}$ ersetzt (Koordinaten des aus der Basis entfernten Vektors $v^{(l)} = A_{b(l)}$ bezüglich der neuen Basis B^+), d. h. wir können Gleichung 3.16 anwenden mit $\widehat{w} = \widehat{A}_{.j}$ (aktuelle Pivotspalte) und wir erhalten

$$\widehat{A}^+_{i,j} = \begin{cases} \frac{1}{\widehat{A}_{i,j}} & i = l \\ -\frac{\widehat{A}_{i,j}}{\widehat{A}_{l,j}} & i \neq l \end{cases} \tag{3.18}$$

Beispiel 3.16 (Problem aus Satz 3.14)
Mit Gleichung 3.18 erhalten wir als neue Spalte:

$$\widehat{A}^+_{1,2} = \frac{1}{\widehat{A}_{1,2}} = \frac{1}{2} \qquad \widehat{A}^+_{2,2} = -\frac{\widehat{A}_{2,2}}{\widehat{A}_{1,2}} = -\frac{3}{2}$$

Berechnung von $\widehat{A}^+$ (ohne Pivotspalte)

Wir benutzen Gleichung 3.17, wobei $w^+ = \widehat{A}^+_{.j}$ die bereits neu berechnete Pivotspalte ist. Weiter ist $\widehat{v} = \widehat{A}_{.\nu}, \nu \in \{1, \dots, n-m\}, \nu \neq j$.

$$\widehat{A}^+_{i,\nu} = \begin{cases} \widehat{A}_{l,\nu}\widehat{A}^+_{l,j} & i = l \\ \widehat{A}_{i,\nu} + \widehat{A}_{l,\nu}\widehat{A}^+_{i,j} & i \neq l \end{cases} \tag{3.19}$$

Beispiel 3.17 (Fortsetzung von Satz 3.16)
Zur Neuberechnung der Spalten 1 und 3 von $\widehat{A}$ erhalten wir aus Gleichung 3.19 mit $\nu = 1, 3$:

$$\begin{aligned}
\widehat{A}^+_{1,1} &= \widehat{A}_{1,1}\widehat{A}^+_{1,2} = -\frac{3}{4} \cdot \frac{1}{2} = -\frac{3}{8} \\
\widehat{A}^+_{1,3} &= \widehat{A}_{1,3}\widehat{A}^+_{1,2} = -\frac{1}{8} \\
\widehat{A}^+_{2,1} &= \widehat{A}_{2,1} + \widehat{A}_{1,1}\widehat{A}^+_{2,2} = \frac{7}{8} \\
\widehat{A}^+_{2,3} &= -\frac{3}{8}
\end{aligned}$$

neues Tableau: Tableau e

Berechnung von $\overline{x}^+$: Anwendung von Gleichung 3.18 mit $w^+ := \widehat{A}^+_{\cdot j}$ und $\widehat{v} = \overline{x}$

(3.20) $$\overline{x}^+_i = \begin{cases} \overline{x}_l \widehat{A}^+_{l,j} & i = l \\ \overline{x}_i + \overline{x}_l \widehat{A}^+_{i,j} & i \neq l \end{cases}$$

Bemerkung 3.18
Zusammen mit Gleichung 3.17 und Gleichung 3.13 erhalten wir:

$$x^+_l = \overline{x}_l \widehat{A}^+_{l,j} = \frac{\overline{x}_l}{\widehat{A}_{l,j}} = \sigma$$

Für $i \neq l$ gilt:

$$x^+_i = \overline{x}_i + \overline{x}_l \widehat{A}^+_{i,j} = \overline{x}_i - \overline{x}_l \frac{\widehat{A}_{i,j}}{\widehat{A}_{l,j}} = \overline{x}_i - \sigma \widehat{A}_{i,j}$$

Beispiel 3.19 (Fortsetzung von Gleichung 3.17)
$\overline{x}^+_1 = \overline{x}_1 \widehat{A}^+_{1,2} = 3/2, \overline{x}^+_2 = \overline{x}_2 + \overline{x}_1 \widehat{A}^+_{2,2} = 1/2$

Berechnung von $\overline{A}^+$: Analog zur Berechnung von $\overline{x}^+$ und $\widehat{A}^+$ erhält man für $\mu = 1, \dots, m$:

(3.21) $$\overline{A}^+_{i,\mu} = \begin{cases} \overline{A}_{l,\mu} \widehat{A}^+_{l,j} & i = l \\ \overline{A}_{i,\mu} + \overline{A}_{l,\mu} \widehat{A}^+_{i,j} & i \neq l \end{cases}$$

Beispiel 3.20 (Fortsetzung von Satz 3.19)
siehe voriges Tableau, Zahlen sind dort mit ergänzt.

Berechnung von $\overline{y}^+$: Nach der Definition ist $\overline{y}^+ = ((A_{B^+})^{-1})^+ c_{B^+} = (\overline{A}^+)^{-1} c_{B^+}$. Mit Gleichung 3.21 erhalten wir für $\mu = 1, \dots, m$:

$$\begin{aligned}
\overline{y}^+_\mu &= \sum_{i=1, i\neq l}^{m} c_{b(i)} \overline{A}^+_{i,\mu} + c_{n(j)} \overline{A}^+_{l,\mu} \\
&= \sum_{i=1, i\neq l}^{m} c_{b(i)} (\overline{A}_{i,\mu} + \overline{A}_{l,\mu} \widehat{A}^+_{i,j}) + c_{n(j)} \overline{A}_{l,\mu} \widehat{A}^+_{l,j} \\
&= \sum_{i=1, i\neq l}^{m} c_{b(i)} \overline{A}_{i,\mu} + c_{b(l)} \overline{A}_{l,\mu} + \sum_{i=1, i\neq l}^{m} c_{b(i)} \overline{A}_{l,\mu} \widehat{A}^+_{i,j} - c_{b(l)} \overline{A}_{l,\mu} + c_{n(j)} \overline{A}_{l,\mu} \widehat{A}^+_{l,j}
\end{aligned}$$

Zusammen mit $\sum_{i=1}^{m} c_{b(i)}\overline{A}_{i,\mu} = \overline{y}_\mu$ und Einsetzen von Gleichung 3.18 folgt:

$$\begin{aligned}
\overline{y}_\mu^+ &= \overline{y}_\mu - \sum_{i=1}^{m} c_{b(i)}\overline{A}_{l,\mu}\frac{\widehat{A}_{i,j}}{\widehat{A}_{l,j}} + c_{n(j)}\overline{A}_{l,\mu}\frac{1}{\widehat{A}_{l,j}} \\
&= \overline{y}_\mu + \frac{\overline{A}_{l,\mu}}{\widehat{A}_{l,j}}\underbrace{\left(c_{n(j)} - \sum_{i=1}^{m} c_{b(i)}\widehat{A}_{i,j}\right)}_{\widehat{s}_j}
\end{aligned}$$

Daraus folgt:

$$\overline{y}_\mu^+ = \overline{y}_\mu + \widehat{s}_j\frac{\overline{A}_{l,\mu}}{\widehat{A}_{l,j}} = \overline{y}_\mu + \widehat{s}_j\overline{A}_{l,\mu}\widehat{A}_{l,j} \tag{3.22}$$

$$\widehat{s}_\nu^+ = \begin{cases} -\widehat{s}_j\widehat{A}_{l,j}^+ & \nu = j \\ \widehat{s}_\nu - \widehat{s}_j\widehat{A}_{l,j}^+\widehat{A}_{l,\nu} & \nu \neq j \end{cases} \tag{3.23}$$

$$\overline{z}^+ = \overline{z} + \widehat{s}_j\overline{x}_l^+ \tag{3.24}$$

3.3.2. Implementierung des Austauschschritts

Sei ein Tableau mit den Daten $b, n, \widehat{A}, \overline{A}, \overline{x}, \widehat{s}, \overline{y}, \overline{z}$ gegeben. Weiterhin seien $j \in \{1, \ldots, n-m\}$ und $l \in \{1, \ldots, m\}$ die Indizes für den Basistausch.

1. Neuberechnung der Pivotspalte nach Gleichung 3.18
2. Neuberechnung von $\overline{x}$ nach Gleichung 3.20
3. Neuberechnung von $\overline{z}$ mit Gleichung 3.24, $\overline{y}$ nach Gleichung 3.22 und $\widehat{s}$ nach Gleichung 3.23
4. Neuberechnung der restlichen Spalten von $\widehat{A}$ nach Gleichung 3.19
5. Neuberechnung von $\overline{A}$ nach Gleichung 3.21
6. Indextausch

3.3.3. Die lexikografische Zusatzregel

Definition 3.21 (Lexikografisch positive Vektoren)
Ein Vektor $x \in \mathbb{R}^n \setminus \{0_n\}$ heißt **lexikografisch positiv**, wofür wir $x \succ 0_n$ bzw. $0_n \prec x$ schreiben, wenn die erste von Null verschieden Komponente größer als Null ist.
Sind $x, y \in \mathbb{R}^n$, dann heißt x lexikografisch größer y, falls $x - y \succ 0$ ist. Analog heißt x lexikografisch kleiner y, falls $x - y \prec 0$ ist.

Beispiel 3.22
Die Vektoren $x = (0, 1, 2)^T, y = (0, 2, -1)^T$ sind lexikografisch positiv. Weiter gilt $x - y = (0, -1, 3)^T \prec 0, y - x = (0, 1, -3)^T \succ 0 \Rightarrow y \prec x$.

Bemerkung 3.23
Zwei Vektoren $x, y \in \mathbb{R}^n$ sind bezüglich der lexikografischen Ordnung immer vergleichbar, d. h. $x \neq y \Rightarrow x \succ y \wedge x \prec y$.

Sei $y \in \mathbb{R}^n$ fest und $K = \{ x \in \mathbb{R} \mid x \succeq y \}$. Dann ist K nicht immer abgeschlossen. Dazu betrachten wir einen Iterationsschritt des Simplexverfahrens. Sei B eine aktuelle Basis, $j \in \{1, \ldots, n-m\}$ mit $s_{N,j} < 0, \sigma = \min\left\{ \frac{\overline{x}_\nu}{\widehat{A}_{\nu,l}} \,\middle|\, \nu \in \{1, \ldots, m\} \text{ mit } \widehat{A}_{\nu,l} > 0 \right\}, J = \left\{ \mu \in \{1, \ldots, m\} \,\middle|\, \widehat{A}_{\mu,j} > 0, \sigma = \frac{\overline{x}_\mu}{\widehat{A}_{\mu,j}} \right\}$

Mit $\overline{A}_i$ bezeichnen wir den i-ten Zeilenvektor von $\overline{A} = A_B^{-1}$. Für die Bestimmung von l benutzen wir die lexikografische Zusatzregel. Wähle $l \in J$ so, dass $\frac{1}{\widehat{A}_{l,j}}\overline{A}_l \prec \frac{1}{\widehat{A}_{\mu,l}}\overline{A}_\mu$ für alle $\mu \in J \setminus \{l\}$.

Bemerkung 3.24
Hat J mehr als ein Element, dann sind die Vektoren $\frac{\overline{A}_\mu}{\widehat{A}_{\mu,j}}, \mu \in J$ paarweise verschieden. Damit ist l durch die lexikografische Zusatzregel eindeutig bestimmt.

Definieren $\overline{J} = \left\{ \nu \in \{1, \ldots, m\} \,\middle|\, \widehat{A}_{\nu,j} > 0 \right\} \Rightarrow J \leq \overline{J}$. Nach der Definition von J und $\overline{J}$ gilt für $\mu \in J$: $\sigma = \frac{\overline{x}_\mu}{\widehat{A}_{\mu,j}}$. Betrachten $\frac{\overline{x}_\nu}{\widehat{A}_{\nu,j}}$ für alle $\nu \in \overline{J} \setminus J \Rightarrow \sigma < \frac{\overline{x}_\nu}{\widehat{A}_{\nu,j}}$. Damit erhält man die äquivalente Form der lexikografischen Zusatzregel: Wähle $l \in \overline{J}$ so, dass gilt:

$$\frac{1}{\widehat{A}_{l,j}}\begin{pmatrix}\overline{x}_l \\ \overline{A}_l\end{pmatrix} \prec \frac{1}{\widehat{A}_{\nu,j}}\begin{pmatrix}\overline{x}_\nu \\ \overline{A}_\nu\end{pmatrix} \qquad \nu \in \overline{J} \setminus \{l\}$$

Wir definieren Q durch $Q := (\overline{x}, \overline{A})$ und $q := \begin{pmatrix}\overline{z} \\ \overline{y}\end{pmatrix}$.

Beispiel 3.25
In Satz 3.14 ist $Q = \begin{pmatrix}3 & | & 1/4 & 0 \\ 5 & | & 0 & 1/4\end{pmatrix}$. Die Zeilenvektoren von Q sind lexikografisch positiv. Mit der neuen Basis B^+ nach dem Basistausch sei $Q^+ := (\overline{x}^+, \overline{A}^+)$ und $q^+ = \begin{pmatrix}\overline{z}^+ \\ \overline{y}^+\end{pmatrix}$.

Lemma 3.26
Wird l für den Basistausch nach der Zusatzregel und sind die Zeilenvektoren von Q lexikografisch positiv, dann gilt:

a) Zeilenvektoren von Q^+ sind lexikografisch positiv

b) $q^+ \prec q$

Beweis: 18
Nach der Voraussetzung sind die Zeilenvektoren Q lexikografisch positiv:

$$\begin{pmatrix} \overline{x}_i \\ A_i \end{pmatrix} \succ 0 \qquad i = 1, \dots, m$$

Für $i = l$ gilt nach Gleichung 3.20 und Gleichung 3.21:

$$\begin{pmatrix} \overline{x}_l^+ \\ A_l^+ \end{pmatrix} = \underbrace{\frac{1}{\widehat{A}_{l,j}}}_{>0} \begin{pmatrix} \overline{x}_l \\ A_l \end{pmatrix} \succ 0$$

Für $i \in \{1, \dots, m\} \setminus \{l\}$ gilt

$$\begin{pmatrix} \overline{x}_i^+ \\ A_i^+ \end{pmatrix} = \begin{pmatrix} \overline{x}_i \\ A_i \end{pmatrix} + \widehat{A}_{i,j}^+ \begin{pmatrix} \overline{x}_l \\ A_l \end{pmatrix} = \begin{pmatrix} \overline{x}_i \\ A_i \end{pmatrix} - \frac{\widehat{A}_{i,j}}{\widehat{A}_{l,j}} \begin{pmatrix} \overline{x}_l \\ A_l \end{pmatrix} \text{ mit } \widehat{A}_{i,j}^+ = -\underbrace{\frac{\widehat{A}_{i,j}}{\widehat{A}_{l,j}}}_{>0}$$

Für $i \notin \overline{J}$ gilt $\widehat{A}_{i,j} \leq 0 \Rightarrow \widehat{A}_{i,j}^+ \geq 0 \Rightarrow \begin{pmatrix} \overline{x}_i^+ \\ A_i^+ \end{pmatrix} \succ 0$ und für $i \in \overline{J}$ schreiben wir die obige Formel um:

$$\begin{pmatrix} \overline{x}_i^+ \\ A_i^+ \end{pmatrix} = \underbrace{\widehat{A}_{i,j}}_{>0} \underbrace{\left(\frac{1}{\widehat{A}_{i,j}} \begin{pmatrix} \overline{x}_i \\ A_i \end{pmatrix} - \frac{1}{\widehat{A}_{l,j}} \begin{pmatrix} \overline{x}_l \\ A_l \end{pmatrix} \right)}_{>0} \succ 0$$

Der Punkt b) folgt aus Gleichung 3.24 und Gleichung 3.22: $q^+ = q + \underbrace{\widehat{s}_j}_{<0} \underbrace{\widehat{A}_{l,j} \begin{pmatrix} \overline{x}_l \\ A_l \end{pmatrix}}_{>0} \Rightarrow q^+ \prec q$ ∎

Satz 3.27
Die Zeilenvektoren der zur Basis $B^{(0)}$ gehörenden Matrix $Q^{(0)}$ seien lexikografisch positiv. Wird l für den Basistausch in jedem Iterationsschritt nach der Zusatzregel gewonnen, dann stoppt das Verfahren nach endlich vielen Schritten, wobei das Optimalitätskriterium oder das Unbeschränktheitskriterium erfüllt sind.

Beweis: 19
Seien $B^{(k)}$ Basen für $k = 1, 2, 3, \dots$ und $Q^{(k)}, q^{(k)}$ die zugehörigen Matrizen Q und Vektoren q. Dann gilt nach Satz 3.26 $q^{(k+1)} \prec q^{(k)}$ für alle k. Nehmen wir an, es gelte $B^{(k+\nu)} = B^{(k)}$ mit einem $\nu > 0$. Nach der Definition von q bzw. $\overline{z}, \overline{y}$ müsste dann gelten: $q^{(k+\nu)} = q^{(k)}$. Dies ist allerdings ein Widerspruch. Deswegen sind alle Basen verschieden. Da es höchstens $\binom{n}{m}$ Basen gibt, stoppt das Verfahren nach endlich vielen Schritten. ∎

Bemerkung 3.28
Sei ein lineares Programm $\min c^T x$ unter den Nebenbedingungen $Ax \leq b, x \geq 0$ mit $b \geq 0$ gegeben. Durch die Einführung von Schlupfvariablen $y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ erhalten wir $\min(c^T, 0_m) \begin{pmatrix} x \\ y \end{pmatrix}$ unter

den Nebenbedingungen $\underbrace{(A, I_m)}_{\tilde{A}} \begin{pmatrix} x \\ y \end{pmatrix} = b, \begin{pmatrix} x \\ y \end{pmatrix} \geq 0$. Startbasis sind die letzten m Spaltenvektoren von $\tilde{A}$. Dann ist $\overline{x} = b$ und $Q = \begin{pmatrix} b_1 & | & 1 & \dots & 0 \\ \vdots & | & \vdots & \ddots & \vdots \\ b_m & | & 0 & \dots & 1 \end{pmatrix}$.

Tabelle einfügen

Beispiel 3.29
Mit der Zusatzregel ergibt sich $\frac{1}{4}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} < \frac{1}{4}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow l = 3$.

3.4. Phase 1 des Simplexverfahrens

Ausgangspunkt ist ein lineares Programm in Standardform. Die Phase 1 löst das Hilfsproblem (HP) und man erhält aus der Lösung des Hilfsproblems

- die Startbasis für die Phase 2 oder
- die Information, dass die zulässige Menge leer ist oder
- die Information, dass der Rang von A kleiner als m ist (redundante Gleichungen)

Das Hilfsproblem

Wir bezeichnen den Einheitsvektor mit $e = (1, 1, \dots, 1)^T$ und die Variablen des Hilfsproblems mit $z = (z_1, \dots, z_n, z_{n+1}, \dots, z_{n+m})^T = (x_1, \dots, x_n, y_1, \dots, y_m)^T$. Das Hilfsproblem lautet

$$\min e^T y = \sum_{i=1}^{m} y_i$$

unter den Nebenbedingungen

$$\begin{pmatrix} x \\ y \end{pmatrix} \in \tilde{\mathcal{F}} := \left\{ z = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+m} \,\middle|\, \begin{pmatrix} x \\ y \end{pmatrix} \geq 0, (A \ I_m) \begin{pmatrix} x \\ y \end{pmatrix} = b \right\} = \left\{ z = \begin{pmatrix} x \\ y \end{pmatrix} \,\middle|\, x, y \geq 0, Ax + y = b \right\}.$$

Ist $\begin{pmatrix} x \\ y \end{pmatrix} \in \tilde{\mathcal{F}}$ und $y = 0_m \Rightarrow Ax = b$, d. h. $x \in \mathcal{F}$. Wir schreiben das Hilfsproblem als lineares Programm in Standardform: Definieren $\tilde{c} = \begin{pmatrix} 0_n \\ e \end{pmatrix}, \tilde{A} = (A, I_m)$ und es ist $\min \tilde{c}^T z$ unter den Nebenbedingungen $z \geq 0, \tilde{A}z = b$. Sei o. B. d. A. $b \geq 0$. Es gilt, $\mathcal{F} \neq \emptyset$. Die Zielfunktion ist durch 0 nach unten beschränkt. Nach dem Hauptsatz der linearen Optimierung existiert eine Ecke, die Lösung ist.

Nun wenden wir die Phase 2 des Simplexverfahrens auf das Hilfsproblem an. Wir wählen die letzten m Spalten von $\tilde{A}$ als Basis. Dann ist:

$$J_B = \{n+1, \dots, n+m\} \qquad J_N = \{1, \dots, n\} \qquad \tilde{A}_B = I_m \qquad \tilde{A}_N = A$$

Die zugehörige Basislösung ist $z_B = \underbrace{\tilde{A}_B^{-1}}_{=I_m} b = b, z_N = 0_n$. Wegen $b \geq 0$ ist $z_B \geq 0 \Rightarrow$ zulässige Basislösung. Einträge für das Starttableau:

$$\begin{aligned}
\tilde{A}_B^{-1} &= I_m \\
\tilde{A}_B^{-1}\tilde{A}_N &= \tilde{A}_N = A \\
\widehat{s}^T &= \tilde{c}_N^T - \tilde{c}_B^T\tilde{A}_B^{-1}\tilde{A}_N = 0_n - e^T I_m A = -e^T A \\
\tilde{c}^T z &= \tilde{c}_B^T z_B = e^T b \\
y^T &= \tilde{c}_B^T \tilde{A}_B^{-1} = e^T
\end{aligned}$$

Das Simplexverfahren mit der Zusatzregel berechnet also nach endlich vielen Schritten eine Ecke für folgendes Starttableau:

	$(1 \dots n)$		$(n+1 \dots n+m)$
$\begin{pmatrix} n+1 \\ \dots \\ n+m \end{pmatrix}$	A	b	I_m
	$-e^T A$	$e^T b$	e^T

Lemma 3.30
Sei $\tilde{z} = \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix}$ eine Lösung des Hilfsproblems. Ist der Optimalwert $\tilde{c}^T\tilde{z} = 0$, dann ist $\tilde{x} \in \mathcal{F}$, also $\mathcal{F} \neq \emptyset$ und $\tilde{y} = 0_m$. Ist der Optimalwert $\tilde{c}^T\tilde{z} > 0$, dann ist $\mathcal{F} = \emptyset$.

Beweis: 20
Sei $\tilde{c}^T\tilde{z} = \sum_{i=1}^m \tilde{y}_i = 0$. Wegen $\tilde{y} \geq 0 \Rightarrow \tilde{y} = 0_m \Rightarrow b = \tilde{A}\tilde{z} = (A, I_m)\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = A\tilde{x} + \tilde{y} = A\tilde{x}, \tilde{x} \geq 0 \Rightarrow$ $\tilde{x} \in \mathcal{F}$, d. h. $\mathcal{F} \neq \emptyset$.

Sei $\mathcal{F} \neq \emptyset$ und $x \in \mathcal{F}$. Dann ist $z = \begin{pmatrix} x \\ 0_m \end{pmatrix} \in \tilde{\mathcal{F}}$. Denn $z \geq 0$ und $\tilde{A}z = Ax + 0_m = b$. Damit ist der Optimalwert Null. ∎

Das Lemma sagt uns: Nach Anwendung der Phase 2 des Simplexverfahrens auf das Hilfsproblem gibt es zwei Möglichkeiten. Für die berechnete Lösung gilt:

1. $\tilde{c}^T\tilde{z} > 0$: *Stoppe* das Verfahren, denn $\mathcal{F} = \emptyset$.
2. $\tilde{c}^T\tilde{z} = 0$: (Damit ist auch $\tilde{y} = 0_m$). Dies ist noch genauer zu untersuchen.

Beispiel 3.31

$$\min 2x_1 + 6x_2 - 7x_3 + 2x_4 + 4x_5$$

unter den Nebenbedingungen

$$\begin{aligned} 4x_1 - 3x_2 + 8x_3 - x_4 &= 12 \\ -x_2 + 12x_3 - 3x_4 + 4x_5 &= 20 \\ x_i \geq 0, i = 1, \dots, 5 \end{aligned}$$

Das Hilfsproblem lautet wie folgt:

$$\min y_1 + y_2$$

unter den Nebenbedingungen

$$\begin{aligned} 4x_1 - 3x_2 + 8x_3 - x_4 &= 12 \\ -x_2 + 12x_3 - 3x_4 4x_5 &= 20 \\ x_i \geq 0, i = 1, \dots, 5, y_1, y_2 \geq 0 \end{aligned}$$

Tableau einfügen

Die Lösung ist:

$$\tilde{z} = \begin{pmatrix} \\ \\ \tilde{x} \\ \\ \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3/2 \\ 0 \\ 1/2 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \tilde{y} = 0_m$$

Es gilt: $\tilde{x} \in \mathcal{F}$. Alle Basisindizes sind kleiner als $n = 5$. Damit ist $\tilde{x}$ Basislösung für das Anfangsproblem ($\tilde{x}$ ist Ecke für das Anfangsproblem.).

Sei $\tilde{z} = \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix}$ die berechnete Lösung des Hilfsproblems. Es gelte: $\tilde{c}^T \tilde{z} = 0, \tilde{y} = 0_m$. Weiter ist:

$$B = \{\tilde{A}_{.b(1)}, \dots, \tilde{A}_{.b(n)}\} N = \{\tilde{A}_{.n(1)}, \dots, \tilde{A}_{.n(n)}\}$$

Tableau einfügen

Letztes Tableau der Phase 1: Dann ist $x \in \mathcal{F}$.

Im schönen Fall gilt, $b(i) \leq n, i = 1, \dots, m.$, d. h. zur Basis gehören nur Spaltenvektoren von A. Dann folgt, dass auch B Basis von A ist und $\tilde{x}$ Basislösung zur Basis B ist. Somit kann B auch als Startbasis für die Phase 2 genutzt werden.

Sei $l \in \{1, \dots, m\}$ ein Index mit $b(l) = \max_{1 \leq i \leq m} b(i)$. Wir unterscheiden zwei Fälle:

1. Es ist $b(l) \leq n$, d. h. alle Basisindizes sind $\leq n \Rightarrow B \subset \{A_{\cdot 1}, \ldots, A_{\cdot n}\} \Rightarrow \tilde{x}$ ist auch Basislösung zu $Ax = b$. Somit ist $\tilde{x}$ zulässige Basislösung für das Ausgangsproblem. Daher können wir mit der Basis B von A die Phase 2 starten. Der obere Teil des Starttableaus entsteht aus dem Endtableau der Phase 1 durch das Streichen derjenigen Spalten, die zu Indizes $> n$ gehören.
2. Es ist $b(l) > n$, d. h. einige Hilfsvariablen sind Basisvariablen.
   a) Es gibt einen Index $j \in \{1, \ldots, n\}$ mit $n(j) \in \{1, \ldots, n\}$ (d. h. die Variable $x_{n(j)}$ ist keine Basisvariable)

   $$\tilde{z} = \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = (\tilde{x}_1, \ldots, \tilde{x}_n, \tilde{y}_1, \ldots, \tilde{y}_m)^T$$

   $b(l) > n \Rightarrow \tilde{y}_{b(l)-n}$ ist Basisvariable und für den j-ten Spaltenvektor der Matrix $\tilde{A} := \tilde{A}_B^{-1}\tilde{A}_N$ gilt: $\widehat{A_{l,j}} \neq 0$.

   In dem Fall können wir nach Satz 3.10 einen Basistausch $b(l) \Leftrightarrow n(j)$ durchführen und wir erhalten eine neue Basis. Nach endlich vielen Schritten sind keine Hilfsvariablen mehr Basisvariablen. Dann gehen wir analog zu Fall 1 vor.

   b) Für alle Indizes $j \in \{1, \ldots, n\}$ mit $n(j) \in \{1, \ldots, n\}$ ist $\widehat{A_{l,j}} = 0$. Wegen $b(l) > n$ gibt es ein $i \in \{1, \ldots, m\}$ mit $b(l) = n + 1 \Leftrightarrow i = b(l) - n$. Dann kann man zeigen: Die i-te Zeile von A ist Linearkombination der restlichen Zahlen, d. h. $\mathrm{rg}(A) < m$ bzw. die i-te Gleichung ist redundant und kann daher gestrichen werden.

Beispiel 3.32

$$\min 2x_1 + 3x_2 + x_3$$

unter den Nebenbedingungen

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

Hilfsproblem für Phase 1
Das Problem aus Satz 3.29 war:

Tabelle einfügen

Wir tauschen $x_5 = y_1$ gegen x_4:

Tabelle einfügen

Das Streichen ergibt das obere Starttableau für die Phase 2. Die letzte Zeile muss neu berechnet werden.

Verfahren 3.33 (Zusammenfassung der Phase 1)

1. Berechne eine Lösung $\tilde{e} = \begin{pmatrix} x \\ \tilde{y} \end{pmatrix}$ vom Hilfsproblem mit der Phase 2. Ist $\tilde{x}^T\tilde{z} > 0 \Rightarrow \mathcal{F} = \emptyset$ und das Verfahren stoppt.

2. Berechne $l \in \{1, \ldots, m\}$ mit $b(l) = \max_{1 \le i \le m} b(i)$.

3. Ist $b(l) \le n$, dann streiche die zu den Hilfsvariablen gehörigen Spalten im linken Tableauteil, berechne die letzte Tableauzeile neu und starte Phase 2.

4. Ist $b(l) > n$, dann führe den Basistausch $b(l) \Leftrightarrow n(j)$ durch, wenn es ein $j \in \{1, \ldots, n\}$ mit $n(j) \in \{1, \ldots, n\}$ und $\widehat{A}_{l,j} \ne \emptyset$ gibt.

5. Falls $\mathrm{rg}(A) < m$, berechne $i = b(l) - n$ und streiche die i-te Gleichungsrestriktion.

Bemerkung 3.34
Gegeben sei ein lineares Problem vom Typ $\min c^T x$ unter den Nebenbedingungen $Ax \le b, x \ge 0$. Wenn wir das in Standardform transformieren, erhalten wir $\min(c^T, 0, \ldots, 0)\begin{pmatrix} x \\ y \end{pmatrix}$ unter den Nebenbedingungen $(A, I_m)\begin{pmatrix} x \\ y \end{pmatrix} = b, x, y \ge 0$. Naheliegend ist es, $J_B = \{n+1, \ldots, n+m\}$ zu wählen. Dies geht nur, wenn $b \ge 0$ ist.

3.5. Das revidierte Simplex-Verfahren

Dieses Verfahren benutzt anstelle von $\widehat{A}$ nur die Pivotspalte $\widehat{w}$ und spart damit Speicherplatz. Es erhebt sich die Frage, wie man dann den Basiswechsel durchführen kann.

Wir wählen einen Index $j \in \{1, \ldots, n-m\}$ mit $\widehat{s}_j < 0$. Wir berechnen die Pivotspalte $\widehat{w} = A_B^{-1} A_{.n(j)}$. Falls $\widehat{w} \le 0$ ist das Unbeschränktheitskriterium erfüllt und das Verfahren stoppt.

Als nächsten wählen wir einen Index $l \in \{1, \ldots, m\}$ mit $\sigma = \frac{\overline{x}_l}{\widehat{w}_l} = \min\{\frac{\overline{x}_\mu}{\widehat{w}_\mu} : 1 \le \mu \le m, \widehat{w}_\mu > 0\}$. Benutzt man beim reduzierten Problem die Suchrichtung d mit $d_j = 1, d_i = 0, i \ne j$, dann ist δ die maximale Schrittweite mit $v(t) = td \in \mathcal{F}_B$.

Die neue Ecke wird definiert durch

$$\begin{aligned} x_N^+ &= v^+ = \delta d \\ x_B^+ &= x_B^* - \delta \widehat{w} \end{aligned}$$

Somit erhalten wir für x^+ die einfache Updateformel:

$$\begin{aligned} x_N^+ &= x_B^* - \delta \widehat{w} \\ x_{n(j)}^+ &= \delta \end{aligned}$$

Für den Vektor $\overline{x} = x_B^*$ bedeutet das:

$$\overline{x}_i^+ = \begin{cases} \bar{l} - \delta \widehat{w}_i & i \ne l \\ \delta & i = l \end{cases} \tag{3.25}$$

Fasst man die Updateformel (Gleichung 3.18) für die Pivotspalte und Gleichung 3.21 für $\overline{A}$ zusammen, folgt:

$$\overline{A}^{+}_{l,\mu} = \frac{\overline{A}_{l,\mu}}{\widehat{w}_l} \qquad \mu = 1, \dots, m \tag{3.26}$$

(3.27)

und für $i \leq l$:

$$\overline{A}^{+}_{i,\mu} = \overline{A}_{i,\mu} - \frac{\widehat{w}_i \overline{A}_{l,\mu}}{\widehat{w}_l} \qquad \mu = 1, \dots, m \tag{3.28}$$

Verfahren 3.35 (revidiertes Simplex-Verfahren, Phase 2)
Gegeben seien eine Basis $B = B^{(0)}$ von A, die Matrix $\overline{A} = A_B^{-1}$ und $\overline{x} = x_B$. Setze $k := 0$.

1. Berechne $\overline{y} = A^T c_B, \widehat{s}$ mit den Komponenten $\widehat{s}_r = c_{n(r)} - \overline{y} A_{.n(r)}$.
2. Optimalitätskriterium: Gilt $\widehat{s} \geq 0$, dann ist die aktuelle Ecke optimal und das Verfahren stoppt.
3. Wähle $j \in \{1, \dots, n-m\}$ mit $\widehat{s}_j < 0$ und berechne $\widehat{w} = \overline{A} A_{.n(j)}$.
4. Unbeschränktheitskriterium: $\widehat{w} \leq 0 \Rightarrow 0$
5. Wähle $l \in \{1, \dots, m\}$ und δ wie oben.
6. Update von $\overline{x}$ und $\overline{A}$ nach Gleichung 3.25, Gleichung 3.26 und Gleichung 3.28.
7. Indextausch: $b(l) \Leftrightarrow n(j)$
8. Setze $k := k + 1$ und gehe zum ersten Punkt.

4. Optimalitätsbedingungen und Dualität

Wir erinnern uns an die Analysisvorlesung. Seien $f\colon \mathbb{R}^n \to \mathbb{R}, g\colon \mathbb{R}^n \to \mathbb{R}^m$ stetig differenzierbar. Nun betrachten wir die Aufgabe $\min f(x)$ unter den Nebenbedingungen $g(x) = 0$.

Falls $\tilde{x}$ Lösung des linearen Problems ist und $g(\tilde{x})$ vollen Rang hat, dann gibt es ein $\lambda \in \mathbb{R}^m$ mit $f'(\tilde{x}) + \lambda^T g'(\tilde{x}) = 0$. Das λ wird als **Lagrangemultiplikator** bezeichnet. Für den Spezialfall $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, g(x) = Ax - b$ gilt:

$$g(x) = 0 \Leftrightarrow Ax = b$$
$$f'(\tilde{x}) + \lambda^T A = 0 \Leftrightarrow \nabla f(\tilde{x}) + A^T \lambda = 0$$

Nutzen dieser Bedigungen:

1. $\tilde{x}$ Lösung $\Rightarrow \exists \lambda\colon \nabla f(\tilde{x}) + A^T\lambda = 0$ (n Gleichungen)
2. $\tilde{x}$ ist Lösung $\Rightarrow$ $\tilde{x}$ ist zulässig $\Rightarrow A\tilde{x} = b$ (m Gleichungen)

Damit haben wir $n{+}m$ Gleichungen für die $n{+}m$ Unbekannten $\tilde{\alpha} \in \mathbb{R}^n, \lambda \in \mathbb{R}^m$.

4.1. Optimalitätsbedingungen

Sei $c \in \mathbb{R}^n, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, m \leq n$ sowie das lineare Problem $\min c^T x$ unter den Nebenbedingungen $x \in \mathcal{F} = \{x \in \mathbb{R}^n\colon x \geq 0, Ax = b\}$.

Satz 4.1 (Kuhn-Tucker)
Ein Punkt $\tilde{x} \in \mathcal{F}$ ist genau dann Lösung des linearen Problem, wenn ein $\tilde{y} \in \mathbb{R}^m, \tilde{s} \in \mathbb{R}^n$ existiert mit:

$$A^T\tilde{y} + \tilde{s} = c \qquad \tilde{x}^T\tilde{s} = 0 \qquad \tilde{s} \geq 0 \tag{4.1}$$

Beweis: 21
„⇒“ Sei $x \in \mathcal{F}$ die Lösung des linearen Problems. Wir betrachten zunächst den Fall $\mathrm{rg}(A) = m$. Da das Problem eine Lösung hat, berechnet das Simplex-Verfahren (nach **??**) in endlich vielen Schritten eine optimale Ecke x zu einer Basis B von A.

Weiter ist die Optimalitätsbedingung $s_N \geq 0$ erfüllt und $s_N^T = c_N^T - c_B^T A_B^{-1} A_N$. mit $\tilde{y}^T :=$ $c_B^T A_B^{-1}$ gilt daher:

$$\begin{aligned} s_N = c_N - A_N^T \tilde{y} &\Leftrightarrow A_N^T \tilde{y} \leq c_N \\ &\Rightarrow \begin{pmatrix} A_B^T \\ A_N^T \end{pmatrix} \tilde{y} = \begin{pmatrix} c_B \\ A_N^T \tilde{y} \end{pmatrix} \leq \begin{pmatrix} c_B \\ c_N \end{pmatrix} \\ &\Leftrightarrow A^T \tilde{y} \leq c \end{aligned}$$

Wir definieren $\tilde{s} := c - A^T \tilde{y}$. Dann gilt $\tilde{s} \geq 0, A^T \tilde{y} + \tilde{s} = c$ und weiter $b^T \tilde{y} = y^T b =$ $c_B^T \underbrace{A_B^{-1} b}_{=x_B} = c_B^T x_B =^{x_N=0} c^T x = c^T \tilde{x} \Rightarrow \tilde{x}^T \tilde{s} = \tilde{x}^T (c - A^T \tilde{y}) = c^T x - (\underbrace{A\tilde{x}}_{=b}) \tilde{y} = c^T \tilde{x} - b\tilde{y} =$ $c^T \tilde{x} - c^T \tilde{x} = 0$.

$$\tilde{x}^T \tilde{s} = 0 \tag{4.2}$$

Ist $rg(A) = m_1 < m$, dann können wir $m - m_1$ redundante Gleichungen streichen. Damit erhält man ein äquivalentes Problem zum (LP) und man kann den obigen Beweis anwenden. Man erhöht Gleichung 4.1 für ein $\tilde{y} \in \mathbb{R}^m$. Setzt man für eine redundante Gleichung mit Index i: $\tilde{y}_i = 0$, dann erhält man Gleichung 4.1 für das Ausgangsproblem.

„$\Leftarrow$“ Seien $\tilde{x} \in \mathcal{F}$ und $\tilde{y} \in \mathbb{R}^m, \tilde{s} \in \mathbb{R}^n$ mit Gleichung 4.1 gegeben. Dann gilt Gleichung 4.2 und es folgt:

$$c^T \tilde{x} = b^T y \tag{4.3}$$

Sei $x \in \mathcal{F}$ beliebig. Wegen Gleichung 4.1 ist $A^T \tilde{y} = c - \tilde{s}$. Wir zeigen $c^T x \geq c^T \tilde{x}$:

$$b^T \tilde{y} = (Ax)^T \tilde{y} = x^T A^T \tilde{y} = x^T (c - \tilde{s}) = x^T c - x^T \tilde{s} = c^T x - \underbrace{x^T \tilde{s}}_{\geq 0}$$

Weiter ist $x \geq 0, \tilde{s} \geq 0 \Rightarrow x^T \tilde{s} \geq 0 \Rightarrow c^T \tilde{x} = c^T x - x^T \tilde{s} \leq c^T x$. ∎

Die Vektoren $\tilde{y}, \tilde{s}$ heißen **Lagrangemultiplikatoren**. Der Beweis des obigen Satzes zeigt:

$$\tilde{x}^T \tilde{s} = 0 \Leftrightarrow c^T \tilde{x} = b^T \tilde{y} \tag{4.4}$$

Aus $\tilde{x}^T \tilde{s} = \sum_{i=1}^n \tilde{x}_i s_i = 0$ und $\tilde{x} \geq 0, \tilde{s} \geq 0 \Leftrightarrow$:

$$\tilde{x}_i \tilde{s}_i = 0 \qquad \forall i = 1, \ldots, n \tag{4.5}$$

d. h. ist $\tilde{x}_i > 0 \Rightarrow \tilde{s}_i = 0$ und $\tilde{s}_i > 0 \Rightarrow \tilde{x}_i = 0$. Die Gleichung 4.5 heißen **Komplementaritätsbedingungen**.

Somit folgt aus Satz 4.1, dass ein Punkt $\tilde{x} \in \mathcal{F}$ genau dann Lösung des linearen Problems ist, wenn $\exists \tilde{y} \in \mathbb{R}^m, \exists \tilde{s} \in \mathbb{R}^n : (x, y, s) = (\tilde{x}, \tilde{y}, \tilde{s})$ Lösung des Systems:

$$\begin{aligned} Ax = b \qquad & x \geq 0 \\ A^T y + s = 0 \qquad & s \geq 0 \\ x_i s_i = 0 \qquad & \forall i = 1, \ldots, n \end{aligned} \tag{4.6}$$

ist. Dabei heißt x **primale Variable** und y, s **duale Variablen**.

Bemerkung 4.2
Der Multiplikator $\tilde{y}$ war formal definiert durch $\tilde{y} = (c_B^T A_B^{-1})^T = (A_B^{-1})^T c_B$. Der Vektor $\tilde{y}$ wird vom Simplex-Verfahren mit berechnet. Es ist der untere rechte Tabelleneintrag.

Der Vektor $\tilde{s}$ war definiert durch:

$$\tilde{s} = c - A^T\tilde{y} \Leftrightarrow \tilde{s}_B = c_B - A_B^T\tilde{y} = c_B - c_B = 0$$
$$\tilde{s}_N = c_N - A_N^T\tilde{y} = c_N - c_B^T A_B^{-1} A_N$$

d. h. $\tilde{s}_B = 0_m$ und $\tilde{s}_n$ ist der vom Simplex-Verfahren berechnete Testvektor. Für eine optimale Basislösung $\tilde{x}$ zur Basis B gilt daher:

$$\tilde{x}^T\tilde{s} = \tilde{x}_B^T\tilde{s}_B + \tilde{x}_N^T\tilde{s}_N = 0$$

4.2. Duale Programme

Wegen Gleichung 4.4 ist das System aus Gleichung 4.6 äquivalent zu:

$$Ax = b \qquad x \geq 0 \tag{4.7}$$
$$A^T y + s = c \qquad s \geq 0$$
$$c^T x = b^T y \tag{4.8}$$

Beispiel 4.3
Wir betrachten das bereits bekannte Mischungsproblem (Satz 1.8). Gegeben sind drei Kornsorten s_1, s_2, s_3 mit den Nährstoffen A, B, C, D.

	s_1	s_2	s_3	Mindestbedarf
A	2	3	7	1250
B	1	1	0	250
C	5	3	0	900
D	0,6	0,25	1	232,5
Preis	41	35	96	

Das Ausgangs- oder primale Problem war die Herstellung einer Mischung, die den Nährstoffbedarf deckt und möglichst billig ist. Als lineares Problem:

$$\min 41x_1 + 35x_2 + 96x_3$$
unter den Nebenbedingungen
$$2x_1 + 3x_2 + 7x_3 \geq 1250$$
$$x_1 + x_2 \geq 250$$
$$5x_1 + 3x_2 \geq 900$$
$$0,6x_1 + 0,25x_2 + x_3 \geq 232,5$$
$$x_i \geq 0$$

Ein zweiter Händler benutzt andere Kornsorten und möchte ein Konkurrenzangebot machen. Die fiktiven Preise sind y_1 für A, y_2 für B, y_3 für C und y_4 für D. Das Ziel des zweiten Händlers ist, die fiktiven Preise so zu berechnen, dass der Gewinn möglichst groß ist und der Preis pro Mengeneinheit höchstens so groß ist, wie der des ersten Händlers. Die Zielfunktion ist:

$$1250y_1 + 250y_2 + 900y_3 + 232,5y_4 = b^T y$$

Als Nebenbedingungen gilt, dass das Angebot konkurrenzfähig sein muss:

$$\begin{aligned} s_1 &: 2y_1 + y_2 + 5y_3 + 0,6y_4 \leq 41 \\ s_2 &: 3y_1 + y_2 + 3y_3 + 0,25y_4 \leq 35 \\ s_3 &: 7y_1 + y_4 \leq 96 \end{aligned}$$

Sinnvollerweise müssen die $y_i \geq 0$ sein. Das duale Problem lautet: max $b^T y$ unter den Nebenbedingungen $A^T \leq c, y \geq 0$.

Die obige Tabelle kann auf Standardform transformiert werden. Damit folgt, $-y_i \leq 0, i = 1, \dots, 4 \Leftrightarrow y_i \geq 0$. Sei $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_4)^T$ Lösung des dualen Problems. Das Angebot des Konkurrenten sei beispielsweise:

	s'_1	s'_2	s'_3
A	3	2	3
B	0,5	4	1
C	6	4	1
D	1	0	0

Dann sind die optimalen Preise für die Kornsorten:

$$\begin{aligned} s'_1 &: 3\tilde{y}_1 + 0,5\tilde{y}_2 + 6\tilde{y}_3 + \tilde{y}_4 \\ s'_2 &: 2\tilde{y}_1 + 4\tilde{y}_2 + 4\tilde{y}_3 \\ s'_3 &: 3\tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 \end{aligned}$$

Allgemeiner kann man ausgehend vom primalen Problem folgendes formulieren:

(LP) $$\min c^T x$$

$$\text{unter den Nebenbedingungen} Ax = b, x \geq 0$$

$$\mathcal{F}_P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$$

(DS) $$\max b^T y$$

$$\text{unter den Nebenbedingungen} \begin{pmatrix} y \\ s \end{pmatrix} \in \mathcal{F}_D = \{y \in \mathbb{R}^m, s \in \mathbb{R}^n : A^T y + s = c, s \geq 0\}$$

Äquivalente Formulierung

(DP) $$\max b^T y$$

(4.9) $$\text{unter den Nebenbedingungen} y \in \tilde{\mathcal{F}}_D = \{y \in \mathbb{R}^m : A^T y \leq c\}$$

Beispiel 4.4
zu Beispiel Dort ist die zulässige Menge von (DP) angegeben: $\tilde{\mathcal{F}}_D = \{y \in \mathbb{R}^n : \tilde{A}^T y \leq \tilde{c}\}$ mit $\tilde{A} = (A, -I_m), \tilde{c} = \begin{pmatrix} c \\ 0_m \end{pmatrix} \Rightarrow \tilde{\mathcal{F}}_D = \{y \in \mathbb{R}^n : A^T y \leq c, -y \leq 0_n\} = \{y \in \mathbb{R}^n : A^T y \leq c, y \geq 0\}$.
Anwendung des Satzes von Kuhn-Tucker auf (DS) bzw. (DP).

richtigen Verweis finden

Satz 4.5

1. Ein Punkt $(\tilde{y}, \tilde{s}) \in \mathcal{F}_D$ ist genau dann Lösung von (DS), wenn es ein $\tilde{x} \in \mathbb{R}^m$ gibt:

$$A\tilde{x} = b \qquad \tilde{x} \geq 0 \qquad \tilde{x}s = 0 \tag{4.10}$$

2. Ein Punkt $\tilde{y} \in \tilde{\mathcal{F}}_D$ ist genau dann Lösung von (DP), wenn es ein $\tilde{x} \in \mathbb{R}^n$ gibt, so dass mit $\tilde{s} = c - A^T\tilde{y}$ die Bedingung oben (Gleichung 4.10) erfüllt ist.

Beweis: 22
Wir müssen (DS) auf ein Problem in Standardform transformieren, indem wir die Variablen y^-, y^+ mit $y = y^+ - y^-, y^+, y^- \geq 0$ einführen. (Variablen des transformieren Problems: s, y^+, y^-) Der Zielfunktionsvektor ist $(-b, b, 0_n)^T$. Das (DS) ist max $b^Ty \Leftrightarrow \min -b^Ty \Leftrightarrow \min -b^T(y^+ - y^-) + 0_n^T s$ und $A^Ty + s = c \Leftrightarrow A^T(y^+ - y^-) + I_s = c = (A^T, -A^T, I_n)(y^+, y^-, s)^T$. Somit ist die Nebenbedingung $(A^T, -A^T, I_n)(y^+, y^-, s)^T = c, y^+, y^-, s \geq 0$. Der Satz von Kuhn-Tucker ergibt, dass ein zulässiger Punkt $(\tilde{y}^+, \tilde{y}^-, \tilde{s})$ genau dann Lösung von $\min(-b^T, b^T, 0_n^T)(y^+, y^-, s)^T$ unter den Nebenbedingungen $(A^T, -A^T, I_n)(y^+, y^-, s)^T = c$, wenn es Vektoren (lagrangemultiplikatoren) $\xi \in \mathbb{R}^n, \sigma_+, \sigma_- \in \mathbb{R}^m, \sigma \in \mathbb{R}^n$ gibt mit:

$$\begin{pmatrix} A \\ -A \\ I_n \end{pmatrix} \xi + \begin{pmatrix} \sigma_+ \\ \sigma_- \\ \sigma \end{pmatrix} = \begin{pmatrix} -b \\ b \\ 0_n \end{pmatrix} \tag{4.11}$$

$$\sigma_+, \sigma_-, \sigma \geq 0 \tag{4.12}$$

$$(\tilde{y}^+)^T\sigma_+ + (\tilde{y}^-)^T\sigma_- + \tilde{s}^T\sigma = 0 \tag{4.13}$$

Für die dritte Gleichung in Gleichung 4.11 ergibt sich $\xi + \sigma = 0 \Leftrightarrow \xi = -\sigma$ weiter folgt nach der Addition der ersten beiden Gleichungen $A\xi - A\xi + \sigma_+ + \sigma_- = -b + b = 0 \Leftrightarrow \sigma_+ + \sigma_- = 0 \overset{\text{Gleichung 4.12}}{\Leftrightarrow} \sigma_+ = 0 = \sigma_-$. Sei $\tilde{x} := \sigma = -\xi$. Dann gilt $A\tilde{x} = b$ (wegen der zweiten Gleichung in Gleichung 4.11), $\tilde{x} \geq 0$ und $\tilde{s}^T\tilde{x} = 0$ wegen Gleichung 4.13. ∎

Bemerkung 4.6
Die Aussagen von Satz 4.1 und Satz 4.5:

- Ein primal zulässiger Punkt $\tilde{x} \in \mathcal{F}_P$ ist genau dann Lösung von (LP), wenn ein dual zulässiger Punkt $(\tilde{y}, \tilde{s}) \in \mathcal{F}_D$ mit $c^T\tilde{x} = b^T\tilde{y}$ existiert.
- Ein dual zulässiger Punkt $(\tilde{y}, \tilde{s}) \in \mathcal{F}_D$ ist genau dann Lösung von (DS), wenn ein primal zulässiger Punkt $\tilde{x} \in \mathcal{F}_P$ mit $b^T\tilde{y} = c^T\tilde{x}$ existiert.

Satz 4.7 (Schwacher Dualitätssatz)
Gegeben seien das Problem (LP) und das dazu duale Problem (DS). Dann gilt:

1. Sind $x \in \mathcal{F}_P$ und $(y, s) \in \mathcal{F}_D$, dann ist $b^Ty \leq c^Tx$.
2. Sind $\tilde{x} \in \mathcal{F}_P$ und $(\tilde{y}, \tilde{s}) \in \mathcal{F}_D$ mit $b^T\tilde{y} = c^T\tilde{x}$, dann ist $\tilde{x}$ Lösung von (LP) und $\tilde{y}$ Lösung von (DS).

Beweis: 23

1. Seien $x \in \mathcal{F}_P$ und $(y, s) \in \mathcal{F}_D$ beliebig. Dann gilt $b^T y = (Ax)^T y = \underbrace{x^T}_{\geq 0} \underbrace{A^T y}_{\leq c} \leq x^T c = c^T x$.
2. Seien $\tilde{x} \in \mathcal{F}_P$ und $(\tilde{y}, \tilde{s}) \in \mathcal{F}_D$ mit $c^T \tilde{x} = b^T \tilde{y}$. Daraus folgt, dass für einen beliebigen Punkt $x \in \mathcal{F}_P$ und $(y, s) \in \mathcal{F}_D$ gilt:

   $$b^T y \leq c^T \tilde{x} = b^T \tilde{y} \leq c^T x \qquad \forall x \in \mathcal{F}_P \, \forall (y, s) \in \mathcal{F}_D$$

   Somit ist $\tilde{x}$ Lösung von (LP) und $(\tilde{y}, \tilde{s})$ Lösung von (DS). ∎

Beispiel 4.8
Wir wenden die Erkenntnisse auf das Mischungsproblem an. Der Vektor x (primal zulässig) definiere eine zulässige Mischung für den ersten Händler. Der Vektor y (dual zulässig) definiere ein zulässiges (fiktives) Angebot des zweiten Händlers. Dann gilt nach dem obigen Satz $b^T y \leq c^T x$, d. h. ein beliebiges Angebot des Konkurrenten muss billiger als jede zulässige Mischung des ersten Händlers sein. Ist $\tilde{x}$ eine zulässige Mischung des ersten Händlers und $\tilde{y}$ ein zulässiges Angebot des Konkurrenten mit $c^T \tilde{x} = b^T \tilde{y}$ (gleiche Kosten), dann ist $\tilde{x}$ optimal für den ersten Händler und $\tilde{y}$ optimal für den zweiten Händler.

Beispiel 4.9 (zum praktischen Nutzen der Dualität)
Sei $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. Diskrete lineare Optimierungsaufgabe:

$$\min_{x \in \mathbb{R}^n} f(x) = \|Ax - b\| \qquad \|y\| = \max_{i=1,\ldots,n} \|y_i\|$$

Dies ist äquivalent zu dem (LP): $\min \delta$ unter den Nebenbedingungen $\delta \geq 0, -\delta e \leq ax - b \leq \delta e$. Dabei ist e der Einheitsvektor. Vergleiche:

$$\begin{gathered} \min(-b^T, b^T, 0) y \text{ unter den Nebenbedingungen} \\ \begin{pmatrix} A^T & -A^T & 0_n \\ e^T & e^T & 1 \end{pmatrix} y = \begin{pmatrix} 0_n \\ 1 \end{pmatrix}, y \geq 0 \end{gathered}$$

Bemerkung 4.10
Es gelte $\mathcal{F}_P \neq \emptyset$ und $\mathcal{F}_D \neq \emptyset$. Dann gilt nach Satz 4.7 für beliebige $x \in \mathcal{F}_P, (y, s) \in \mathcal{F}_D$: $b^T y \leq c^T x$ und es folgt, dass die Zielfunktion von (LP) ist auf $\mathcal{F}_P$ nach unten beschränkt. Weiterhin folgt nach dem Hauptsatz der linearen Optimierung, dass das (LP) eine Lösung hat. Dies ist analog, dass (DS) eine Lösung hat.

Satz 4.11 (Starker Dualitätssatz)
Gegeben seien das lineare Problem (LP) und das duale Problem (DS). Ist $\mathcal{F}_P \neq \emptyset$ und $\mathcal{F}_D \neq \emptyset$, dann besitzen beide Probleme eine Lösung und die Optimalwerte sind gleich.

Beweis: 24
Zur Existenz der Lösung vergleiche Satz 4.10. Sei $\tilde{x} \in \mathcal{F}_P$ Lösung von (LP). Nach dem Satz 4.1 gibt es $(\tilde{y}, \tilde{s}) \in \mathcal{F}_P$ mit $\tilde{x}^T \tilde{s} = 0 \Leftrightarrow b^T \tilde{y} = c^T \tilde{x}$. Somit folgt nach dem schwachen Dualitätssatz, dass $(\tilde{y}, \tilde{s})$ Lösung von (DS) ist. ∎

4.3. Sensitivitätsanalyse

Beispiel 4.12 (Produktionsplanungssystem)
Aus drei Betriebsmitteln B_1, B_2, B_3 sollen vier Produkte P_1, P_2, P_3, P_4 hergestellt werden:

	P_1	P_2
B_1	1	1
B_2	7	5
B_3	3	5
Gewinn	4	5

Als Gleichung $\max 4x_1 + 5x_2 + 9x_3 + 11x_4$ unter den Nebenbedingungen

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &\leq 15 \\ 7x_1 + 5x_2 + 3x_3 + 2x_4 &\leq 120 \\ 3x_1 + 5x_2 + 10x_3 + 15x_4 &\leq 100 \\ x_i &\geq 0, i = 1, \dots, 4 \end{aligned}$$

Die Standardform ist gegeben durch:

$$c = (-4, -5, -9, -11, 0, 0, 0)^T$$

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 7 & 5 & 3 & 2 & 0 & 1 & 0 \\ 3 & 5 & 10 & 15 & 0 & 0 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 15 \\ 120 \\ 100 \end{pmatrix}$$

Das Simplex-Verfahren berechnet:

	5	2	7	4		1	6	3
1	10/7	5/7	-1/7	-5/7	50/7	10/7	0	-1/7
6	-61/7	-6/7	4/7	13/7	325/7	-61/7	1	4/7
3	-3/7	2/7	1/7	12/7	55/7	-3/7	0	1/7
	13/7	3/7	5/7	11/7	-695/7	-13/7	0	-5/7

Die optimale primale Lösung ist $\tilde{x} = (50/7, 0, 55/7, 0, 0, 325/7, 0)^T$ und die optimale duale Lösung ist $\tilde{y} = (-13/7, 0, -5/7)$.

Für den Anwender ergeben sich also folgende Schritte:

1. Schritt Berechnung einer Optimallösung

2. Schritt Sensitivitätsanalyse, d. h. Untersuchung der Abhängigkeit der Lösung und des Optimalwertes von Parametern.

Allgemein betrachten wir die folgende Situation:

$((\mathrm{LP})_v)$ $\qquad \min c^T x$ unter den Nebenbedingungen

$$x \in \mathcal{F}(v) = \{x \in \mathbb{R}^n : x \geq 0, Ax = b + v\}$$

(4.14)

Frage: Wie hängen Lösung und Optimalwert von $v \in \mathbb{R}^n$ ab? Das duale Problem lautet:

$$((\mathrm{DP})_v) \qquad \max(b+t)^T y \text{ unter den Nebenbedingungen } \quad y \in \tilde{\mathcal{F}}_D = \{y \in \mathbb{R}^n : A^T y \le c\} \tag{4.15}$$

Sei x eine Lösung des linearen Problems für $v = 0_m$ zur Basis B und $\tilde{y}$ der zugehörige Multiplikator. Dann ist:

$$\tilde{x}_B = A_B^{-1} b \qquad \tilde{y} = A_B^{-1} c_B$$

Für $v \in \mathbb{R}^m$ definieren wir:

$$x_B(v) = A_B^{-1}(b+v) = \tilde{x}_B + A_B^{-1} v \qquad x_N(v) = 0_{n-m}$$

Satz 4.13
Unter obiger Voraussetzung gelte $x_B(v) \ge 0$ (und damit auch $x(v) \ge 0$). Dann ist $x(v)$ Lösung von $(LP)_v$. Für den Optimalwert gilt:

$$c^T x(v) = c^T \tilde{x} + \tilde{y}^T v \tag{4.16}$$

Weiter ist auch $\tilde{y}$ Lösung von $(DP)_v$.

Beweis: 25
Nach der Definition von $x(v)$ gilt: $Ax(v) = A_B x_B(v) = b + v$. Nach der Voraussetzung ist $x(v) \ge 0 \Rightarrow x(v) \in \mathcal{F}_v$. Weiter ist $\tilde{y} \in \tilde{\mathcal{F}}_D$ und es gilt nach dem starken Dualitätssatz $c^T \tilde{x} = b^T \tilde{y}$ und $\tilde{y} = (A_B^{-1})^T c_B \Rightarrow c^T x(v) = c_B^T x_B(v) = \underbrace{c_B^T \tilde{x}_B}_{=c^T \tilde{x} = b^T \tilde{y}} + c_B^T A_B^{-1} v = (b+t)^T \tilde{y}$. Damit gilt $x(v) \in \mathcal{F}_v, \tilde{y} \in \tilde{\mathcal{F}}_D$ und $c^T x(v) = (b+v)^T \tilde{y}$ und es folgt die Behauptung. ∎

Bemerkung 4.14
Bleibt $x(v) \ge 0 \Rightarrow x(v)$ ist Lösung von $(LP)_v$ zur Basis B. Die Gleichung 4.16 zeigt: Der Multiplikator $\tilde{y}$ (duale Lösung) bestimmt, wie sich der Optimalwert in Abhängigkeit von v ändert. Die $\tilde{y}_i$ mit $i = 1, \dots, n$ heißen **Spaltenpreise**.

Folgendes ist für die obige Theorie wichtig:
Die Referenzlösung für $v = 0_m$ ist eine Basislösung an der Basis B von A und damit Ecke der zulässigen Menge. Benutzt man das Simplex-Verfahren zur Lösung des Referenzproblems, so erhält man immer eine optimale Basislösung.

Beispiel 4.15 (Fortsetzung von Satz 4.12)
Die Kapazität (Vorrat) b_1 von B_1 soll geändert werden. Dazu definieren wir $v_\varepsilon = \varepsilon \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Somit gilt:

$$x_B(v_\varepsilon) = \tilde{x}_B + \varepsilon A_B^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 50/7 \\ 325/7 \\ 55/7 \end{pmatrix} + \varepsilon \begin{pmatrix} 10/7 \\ -61/7 \\ -3/7 \end{pmatrix}$$

Für $-5 \le \varepsilon \le 325/61$ ist $x_B(v_\varepsilon) \ge 0$ und damit also Lösung von $(LP)_{v_\varepsilon}$. Der Optimalwert ist $c^T \tilde{x} + \tilde{y}^T v_\varepsilon = c^T \tilde{x} + \varepsilon \tilde{y}_1 = -695/7 - 13/7 \varepsilon$.

5. Affine Skalierung

(einfachstes Innere-Punkte-Verfahren)

5.1. Einführung

Das Simplex-Verfahren wurde gegen Ende der 50er Jahre von George B. Dantzig entwickelt. Sei $S(n, m)$ die Anzahl der Iterationen des Simplex-Verfahrens. Dann gilt $s(n, m) \leq \binom{n}{m}$. Im schlechtesten Fall werden alle Ecken durchlaufen.

Beispiel 5.1 (Klee-Minty-Würfel)
Wir betrachten das lineare Problem min $\sum_{j=1}^{n} 10^{n-j} x_j$ unter den Nebenbedingungen $x_1 \leq 1, 2\sum_{j=1}^{i-1} 10^{i-j} x_j + x_i \leq 100^{i-1}$ für alle $i = 2, \ldots, n$. Die zulässige Menge ist dabei ein verzerrter Würfel. Hierbei wächst die Anzahl der Iterationen im schlechtesten Fall exponentiell. Dies wurde 1972 von Klee und Minty herausgefunden. Später fand man auch ähnliche Beispiele für Zeilen- und Spaltenauswahlregeln.

5.2. Das Konzept des Verfahrens

Wir betrachten das lineare Problem:

$$\text{((LP))} \qquad \begin{aligned} &\min c^T x \text{ unter den Nebenbedingungen} \\ &x \in \mathcal{F}_P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\} \end{aligned}$$

Definition 5.2 (Innerer zulässiger Punkt)
Ein Vektor $x \in \mathbb{R}^n$ heißt **innerer zulässiger Punkt** von (LP), wenn $x \in \mathcal{F}_P$ und $x > 0$ ist.

Bemerkung 5.3
Das Innere-Punkt-Verfahren startet mit einem inneren zulässigen Punkt $x^{(0)}$ und berechnet eine Folge $\{x^{(k)}\}_{k\in\mathbb{N}}$ innerer zulässiger Punkte. Wir bezeichnen die Menge der inneren zulässigen Punkte mit $\mathcal{F}_P^0$. Sei nun $x^{(k)} \in \mathcal{F}_P^0$ nicht optimal. Dann stellt sich die Frage, wann man ausgehend von $x^{(k)}$ einen besseren Punkt $x^{(k+1)}$ berechnen kann.

Definition 5.4 (Zulässige Richtung)
Eine Richtung $d \in \mathbb{R}^n$ heißt **zulässige Richtung** von $x^{(k)}$, wenn ein $\overline{\sigma} > 0$ existiert, so dass:

$$x^{(k)} + \sigma d \in \mathcal{F}_P \qquad \forall \sigma \in [0, \overline{\sigma}] \tag{5.1}$$

Bemerkung 5.5
Wenn $x^{(k)} > 0$ ist, dann gilt $x^{(k)} + \sigma d > 0$, falls $\overline{\sigma}$ hinreichend klein ist.

Um Gleichung 5.1 sicher zu stellen, muss noch folgendes gelten:

$$A(x^{(k)} + \sigma d) = b = Ax^{(k)} + \sigma Ad \Leftrightarrow Ad = 0 \tag{5.2}$$

Somit sind alle Richtungen $d \in \mathbb{R}^n$ mit $Ad = 0$ zulässige Richtungen.

Definition 5.6 (Zulässige Abstiegsrichtung)
Eine Richtung $d \in \mathbb{R}^n$ heißt **zulässige Abstiegsrichtung**, wenn d zulässige Richtung ist und es gilt:

$$c^T(x^{(k)} + \sigma d) < c^T x^{(k)} \qquad \forall \sigma \in [0, \overline{\sigma}]$$

Um eine zulässige Abstiegsrichtung zu berechnen, betrachten wir das Problem:

$$\begin{aligned} &\min c^T d \text{ unter den Nebenbedingungen} \\ &Ad = 0, \|d\| \leq 1 \end{aligned} \tag{(AS)}$$

Die zulässige Menge von (AS) ist nicht leer und kompakt. Damit folgt nach dem Satz von Weierstraß, dass (AS) eine Lösung hat.

Lemma 5.7
Sei $\overline{x} \in \mathcal{F}_P^0$ keine Lösung von (LP). Weiter sei $\overline{d}$ eine Lösung von (AS). Dann ist $\overline{d}$ zulässige Abstiegsrichtung in $\overline{x}$.

Beweis: 26
Sei $A\overline{d} = 0 \Rightarrow \overline{d}$ ist zulässige Richtung. Nun ist noch zu zeigen, dass $\overline{d}$ eine Abstiegsrichtung ist. Wir nehmen an, $c^T\overline{d} \geq 0$. Sei $x \in \mathcal{F}_P^0$ beliebig und $d := \frac{x-\overline{x}}{\|x-\overline{x}\|}$. Dann gilt $\|d\| = 1, Ad = \frac{1}{\|x-\overline{x}\|}(\underbrace{Ax}_{=b} - \underbrace{A\overline{x}}_{=b}) = 0$. Dann folgt, dass d zulässige Lösung für (AS) $\Rightarrow c^T d \geq c^T\overline{d} \Rightarrow c^T d = \frac{d}{\|x-\overline{x}\|}(c^T x - c^T\overline{x}) \geq 0 \Rightarrow c^T x \geq c^T\overline{x} \Rightarrow \overline{x}$ ist Lösung des (LP). Dies steht aber in Widerspruch zur Voraussetzung. ∎

Im k-ten Iterationsschritt wollen wir $\|\cdot\|_b$ als Norm wählen.

Definition 5.8 (Ellipsoid)
Sei $Q \in \mathbb{R}^{n\times n}$ mit $\mathrm{rg}(Q) = n, x \in \mathbb{R}^n, \rho < 0$. Dann heißt die durch

$$E(Q, x, \rho) = \{y \in \mathbb{R}^n : (y-x)^T(Q^{-1})^T Q(y-x) \leq \rho^2\}$$

definierte Menge **Ellipsoid** mit Mittelpunkt x und Radius ρ.

Bemerkung 5.9
Ist $Q = I_n \Rightarrow E(I_n, x, \rho) = \{y \in \mathbb{R}^n : (y-x)^T(y-x) \leq \rho^2\} = \{y \in \mathbb{R}^n : \|y-x\|^2 \leq \rho^2\}$
$Also ist E(I_n, x, \rho)$ die abgeschlossene Kugel um x mit Radius ρ.

Sei allgemein $Q \in \mathbb{R}^{n\times n}$: $\operatorname{rg}(Q) = n \Rightarrow \operatorname{rg}(Q^{-1}) = n$. Die Matrix $(Q^{-1})^T Q$ ist symmetrisch und positiv definit. Man kann auf dem $\mathbb{R}^n$ eine Norm $\|\cdot\|_Q$ definieren:

$$\|z\|_Q :\ = \sqrt{z^T(Q^{-1})^T Q^{-1} z} = \|Q^{-1}z\|$$

Mit dieser Norm ist $E(Q, x, \rho) = \{y \in \mathbb{R}^n : \|y - x\|_Q \leq \rho\}$, d. h. ein Ellipsoid ist eine abgeschlossene Kugel bezüglich der durch Q definierten Norm.

Sei $x^{(k)} \in \mathcal{F}_P^0 \Rightarrow x^{(k)} > 0$. Damit folgt, dass die Matrix mit den Diagonalelementen $x^{(k)}$ invertierbar ist ($X_k = \operatorname{diag}(x^{(k)})$) und $X_k^{-1} = \operatorname{diag}(1/x_1^{(k)}, \ldots, 1/x_n^{(k)})$.

Das Verfahren der affinen Skalierung benutzt im k-ten Iterationsschritt beim Problem (AS) die Norm $\|\cdot\|_{X_k}$, d. h. zur Berechnung einer zulässigen Abstiegsrichtung lösen wir das Problem:

$$\text{((AS)}_\text{k}) \qquad \min_{d\in\mathbb{R}^n} c^T d \text{ unter den Nebenbedingungen}$$
$$Ad = 0, \|d\|_{X_k} \leq 1$$

Um (AS)_k einfach lösen zu können, führen wir eine „affine Skalierung" der Variablen durch:

$$u = X_k^{-1} d \Leftrightarrow d = X_k u$$

Dann gilt:

$$\begin{aligned} \|d\|_{X_k}^2 &= \|X_k u\|_{X_k}^2 = (X_k u)^T X_k^{-1} X_k^{-1} X_k u \\ &= u^T u = \|u\|^2 \\ \Rightarrow \|d\|_{X_k} &= \|u\| \end{aligned} \tag{5.3}$$

Damit erhalten wir aus (AS)_k das Problem:

$$\text{((AS}_\text{u})_\text{k}) \qquad \min_{u\in\mathbb{R}^n} (X_k c)^T u \text{ unter den Nebenbedingungen}$$
$$AX_k u = 0, \|u\| \leq 1$$

Beispiel 5.10 (Frannie)
$x^{(k)} = (3, 60, 21)^T$

5.3. Berechnung der Suchrichtung

Satz 5.11
Sei $x^{(k)} \in \mathcal{F}_P^0$ und $\operatorname{rg}(A) = m < n$. Wir definieren:

$$y^{(k)} :\ = (AX_k^2 A^T)^{-1} A X_k^2 c \qquad s^{(k)} :\ = X_k(c - A^T y^{(k)}) \tag{5.4}$$

Wenn $s^{(k)} = 0_n$ ist, dann ist $x^{(k)}$ Lösung von (LP). Andernfalls ist:

$$u^{(k)} :\ = -\frac{s^{(k)}}{\|s^{(k)}\|} \tag{5.5}$$

eindeutig bestimmte Lösung von $\text{(AS}_\text{u})_\text{k}$.

Beweis: 27
Die Matrix $\underbrace{AX_k^2A^T}_{=AX_k(AX_k)^T} \in \mathbb{R}^{m\times m}$ ist positiv definit,denn $y^T(AX_k)(AX_k)^Ty = (y^TAX_k)(y^TAX_k)^T = \|(AX_k)^T\|^2 > 0$. Somit ist die Formel für $y^{(k)}$ sinnvoll.

Sei $s^{(k)} = 0_n$. Dann gilt nach Definition $A^Ty^{(k)} = c$. Wegen $x^{(k)} \in \mathcal{F}_P^0$ folgt, $Ax^{(k)} = b, x^{(k)} \geq 0 \Rightarrow Ay^{(k)} + s^{(k)} = c$ mit $s^{(k)} \geq 0$. Somit folgt, dass $(x, y, s) = (x^{(k)}, y^{(k)}, s^{(k)})$ eine Lösung des Systems Gleichung 4.1 ist und nach dem Satz von Kuhn-Tucker ist $x^{(k)}$ eine Lösung von (LP).

Beweis zu Ende bringen

Die Lösungsformel für $y^{(k)}$ ist eine theoretische Formel. Zur praktischen Berechnung von $y^{(k)}$ löst man das lineare Gleichungssystem

$$\underbrace{AX_k^2A^T}_{=(AX_k)(AX_k)^T} y = AX_k^2c \tag{5.6}$$

Dies kann entweder durch das Choleskyverfahren geschehen. Eine weitere Lösungsmöglichkeit erhält man durch folgende Betrachtungen. Die Form $(AX_k)(AX_k)^Ty = AX_k^2c$ des Systems Gleichung 5.6 definiert die Normalengleichung für das überbestimmte Gleichungssystem $(AX_k)^Ty = X_kc$. Definiere als Lösung einen Vektor $\tilde{y}$, der die Zielfunktion $\|(AX_k)^Ty - X_kc\|$ minimiert.

Satz 5.12
Sei $x^{(k)} \in \mathcal{F}_P^0$ und es gelte $\operatorname{rg}(A) = m < n$. Weiterhin seien $y^{(k)}, s^{(k)}$ durch Gleichung 5.4 definiert. Ist $s^{(k)} = 0$, dann ist $x^{(k)}$ Lösung von (LP). Andernfalls ist

$$d^{(k)} = -\frac{X_ks^{(k)}}{\|s^{(k)}\|} \tag{5.7}$$

die eindeutige Lösung von $(\mathrm{AS})_\mathrm{k}$.

Beweis: 28
Für $s^{(k)} = 0$ siehe oben. Sei nun $s^{(k)} \neq 0$. Dann ist d genau dann Lösung von $(\mathrm{AS})_\mathrm{k}$, wenn $u = X_k^{-1}d$ Lösung von $(\mathrm{AS_u})_\mathrm{k}$ ist, d. h. $d^{(k)} = X_ku^{(k)}$. ∎

5.4. Wahl der Schrittweite

Für die Wahl der Schrittweite existieren zwei Varianten:

- Short-Step-Variante: Man benutzt in jedem Schritt eine feste Schrittweite $0 < \sigma < 1$.
- Long-Step-Variante: Man benutzt in jedem Schritt die größtmögliche Schrittweite.

5.4.1. Die Short-Step-Variante

Lemma 5.13
Sei $\mathrm{rg}(A) = m$. Ist $x^{(k)} \in \mathcal{F}_P^0$, $s^{(k)} \neq 0$, $d^{(k)}$ durch Gleichung 5.7 definiert und $0 < \sigma < 1$. Dann ist $x = x^{(k)} + \sigma d^{(k)} \in \mathcal{F}_P^0$.

Summen richtig einfügen

BEWEIS: 29
Es gilt $Ax^{(k)} = b, Ad^{(k)} = 0 \Rightarrow Ax = Ax^{(k)} + \sigma Ad^{(k)} = b$. Nun bleibt zu zeigen, dass $x \geq 0$. Für einen beliebigen Index $i \in \{1, \dots, n\}$ gilt:

$$
\begin{aligned}
&= \sigma^2 (d^{(k)})^T X_k^2 d^{(k)} = \sigma^2 (X_k^{-1} d^{(k)})^T (X_k^{-1} d^{(k)}) \\
&= \sigma^2 \|X_k^{-1} d^{(k)}\|^2 = \sigma^2 \|d^{(k)}\|_{X_k}^2 \leq \sigma^2
\end{aligned}
$$

Wegen $x^{(k)} < 0$ und $0 < \sigma < 1$ folgt:

$$
\begin{aligned}
&|x_i x_i^{(k)}| \leq \sigma x_i^{(k)} < x_i^{(k)} \qquad 1 \leq i \leq n \\
&x_i < x_i^{(k)} \Rightarrow |x_i - x_i^{(k)}| - x_i^{(k)} - x_i < x_i^{(k)} \Rightarrow x_i > 0 \\
&x_i \geq x_i^{(k)} \Rightarrow x_i \geq x_i^{(k)} > 0
\end{aligned}
$$

■

5.4.2. Die Long-Step-Variante

Man wählt in jedem Iterationsschritt möglichst große Schrittweite.

Lemma 5.14
Es gelte $\mathrm{rg}(A) = m$ und $x^{(k)} \in \mathcal{F}_P^0$ sei keine Lösung von (LP). Dann ist $s^{(k)} \neq = 0$ und $d^{(k)}$ sei durch Gleichung 5.7 definiert. Ist $d^{(k)} > 0$[1], dann ist die Zielfunktion von (LP) auf der zulässigen Menge $\mathcal{F}_P$ nicht nach unten beschränkt.

BEWEIS: 30
Für $\sigma > 0$ sei $x(\sigma) := x^{(k)} + \sigma d^{(k)}$. Dann gilt $x(\sigma) \geq 0$ für alle $\sigma \geq 0$ und $Ax(\sigma) = \underbrace{Ax^{(k)}}_{=b} + \sigma \underbrace{Ad^{(k)}}_{=0} = b$ für alle $\sigma \geq 0$, d. h. $x(\sigma) \in \mathcal{F}_P^0$. Nach Satz 5.7 ist $d^{(k)}$ als Lösung von $(\mathrm{AS})_\mathrm{k}$ zulässige Abstiegsrichtung, d. h. $c^T d^{(k)} < 0 \Rightarrow c^T x(\sigma) = c^T x^{(k)} + \sigma c^T c^{(k)} \xrightarrow{\sigma \to \infty} -\infty$.

Ist der aktuelle Iterationspunkt $x^{(k)}$ keine Lösung von (LP), dann gibt es zwei Möglichkeiten:

1. $d^{(k)} \geq 0$ bzw. $s^{(k)} \leq 0 \Rightarrow$ Das lineare Problem hat keine Lösung.
2. $\exists j \in \{1, \dots, n\}$: $d_j^{(k)} < 0$

Fall zwei zu Ende führen

■

[1] genau dann, wenn $s^{(k)} \leq 0$

5.5. Abbruchkriterium

Nach Satz 5.11 ist $x^{(k)}$ Lösung von (LP), wenn $s^{(k)} \geq 0$. Es gilt aber $x^{(k)} > 0$. In einer Lösung $\tilde{x}$ von (LP) gilt in der Regel $\tilde{x}_i = 0$ für mindestens einen Index $i \in \{1, \dots, n\}$, d. h. die Bedingung $s^{(k)} = 0_n$ ist in der Regel nicht erfüllt.

Sei nun $\varepsilon \geq 0$. Ein Punkt $\overline{x} \in \mathcal{F}_P$ heißt **ε-Minimum** von (LP), wenn $c^T\overline{x} \leq c^T x + \varepsilon$ für alle $x \in \mathcal{F}_P$ gilt. Hat (LP) eine Lösung $\tilde{x}$, dann muss $c^T\overline{x} \leq c^T \tilde{x}\varepsilon$ gelten. Entsprechend heißt ein Punkt $\overline{y} \in \mathcal{F}_D$ ε-Maximum von (DP), wenn $b^T\overline{y} \geq b^T y - \varepsilon$ für alle $y \in \mathcal{F}_D$ gilt. Hat (DP) eine Lösung $\tilde{y}$, muss gelten $b^T\overline{y} \geq b^T\tilde{y} - \varepsilon$.

Unser Ziel ist nun, dass Verfahren der affinen Skalierung abzubrechen, wenn $x^{(k)}$ ein ε-Minimum von (LP) ist.

Lemma 5.15
Es gelte $\mathrm{rg}(A) = m, \varepsilon \geq 0, x^{(k)} \in \mathcal{F}_P^0$. Weiter seien $y^{(k)}, s^{(k)}$ durch Gleichung 5.4 bestimmt. Gilt mit $e = (1, \dots, 1)^T$:

$$s^{(k)} = 0 \qquad e^T s^k \leq \varepsilon \tag{5.8}$$

dann ist $x^{(k)}$ ein ε-Minimum von (LP) und $y^{(k)}$ ein ε-Maximum von (DP).

Beweis: 31
Sei $s\colon = X_k^{-1} s^{(k)}$. Dann ist:

1. $s \geq 0, s = c - A^T y^{(k)} \Rightarrow A^T y^{(k)} + s = c, s \geq 0 \Rightarrow y^{(k)} \in \mathcal{F}_D$. Nach Voraussetzung ist $x^{(k)} \in \mathcal{F}_P$, d. h. $\mathcal{F}_P \neq \emptyset \neq \mathcal{F}_D$. Nach dem starken Dualitätssatz folgt, dass es eine Lösung von $\tilde{x}$ von (LP) und eine Lösung $\tilde{y}$ von (DP) gibt und außerdem gilt $c^T\tilde{x} = b^T\tilde{y}$. Wegen Gleichung 5.8 gilt weiter:
$$\begin{aligned} 0 &\leq (x^{(k)})^T s = e^T x_k s = e^T s^{(k)} \leq \varepsilon \\ \Rightarrow 0 &\leq (x^{(k)})^T s = (x^k)^T (c - A^T y^{(k)}) \\ &= c^T x^{(k)} - (x^{(k)})^T A^T y^{(k)} \\ &= c^T x^{(k)} - (Ax^{(k)}) y^{(k)} \\ &= c^T x^{(k)} - b^T y^{(k)} \leq \varepsilon \end{aligned}$$
und es folgt im Punkt:
2. $c^T\tilde{x} \leq c^T x^{(k)} \leq b^T y^{(k)} + \varepsilon$
Nach dem schwachen Dualitätssatz ist $b^T y^{(k)} \leq c^T\tilde{x} \Rightarrow c^T x^{(k)} + \varepsilon$, d. h. $x^{(k)}$ ist ε-Minimum von (LP). Wegen $c^T\tilde{x} = b^T\tilde{y}$ folgt weiter $b^T\tilde{y} - \varepsilon = c^T\tilde{x} - \varepsilon \leq b^T y^{(k)}$, d. h. $y^{(k)}$ ist ε-Maximum von (DP). ∎

Für das verfahren der affinen Skalierung gibt man ein Genauigkeitsschranke $\varepsilon > 0$ vor. Dann gibt es zwei Abbruchkriterien:

1. Die Bedingung in Gleichung 5.8 ist erfüllt. Dann ist $x^{(k)}$ ein ε-Minimum von (LP).
2. Es gilt $d^{(k)} \geq 0$. dann ist die Zielfunktion auf $\mathcal{F}_P$ nicht nach unten beschränkt.

5.6. Das Verfahren

Verfahren 5.16 (Affine Skalierung)
Gegeben sei ein Startpunkt $x^0 \in \mathcal{F}_P^0, 0 < \beta < 1$ und $\varepsilon > 0$. Setze $k \colon = 0$.

1. Setze $x_k = \operatorname{diag}(x^{(k)})$ und berechne
$$y^{(k)} = (Ax_k^2 A^T)^{-1} Ax_k^2 c \qquad s^{(k)} = x_k(c - A^T y^{(k)})$$
durch Lösung des linearen Ausgleichsproblems:
$$\min_{y \in \mathbb{R}^m} \|Ax_k^T y - x_k c\|^2$$
2. Falls $s^{(k)} \geq 0$ und $e^T s^{(k)} \leq \varepsilon$ stoppt das Verfahren, da $x^{(k)}$ ein ε-Minimum von (LP) ist.
3. Fall $s^{(k)} \leq 0$ stoppt das Verfahren, da die Zielfunktion auf $\mathcal{F}_P$ nicht nach unten beschränkt ist.
4. Berechne die Suchrichtung $d^{(k)} = -\frac{x_k^{(k)}}{\|s^k\|}$ und bei der Short-Step-Variante $x^{(k+1)} \colon = x^{(k)} + \beta d^{(k)}$ sowie bei der Long-Step-Variante $x^{(k+1)} \colon = x^{(k)} + \beta \sigma_k d^{(k)}$ mit $\sigma_k = \min\{-\frac{x_i^{(k)}}{d^{(k)}} \colon 1 \in \{1, \dots, n\}, d^{(k)} < 0\}$.
5. Setze $k \colon = k + 1$ und gehe zu Punkt 1. Zusätzliches Abbruchkriterium: $\|x^{(k+1)} - x^{(k)}\| = \beta \sigma_k d^{(k)} \leq \sigma$.

Beispiel 5.17 (Frannie)
Lösung: $(6, 0, 0)^T, x^{(0)=(2,1,1)^T, \varepsilon = 0{,}01}$
Short-Step-Variante: Stopp nach 16 Iterationen
Long-Step-Variante: Stopp nach 11 Iterationen

Beispiel 5.18

$$\begin{aligned} \min -x_1 - 3x_2 \text{ unter den Nebenbedingungen} \\ x_1 + 3x_2 \leq 13 \\ 3x_1 x_2 \leq 15 \\ -x_1 + x_2 \leq 3 \\ x_1, x_2 \geq 0 \end{aligned}$$

Die Lösung mit der Long-Step-Variante ist $x^{(0)} = (2, 2, 5, 7, 3)^T, \varepsilon = 0{,}01$. Es stoppt nach 6 Iterationen mit $x^{(6)} = (2{,}46\dots, 3{,}51\dots, 7{,}8 \cdot 10^{-8}, 4{,}08\dots, 1{,}95\dots)^T$. Weiter ist:

$$s^{(6)} = \begin{pmatrix} 4{,}4 \cdot 10^{-16} \\ 0{,}7 \cdot 10^{-8} \\ 8{,}8 \cdot 10^{-16} \\ 6{,}4 \cdot 10^{-16} \end{pmatrix} \qquad c - A^T y^{(6)} = \begin{pmatrix} 2{,}2 \cdot 10^{-16} \\ 0 \\ 1 \\ 2{,}1 \cdot 10^{-16} \\ 3{,}2 \cdot 10^{-16} \end{pmatrix} = \widetilde{s}^{(6)}$$

$$\tilde{x}^T \tilde{s} = 0 \qquad \tilde{x} + \tilde{s} > 0$$

5.7. Strikte Komplementarität

Goldman-Tucker-Theorem Falls (LP) eine Lösung hat, folgt, dass es mindestens eine strikt komplementäre Lösung des Systems

$$Ax = b \quad A^T y + s = c \quad x_i s_i = 0 \qquad x, s \geq 0, i = 1, \dots, n \tag{5.9}$$

gibt.

Strikte Komplementaritätsbedingung Eine Lösung $(\tilde{x}, \tilde{y}, \tilde{s})$ von Gleichung 5.9 heißt **strikt komplementär**, wenn gilt:

$$\tilde{x} + \tilde{s} > 0 \tag{5.10}$$

Wir bezeichnen die Lösungsmenge des primalen Problems mit S_P und die Lösungsmenge des dualen Problems mit S_D.

$$\begin{aligned} J &= \{1, \dots, n\} \\ B &= \{i \in J \colon x_j \neq 0 \,\forall x \in S_P\} \\ N &= \{j \in j \colon s_j \neq 0 \,\forall (y, s) \in S_D\} \end{aligned} \tag{5.11}$$

Bemerkung 5.19
Sei $x \in S_P$ beliebig. Dann ist $x_N = 0$. Denn zu jedem Index $j \in N$ gibt es nach der Definition von N ein $(y, s) \in S_D$ mit $s_j \neq 0$. Es ist $c^T x = b^T y \overset{\text{Gleichung 5.2}}{\Leftrightarrow} x^T s = 0 \Leftrightarrow x_i s_i = 0, i = 1, \dots, n$. Da (x, y, s) Lösung von Gleichung 5.9 ist, folgt $x_j s_j = 0 \Rightarrow x_j = 0$.

Entsprechend gilt für beliebige $(y, s) \in S_D \colon s_B = 0$. Denn zu beliebigem $j \in B$ gibt es ein $x \in S_P$ mit $x_j \neq 0$. Da (x, y, s) Lösung von Gleichung 5.9 ist $x_j s_j = 0 \Rightarrow s_j = 0$. Zu jedem Index $j \in B$ gibt es ein $x^{(j)} \in S_P$ mit $x_j^{(j)} \neq 0$. Da S_P konvex ist, folgt,

$$x \colon = \frac{1}{|B|} \sum_{j \in B} x^{(j)} = \sum_{j \in B} \alpha_j x^{(j)} \text{ mit } \alpha_j = \frac{1}{|B|}$$

Es gilt $\alpha_j \geq 0$, $\sum_{j \in B} \alpha_j = 1$ und weiter ist $x_B > 0$, d. h. ist $B \neq \emptyset$, dann gibt es eine Lösung $x \in S_P$ mit $x_B > 0$.

Genauso zeigt man: Ist $N \neq \emptyset$, dann gibt es eine Lösung $(y, s) \in S_D$ mit $s_N > 0$. Wenn die folgenden Indexmengen B, N die Eigenschaften $b \cap N = \emptyset, B \cup N = J = \{1, \dots, n\}$ haben, folgt, dass (x, y, s) die strikt komplementäre Lösung ist. Somit braucht zum Beweis des Goldman-Tucker-Theorems nur die obigen Eigenschaften zeigen.

Zu „$B \cap N = \emptyset$“: Sei j ein Index mit $j \in B \cap N$. Dann gibt es ein $x \in S_P$ mit $x_j \neq 0$ und ein $(y, s) \in S_D$ mit $s_j \neq 0 \Rightarrow x_j s_j \neq 0$ ↯

Lemma 5.20
Sei G eine $p \times n$-Matrix, $g \in \mathbb{R}^n$. Dann hat entweder das System

$$G^T v = g \qquad v \geq 0 \tag{5.12}$$

eine Lösung $v \in \mathbb{R}^p$ oder das System

$$Gd \geq 0 \qquad g^T d < 0 \tag{5.13}$$

eine Lösung $d \in \mathbb{R}^n$. Aber beide Systeme sind nicht gleichzeitig lösbar.

Korollar 5.21
Sei G eine $p \times n$-Matrix, H eine $q \times n$-Matrix, $g \in \mathbb{R}^n$. Dann hat entweder das System

$$G^T v + H^T w = g \qquad v \geq 0 \tag{5.14}$$

eine Lösung $v \in \mathbb{R}^p$ oder das System

$$Gd \geq 0 \qquad Hd = 0 \qquad g^T d < 0 \tag{5.15}$$

eine Lösung $d \in \mathbb{R}^n$

$$B = \{j \in J : x_j \neq 0, x \in S_P\} \tag{5.16}$$

$$N = \{j \in J : s_j \neq 0, (y, s) \in S_D\} \tag{5.17}$$

Bemerkung 5.22
Sei $x \in S_P$ beliebig. Dann ist $x_N = 0$. Denn zu jedem Index $j \in N$ gibt es (nach der Definition von N) ein $(y, s) \in S_D$ mit $s_j \neq 0$. Da (x, y, s) Lösung von Gleichung 5.9 ist $x_j = 0$.

Satz 5.23 (Goldman-Tucker-Theorem)
Es sei $S_P \neq 0$ (Dann ist auch $S_D \neq 0$). Dann gilt $B \cup N = \{1, \dots, n\}$, d. h. es gibt mindestens eine strikt komplementäre Lösung.

Beweis: 32
Wir zeigen, dass die Indexmenge $I := J \setminus (B \cup N)$ leer ist. Annahme: Es gibt $k \in I$. Das System $A_{.k}^T w < 0, A_B^T w = 0, -A_{.j}^T \geq 0$ für $j \in I \setminus \{k\}$ hat eine Lösung $w \in \mathbb{R}^n$.

Sei $(\tilde{x}, \tilde{y}, \tilde{s})$ eine primal duale Lösung mit $\tilde{s}_w > 0$. Wir definieren (y^*, s^*) durch $y^* = \tilde{y} + \varepsilon w$, $s^* = c - A^T y^* = c - A^T \tilde{y} - \varepsilon A^T w = \tilde{s} - \varepsilon A^T w$. Dann gilt

$$\begin{aligned}
s_k^* &= \underbrace{\tilde{s}_k}_{\geq 0} - \varepsilon \underbrace{A_k^T w}_{<0} > 0 \qquad && \varepsilon > 0 \\
s_j^* &= \underbrace{\tilde{s}_j}_{\geq 0} - \underbrace{\varepsilon A_k^T w}_{\geq 0} \geq 0 \qquad && \varepsilon > 0, j \in I \setminus \{k\} \\
s_B^* &= \tilde{s}_B - \varepsilon A_B^T w = \tilde{s}_B = 0 \qquad && \varepsilon > 0
\end{aligned}$$

Weiter ist $s_N^* = \tilde{s}_N - \varepsilon A_N^T w > 0$ für $\varepsilon > 0$ hinreichend klein. Insgesamt ist $s^* \geq 0$ für hinreichend kleines $\varepsilon > 0$. Weiter ist $s^* = c - A^T y^* \Rightarrow (y^*, s^*) \in \mathcal{F}_D$. Es gilt sogar, $(y^*, s^*) \in S_D$. Denn für beliebiges $x^* \in S_P$ ist $x_i^* s_i^* = 0$. Für $i \in B$ gilt dies wegen $s_B^* = 0$. Für $i \in N$ gilt dies wegen $x_N^* = 0$. Für $i \in I$ ist $x_i^* = 0$ und wegen $s_k^* > 0$ muss $k \in N$ sein. ∎

Beispiel 5.24
Simplexverfahren mit Startpunkt $x^{(0)} = (0, 0, 13, 15, 2)^T$. Nach 2 Iteratioen: $x = (1, 4, 0, 8, 0)^T$, $y = (-1, 0, 0)^T$. Mit $s = c - A^T y(0, 0, 1, 0, 0)^T$ erhalten wir eine primal duale Lösung. Wegen $x_5 = s_5 = 0$ ist die Lösung *nicht* strikt komplementär. Die Lösung des linearen Problems ist $\{(1+3t, 4-t, 0, 8-8t, 4t)^T : t \in [0, 1]\}$. Die Punkte $(x(t), y, s)$ sind primal duale Lösungen. Für $t \in (0, 1)$ gilt: $x(t)_i > 0, s_i = 0, i = 1, 2, 4, 5, x(t)_3 = 0, s < 0 \Rightarrow$ für $t \in (0, 1)$ sind die Lösungen strikt komplementär. Für $t = 0$ ist $x(t)_5 = s_5 = 0$ und für $t = 1$ ist $x(t)_4 = s_4 = 0$. Beide Lösungen sind nicht strikt komplementär.

Korollar 5.25
Es sei $s_P \neq 0$ oder äquivalent $s_D \neq 0$. Dann gilt $s_P = \{x \in \mathcal{F} : x_N = 0\}$ und $s_D = \{(y, s) \in \mathcal{F}_D : s_B = 0\}$.

Beweis: 33
Nach Satz 5.23 (Goldman-Tucker) gilt $B \cup N = \{1, \ldots, n\}$ und es gibt mindestens eine strikt komplementäre Lösung $(\tilde{x}, \tilde{y}, \tilde{s})$.

Sei $x \in \mathcal{F}_P$ mit $x_N = 0$ beliebig. Wegen $\tilde{s}_B = 0$ (siehe Satz 5.22) folgt $x_i \tilde{s}_i = 0$. Damit folgt weiter, dass $(x, \tilde{y}, \tilde{s})$ Lösung des Systems Gleichung 5.9 ist und dass $x \in S_P$ ist.

Sei umgekehrt $x \in S_P$ beliebig. Dann gilt $x_i \tilde{s}_i = 0$. Weiter ist $\tilde{s}_N > 0 \Rightarrow x_N = 0$. ∎

Bemerkung 5.26
Sei $(\tilde{x}, \tilde{y}, \tilde{s})$ eine primal duale Lösung und ist es strikt komplementär, dann folgt, dass $\tilde{s}_N, \tilde{x}_B > 0$. Hat das (LP) eine eindeutig bestimmte Lösung, dann ist $\tilde{x}$ Ecke von $\mathcal{F}_P$, d. h. es gibt eine Basis B, so dass $\tilde{x}$ zulässige Basislösung zur Basis B ist.

Sei J_B die Basisindizes. Sei N die Menge der Nichtbasisvektoren und J_N die zulässige Indexmenge. Da $\tilde{x}$ Basislösung zur Basis B ist, gilt $\widehat{x}_N = 0, J_N = N, J_B = B$. Aus dem Goldman-Tuckertheorem folgt, wenn die Lösung eindeutig bestimmt ist, muss die Lösung strikt komplementär sein. Wegen $\tilde{x}_B = \tilde{x}$ ist die Lösung nicht entartet. Ist $\tilde{s}_N > 0$, folgt, dass die Lösung von (LP) eindeutig bestimmt ist.

Wie ist hier der Index?

5.8. Zur Konvergenz des Verfahrens

Typische Voraussetzungen:

V1 $m < n$ und A hat vollen Rang m.

V2 Das primale Problem hat keine entartete Ecke.

V3 Das duale Problem hat keine entartete Basislösung.

Die Nebenbedingungen sind $A^T y + s = c$ bzw. $\tilde{A}\begin{pmatrix} y \\ s \end{pmatrix} = c$ und $\tilde{A} = (A^T, 1_n)$.

Sei B eine Basislösung von $\tilde{A}$. Eine Basislösung $z = \begin{pmatrix} y \\ s \end{pmatrix}$ zur Basis B heißt **nicht entartet**, falls gilt: $z_i \neq 0$ für alle $i \in J_B$.

Lemma 5.27
Sei obige Voraussetzung (V2) erfüllt. Dann gibt es ein $\varepsilon_1 > 0$, so dass gilt: Zu jedem $x \in \mathcal{F}_P$ gibt es eine Basis B von A mit $x_B > \varepsilon_1$.

Beweis: 34
Aus dem Satz 2.37 folgt, dass $\mathcal{F}_P$ endlich viele Ecken $v^{(1)}, \ldots, v^{(k)}$ hat. Wegen der Voraussetzung (V2) hat jedes $v^{(i)}$ mindestens m positive Komponenten. Wegen $v_N^{(i)} = 0_{n-m}$ folgt, dass $v^{(i)}$ hat genau m positive Komponenten hat. Wir definieren $\omega := \min\{v_j^{(i)} : v_j^{(i)} > 0, i = 1, \ldots, k, j = 1, \ldots, n\}$. Sei nun $x \in \mathcal{F}_P$ beliebig. Nach Gleichung 2.1 kann man x in folgender Form darstellen:

$$x = \sum_{i=1}^{k} \alpha_i v^{(i)} + d \qquad \alpha_i \geq 0, i = 1, \ldots, k, \sum_{i=1}^{k} \alpha_i = 1, d \geq 0, Ad = 0$$

Sei $l \in \{1, \ldots, k\}$ ein Index mit $\alpha_l = \max_{1 \leq i \leq k} \alpha_i$. Dann ist $\alpha_l \geq 1/k \Rightarrow x = \sum_{i=1}^{k} \alpha_i v^{(i)} + d \geq \alpha_l(v^{(l)}) \geq 1/k v^{(l)}$. Nach der Definition von ω ist $v_B^{(l)} \geq \omega \Rightarrow x_B \geq 1/k v^{(l)} \geq \omega/k \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} =: \varepsilon_1 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$. ∎

Satz 5.28 (Konvergenzsatz 1)
Es gelten die Voraussetzungen, dass $m < n$, A vollen Rang m und das duale Problem keine entartete Basislösung hat. Dann konvergiert die von der Short-Step-Variante des Verfahrens berechnete Folge $\{x^{(k)}, y^{(k)}, s^{(k)}\}$ mit $\tilde{s}^{(k)} = x_k^{-1} s^{(k)}$ gegen eine strikt komplementäre Lösung $(\tilde{x}, \tilde{y}, \tilde{s})$. Weiter gilt:

$$0 \leq c^T x^{(k+1)} - c^T \tilde{x} \leq \left(1 - \frac{3}{\sqrt{n - m + \delta_k}}\right)(c^T x^{(k)} - c^T \tilde{x})$$

wobei $\delta_k \xrightarrow{k \to \infty} 0$ gilt.

Satz 5.29 (Konvergenzsatz 2)
Es gelte, dass $m < n$, A vollen Rang m und das primale Problem keine entartete Ecke hat. Dann konvergiert die von der Long-Step-Variante berechnete Folge gegen eine strikt komplementäre Lösung und es gilt:

$$0 \leq c^T x^{(k+1)} - c^T \tilde{x} \leq \left(1 - \frac{13}{\sqrt{n}}\right) c^T (x^{(k)} - \tilde{)}$$

Bemerkung 5.30
Man kann auf die zweite Voraussetzung verzichten, wenn man $b \leq 2/3$ fordert.

5.8.1. Startpunktberechnung

Frage: Wie berechne ich einen Startpunkt $x^{(0)} \in \mathcal{F}_p^0$? Mit einer zusätzlichen Variable t und einem festen Punkt $z > 0_n$ betrachten wir das Hilfsproblem $\min_{x \in \mathbb{R}^n, t \in \mathbb{R}} t$ unter den Nebenbedingungen $Ax - t(Az - b)b, x, t \geq 0$. Es gilt, $(z, 1)$ ist zulässig für $(LP)_0$.

Sei $\mathcal{F}_p \neq 0$. Dann gilt für die Lösungsmenge $(LP)_0$:

$$S_0 = \{(x, 0) : x \in \mathcal{F}_p\} \tag{5.18}$$

Ist $Az = b$, dann wählen wir $x^{(0)} = z$, d. h. wir müssen $(LP)_0$ nur lösen, wenn $Az \neq t$ ist. Die Standardform von $(LP)_0$ ist:

$$\tilde{A} = (A, Az - b) \qquad \tilde{c} = \begin{pmatrix} 0_n \\ 1 \end{pmatrix}$$

Falls gilt: $\operatorname{rg} A = m \Rightarrow \operatorname{rg} \tilde{A} = m$.

Das Verfahren der affinen Skalierung berechnet eine strikt komplementäre Lösung $(\tilde{x}, \tilde{t}, \tilde{y}, \tilde{s})$, $\tilde{s} \in \mathbb{R}^{n+1}$. Falls $\tilde{t} > 0 \Rightarrow \mathcal{F}_p = \emptyset$.

Sei $(\tilde{x}, \tilde{t}, \tilde{y}, \tilde{s})$ eine primal-duale Lösung, die beispielsweise mit der Long-Step-Variante berechnet wurde. Ist $\tilde{t} > 0 \Rightarrow \mathcal{F}_p = \emptyset$. Sei $\tilde{t} = 0 \Rightarrow A\tilde{x} = b, \tilde{x} \geq 0$, d. h. $\tilde{x} \in \mathcal{F}_p$ und $\mathcal{B}^* = \{i : \tilde{x}_i > 0\}, \mathcal{N}^* = \{i : \tilde{x}_i = 0\}$.

1. Fall $\mathcal{N}^* = \emptyset \Rightarrow \tilde{x} > 0$
2. Fall $\mathcal{N}^* \neq \emptyset$. Seien $\mathcal{B}$ und $\mathcal{N}$ die durch Gleichung 5.16 definierten Mengen. Dann gilt $\tilde{t} = \tilde{x}_{n+1}, \mathcal{B}^* = \mathcal{B}$

A. Klausuraufgaben

(SS 2006)

1. Gegeben sei das Problem max $c^T x$ unter den Nebenbedigungen $Ax \leq b, x \in \mathbb{R}^n$.

   a) Transformieren Sie es in Standardform mit der zulässigen Menge $\mathcal{F}$.

   b) Zeigen Sie, dass für beschränktes $\mathcal{F}$ das Problem mind. eine Lösung besitzt.

   c) Welche wichtige Eigenschaft besitzt die Menge $\mathcal{F}$? Was versteht man unter einer „Ecke" der Menge $\mathcal{F}$? Geben Sie darauf aufbauend eine Charakterisierung jedes einzelnen Punktes aus $\mathcal{F}$ an.

   d) Sei nun $\mathcal{F}^* = \{x \in \mathcal{F} | b_l \leq x \leq b_u\}$ mit $b_l, b_u \in \mathbb{R}^n$ mit $b_l < b_u$ und komponentenweise geltenden Relationen. Welche Inklusion(en) ($\subset, \subseteq, \subsetneq, \supset, \supseteq, \supsetneq$) gilt bzw. gelten zwischen $\mathcal{F}$ und $\mathcal{F}^*$.

2. Die Niveaumenge einer Funktion $f\colon D(f) \to \mathbb{R}$ zum Niveau $\alpha \in \mathbb{R}$ ist gegeben durch

$$\mathcal{N}(f, \alpha) = \{x \in D(f) | f(x) \leq \alpha\}$$

   Zeigen Sie, dass die Niveaumenge einer *konvexen Funktion* ebenfalls konvex ist, vorausgesetzt sie ist nichtleer. Definieren Sie dafür zunächst den Begriff einer *konvexen Menge* und einer *konvexen Funktion*.

3. Betrachten Sie das Problem

$$\begin{aligned}&\min c^T x\\ &\text{Nb.} Ax = b, \quad 0 \leq x \in \mathbb{R}^n\end{aligned}$$

   oder äquivalent

$$\begin{aligned}&\min c^T x\\ &\text{Nb.} \sum_{i=1}^{n} A_{.i} x_i = b, \quad 0 \leq x \in \mathbb{R}^n\end{aligned}$$

   wobei $A_{.i}$ die i-te Spalte der Matrix A ist.

   Was versteht man unter einer Basis und einer Basislösung? Wann ist eine Basislösung zulässig?

4. Lineare Optimierungsprobleme spielen bei der Auswertung von Messungen eine wichtige Rolle. Wir betrachten hierzu die folgende Aufgabenstellung. Es seien Messpunkte $(x_i, y_i) \in \mathbb{R}^2, i = 1, \ldots, n$, gegeben. Zwischen den x- und den y-Werten wird ein linearer Zusammenhang vermutet. Daher soll zu den Messpunkten eine „optimale" Gerade bestimmt werden. Benutzen Sie als Optimalitätskriterium die Summe der Abweichung an den Messpunkten. Definiert man eine Gerade $g(x) = ax + b$ durch die beiden Parameter a und b, dann ist eine Paar $(a, b) \in \mathbb{R}^2$ gesucht, so dass

$$\sum_{i=1}^{n} |ax_i + b - y_i|$$

minimal wird, d. h. es ist das Optimierungsproblem

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} |ax_i + b - y_i|$$

zu lösen.

   a) Formulieren Sie diese Aufgabe als lineares Problem in kanonischer Form.

   b) Transformieren Sie das Problem in Standardform.

   c) Welcher Zusammenhang besteht zwischen den zulässigen Mengen bzw. den Ecken der Mengen der beiden Probleme.

   d) Geben Sie das duale Problem an.

   e) Stellen Sie das Hilfsproblem für die Phase 1 des Simplex-Verfahrens auf.

5. Gegeben sei mit einem $\varepsilon \in \mathbb{R} \setminus \{0\}$ das lineare Optimierungsproblem

$$\begin{aligned} \max\, & 4x_1 + 5x_2 + 9x_3 + 18x_4 && \\ \text{Nb.}\, & x_1 + x_2 + x_3 + \varepsilon x_4 && = 15 \\ & 7x_1 + 5x_2 + 3x_3 && \leq 120 \\ & 3x_1 + 5x_2 + 10x_3 && \leq 100 \\ & x_i \geq 0 && i = 1, \ldots, 4 \end{aligned}$$

   a) Transformieren Sie das Problem in Standardform!

   b) Bestimmen Sie eine Basis. Für welche ε ist die zugehörige Basislösung zulässig? Berechnen Sie in diesem Fall das Starttableau zu dieser Basis.

   c) Bestimmen Sie alle ε, für welche die Basislösung optimal ist.

Index