

Innere-Punkte-Verfahren

Neben den diversen Varianten des Simplex-Verfahrens gibt es eine zweite große Klasse von Verfahren zur Lösung von linearen Programmen, nämlich die Innere-Punkte-Verfahren. Wir beginnen in Abschn. 1 mit der Wiederholung wichtiger Grundlagen sowie Bemerkungen zum Vergleich zwischen Simplex-Verfahren und Innere-Punkte-Verfahren. Anschließend stellen wir in Abschn. 2 das Kurz-Schritt-Verfahren vor, welches bereits alle grundlegenden Ideen weiterführender Innere-Punkte-Verfahren verwendet. In der Praxis kommt das Kurz-Schritt-Verfahren aufgrund ungünstiger Konvergenzeigenschaften jedoch selten zum Einsatz. Daher erweitern wir das Verfahren in Abschn. 3 zum heuristischen Prädiktor-Korrektor-Verfahren, welches theoretisch gesehen schlechte Eigenschaften besitzt, in vielen praktischen Anwendungen jedoch äußerst effizient eingesetzt werden kann. Eine signifikante Reduktion der Laufzeit beider Verfahren kann erzielt werden, indem auftretende lineare Gleichungssysteme aufgrund ihrer speziellen Struktur geeignet gelöst werden. Ergebnisse hierzu fassen wir in Abschn. 4 zusammen. Abschließend präsentieren wir in Abschn. 5 als Anwendungsbeispiel Zwei-Personen Nullsummenspiele aus der Spieltheorie, welche effizient mit dem Prädiktor-Korrektor-Verfahren gelöst werden können.

1 Grundlagen

In diesem Kapitel leiten wir *Innere-Punkte-Verfahren* zur Lösung von linearen Programmen in Standardform her. Dazu werden Grundlagen der linearen Optimierung vorausgesetzt, dennoch wiederholen wir zentrale Definitionen und Ergebnisse, damit auch dieses Kapitel in sich schlüssig ist.

Im Vergleich zwischen Simplex-Verfahren und Innere-Punkte-Verfahren sind grundsätzlich folgende Hinweise zu beachten: Das Simplex-Verfahren ist unter Anwendung geeigneter Pivot-Regeln stets endlich und terminiert mit einer Optimallösung bzw. der Aussage, dass das Problem unzulässig oder unbeschränkt ist. Allerdings

gibt es spezielle Beispiele, welche eine in der Anzahl der Variablen exponentielle Laufzeit vorhersagen. Im Vergleich dazu lässt sich zeigen, dass Innere-Punkte-Verfahren in aller Regel eine polynomielle Laufzeit haben. Andererseits existieren Beispiele, in denen die Verfahren nicht oder nur sehr langsam konvergieren. Somit kann keine klare Antwort gegeben werden, ob Simplex-Varianten oder Innere-Punkte-Verfahren in der Praxis effizienter sind. Genauer hängt es stark von der Struktur der Eingabedaten ab, ob das eine oder das andere Verfahren bevorzugt werden sollte.

Alle wesentlichen Ideen und Herleitung aus diesem Kapitel können in üblichen Lehrbüchern zur Optimierung bzw. zum Operations Research vertieft werden, etwa in Jarre und Stoer (2004) oder Domschke et al. (2015). In Hamacher und Klamroth (2006) wird hingegen ein anderer Ansatz zur Herleitung von Inneren-Punkte-Verfahren verfolgt, nämlich ein Ansatz unter Verwendung der projektiven Geometrie, wie es in Karmarkar (1984) als eines der ersten Innere-Punkte-Verfahren vorgeschlagen wurde.

=

Definition 1.1 Ein lineares Programm in Standardform wird gegeben durch

$$\min c^\top \cdot x, \quad \text{sodass} \quad A \cdot x = b \quad \text{und} \quad x \geq 0. \quad (1)$$

Dabei sind $c \in \mathbb{R}^n$ und $b \in \mathbb{R}^m$ Vektoren sowie $A \in \mathbb{R}^{m \times n}$ eine Matrix.

Wir werden im Folgenden stets die Annahme treffen, dass der Rang der Matrix A gleich m ist und dass $n \geq m$ gilt. Diese Annahmen sind plausibel, denn anderenfalls können wir das Gleichungssystem $A \cdot x = b$ mittels Anwendung elementarer Zeilenoperationen in ein äquivalentes System überführen, welchen den Annahmen genügt.

Es gibt unterschiedliche Möglichkeiten zur Motivation und Herleitung von Inneren-Punkte-Verfahren. Wir wählen in diesem Kapitel den Weg über die Dualität und wiederholen daher folgende Definition, wobei $y \leq 0$ eine in der Optimierung üblichen Schreibweise für im Vorzeichen unbeschränkte Variablen $y \in \mathbb{R}^m$ ist:

=

Definition 1.2 Gegeben sei ein lineares Programm in Standardform, d.h.

$$\min c^\top \cdot x, \quad \text{sodass} \quad A \cdot x = b \quad \text{und} \quad x \geq 0 \quad (2)$$

mit $c \in \mathbb{R}^n$, mit $b \in \mathbb{R}^m$ sowie mit $A \in \mathbb{R}^{m \times n}$. Dabei gelte $n \geq m$ und der Rang von A sei gleich m . Das zugehörige **duale** lineare Programm lautet

$$\max b^\top \cdot y, \quad \text{sodass} \quad A^\top \cdot y + s = c \quad \text{und} \quad s \geq 0, y \leq 0. \quad (3)$$

Die beiden linearen Programme werden als **Primales** und **Duales** bezeichnet.

Dabei wurde eine leicht andere aber dennoch identische Darstellung des Dualen im Vergleich zur Definition aus dem Kapitel über Simplex-Verfahren verwendet. Weiterhin wiederholen wir folgende Aussage zur starken Dualität bzw. zum komplementären Schlupf:

Satz 1.1 Gegeben sei ein lineares Programm in Standardform sowie das zugehörige duale lineare Programm. Weiter sei x zulässig für das Primale und y zulässig für das Duale. Dann gilt

$$x^\top \cdot (c - A^\top \cdot y) = x^\top \cdot s = 0$$

genau dann, wenn x eine Optimallösung des Primales und y eine Optimallösung des Dualen ist.

Wir haben damit bereits die wesentlichen Grundlagen zur Herleitung der Inneren-Punkte-Verfahren wiederholt und führen noch folgende Notation ein, bevor wir die Ergebnisse zusammenfassen:

Notation 1.3 Zu einem Vektor $x \in \mathbb{R}^n$ und einem Vektor $s \in \mathbb{R}^n$ definieren wir die Diagonalmatrizen

$$X = \text{diag}(x) \in \mathbb{R}^{n \times n} \quad \text{und} \quad S = \text{diag}(s) \in \mathbb{R}^{n \times n}.$$

Die Matrizen X und S sind regulär, falls $x > 0$ und $s > 0$ gilt.

Die Ergebnisse und Bezeichnungen zuvor liefern schließlich folgende Aussage:

Satz 1.2 Gegeben sei ein lineares Programm in Standardform, d.h.

$$\min c^\top \cdot x, \quad \text{sodass} \quad A \cdot x = b \quad \text{und} \quad x \geq 0$$

mit $c \in \mathbb{R}^n$, mit $b \in \mathbb{R}^m$ sowie mit $A \in \mathbb{R}^{m \times n}$. Dabei gelte $n \geq m$ und der Rang von A sei gleich m . Weiterhin nehmen wir an, dass eine beschränkte Optimallösung existiert.

Falls $(x, y, s) \in \mathbb{R}^{2n+m}$ das nichtlineare Gleichungssystem

$$g(x, y, s) = \begin{pmatrix} A \cdot x - b \\ A^\top \cdot y + s - c \\ X \cdot s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (4)$$

löst und zusätzlich $x \geq 0$ sowie $s \geq 0$ gilt, dann ist x eine Optimallösung

für das Primale und y eine Optimallösung für das Duale. Dabei ist X die Diagonalmatrix zum Vektor x .

Damit sind alle Grundlagen wiederholt, sodass wir in den folgenden Abschnitten Innere-Punkte-Verfahren herleiten können.

2 Kurz-Schritt-Verfahren

Die zentrale Idee aller Inneren-Punkte-Verfahren ist folgende: Es wird eine Nullstelle der Funktion $g(x, y, s)$ aus Gl. (4) unter Anwendung des Newton-Verfahrens zur Lösung von nichtlinearen Gleichungssystem bestimmt, sodass damit gleichzeitig eine Lösung x des Primales sowie eine Lösung y des Dualen berechnet wird. Allerdings ergeben sich dabei beachtliche Probleme:

- (1) Zunächst muss $x \geq 0$ sowie $s \geq 0$ gelten, damit es sich bei der Lösung des Gleichungssystem tatsächlich auch um Lösungen der linearen Programme handelt. Diese Bedingungen berücksichtigt das Newton-Verfahren im Allgemeinen nicht.
- (2) Darüber hinaus ist die zur Anwendung des Newton-Verfahrens benötigte Jacobi-Matrix $Dg(x, y, s)$ im Allgemeinen nur dann regulär, falls $x > 0$ sowie $s > 0$ gilt. Details zu dieser Aussage werden wir später genauer formulieren.

Insbesondere die Bedingungen $x > 0$ und $s > 0$ widersprechen den Gleichungen

$$X \cdot s = 0.$$

Um dieses Problem zu umgehen, führen wir einen Parameter $\mu > 0$ ein und betrachten anstelle von (4) das nichtlineare Gleichungssystem

$$f(x, y, s) = \begin{pmatrix} A \cdot x - b \\ A^\top \cdot y + s - c \\ X \cdot s - \mu \cdot e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (5)$$

mit $e = (1, \dots, 1) \in \mathbb{R}^n$. Somit konvergiert die Funktion $f(x, y, s)$ gegen $g(x, y, s)$, je kleiner μ gewählt wird.

Jede Nullstelle von $f(x, y, s)$ erfüllt damit zunächst einmal die Bedingung, dass keine Einträge der Vektoren $x \in \mathbb{R}^n$ und $s \in \mathbb{R}^n$ gleich null sind. Daher stammt schließlich auch der Name der Inneren-Punkte-Verfahren: Wir werden iterative Verfahren herleiten, bei denen wir in jedem Schritt aufgrund von $\mu > 0$ Lösungen im Inneren der Mengen

$$\{x \geq 0 : A \cdot x = b\} \quad \text{und} \quad \{s \geq 0 : A^\top \cdot y + s = c\} \quad (6)$$

erhalten. Um diesem Ziel näher zu kommen, untersuchen wir die Jacobi-Matrix von $f(x, y, s)$: Mit der $(n \times n)$ -Einheitsmatrix I_n erhalten wir als Block-Matrix

$$Df(x, y, s) = \begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I_n \\ S & 0 & X \end{pmatrix} \in \mathbb{R}^{(2n+m) \times (2n+m)}. \quad (7)$$

Dabei ist zu beachten, dass y keinen Einfluss auf $Df(x, y, s)$ nimmt, aufgrund der Definition der Matrizen X und S hängt $Df(x, y, s)$ jedoch von x und s ab. Zur Anwendung des Newton-Verfahrens wird eine reguläre Jacobi-Matrix benötigt. Hierzu gilt das bereits zuvor erwähnte Ergebnis:

Satz 2.1 Gegeben sei eine Matrix $A \in \mathbb{R}^{m \times n}$. Dabei gelte $n \geq m$ und der Rang von A sei gleich m . Weiter seien $x \in \mathbb{R}^n$ und $s \in \mathbb{R}^m$ zwei Vektoren mit $x > 0$ und $s > 0$. Unter diesen Voraussetzungen ist die Jacobi-Matrix

$$Df(x, y, s) = \begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I_n \\ S & 0 & X \end{pmatrix} \in \mathbb{R}^{(2n+m) \times (2n+m)}$$

regulär.

Herleitung Wir haben damit eine Aussage, unter welchen Voraussetzungen ein nächster Iterationsschritt des Newton-Verfahrens überhaupt durchgeführt werden kann. Genauer sei

$$(x_{(0)}, y_{(0)}, s_{(0)}) \in \mathbb{R}^{(2n+m)}$$

eine erste Approximation einer Nullstelle von $f(x, y, s)$ mit $x_{(0)} > 0$ und $s_{(0)} > 0$. Dann erhalten wir im nächsten Schritt des Newton-Verfahrens

$$(x_{(1)}, y_{(1)}, s_{(1)}) = (x_{(0)}, y_{(0)}, s_{(0)}) + (x_{(0)}^N, y_{(0)}^N, s_{(0)}^N) \in \mathbb{R}^{(2n+m)},$$

wobei $(x_{(0)}^N, y_{(0)}^N, s_{(0)}^N)$ die Lösung des linearen Gleichungssystems

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I_n \\ S_{(0)} & 0 & X_{(0)} \end{pmatrix} \cdot \begin{pmatrix} x_{(0)}^N \\ y_{(0)}^N \\ s_{(0)}^N \end{pmatrix} = \begin{pmatrix} b - A \cdot x_{(0)} \\ c - A^\top \cdot y_{(0)} - s_{(0)} \\ \mu \cdot e - X_{(0)} \cdot s_{(0)} \end{pmatrix} = -f \begin{pmatrix} x_{(0)} \\ y_{(0)} \\ s_{(0)} \end{pmatrix}$$

ist. Dabei ist $X_{(0)}$ die Diagonalmatrix zum Vektor $x_{(0)}$ und analog $S_{(0)}$ die Diagonalmatrix zum Vektor $s_{(0)}$. Dank der Voraussetzungen $x_{(0)} > 0$ und $s_{(0)} > 0$ ist dieses Gleichungssystem nach Satz 2.1 eindeutig lösbar.

Allerdings können die für den folgenden Schritt des Newton-Verfahrens notwendigen Bedingungen $x_{(1)} > 0$ und $s_{(1)} > 0$ im Allgemeinen nicht sichergestellt werden.

Um diese Voraussetzungen dennoch zu gewährleisten, werden wir später einige Annahme an die Startlösung $(x_{(0)}, y_{(0)}, s_{(0)})$ stellen. Zudem wird kein fester Parameter μ gewählt, sondern μ wird von Schritt zu Schritt verkleinert, sodass wir schließlich gegen eine Nullstelle der Funktion $g(x, y, s)$ aus Gl. (4) konvergieren, für welche $x \geq 0$ und $s \geq 0$ gilt.

Wir erhalten damit ein iteratives Verfahren, welches unter gewissen Voraussetzungen (die wir in der folgenden Zusammenfassung im Detail formulieren werden) gegen eine Optimallösung des linearen Programms konvergiert. Dabei bilden die Lösungen

$$(x_{(k)}, y_{(k)}, s_{(k)}) \in \mathbb{R}^{(2n+m)}$$

der einzelnen Iterationen den sogenannten **zentralen Pfad**, welcher sich dank $\mu_{(k)} > 0$ stets im Inneren der zulässigen Mengen \mathcal{G} befindet.

?

Aufgabe 2.1 Angenommen, $(x_{(1)}, y_{(1)}, s_{(1)})$ ist die Approximation nach dem ersten Schritt des Newton-Verfahrens. Zeige, dass dann

$$A \cdot x_{(1)} = b \quad \text{sowie} \quad A^\top \cdot y_{(1)} + s_{(1)} = c$$

gilt. Welche Aussage dazu kann über die folgenden Schritte getroffen werden?

Als Ergebnis der Herleitung zuvor fassen wir bereits eine grundlegende Version von Inneren-Punkte-Verfahren zusammen. Dabei geben wir ohne weitere Herleitung Voraussetzungen an die Startlösung sowie eine Vorschrift zur Verwendung von $\mu_{(k)}$ an, sodass das Verfahren terminiert und konvergiert. Weitere Details und Herleitungen dazu können beispielsweise in Jarre und Stoer (2004) nachgeschlagen werden. Zudem werden wir erst in Abschn. 4 im Detail darauf eingehen, wie das (große) lineare Gleichungssystem aus Schritt (1) des Verfahrens effizient gelöst werden kann.

#

Zusammenfassung 2.1 (Kurz-Schritt-Verfahren) Gegeben sei ein lineares Programm in Standardform, d.h.

$$\min c^\top \cdot x, \quad \text{sodass} \quad A \cdot x = b \quad \text{und} \quad x \geq 0$$

mit $c \in \mathbb{R}^n$, mit $b \in \mathbb{R}^m$ sowie mit $A \in \mathbb{R}^{m \times n}$. Dabei gelte $n \geq m$ und der Rang von A sei gleich m . Weiterhin nehmen wir an, dass eine beschränkte Optimallösung existiert. Darüber hinaus seien

$$x_{(0)} > 0, \quad y_{(0)} \leq 0, \quad s_{(0)} > 0 \quad \text{und} \quad \mu_{(0)} > 0$$

derart gewählt, sodass

$$A \cdot x_{(0)} = b \quad \text{und} \quad A^\top \cdot y_{(0)} + s_{(0)} = c$$

gilt und weiter die Bedingung

$$\left\| X_{(0)} \cdot s_{(0)} - \mu_{(0)} \cdot e \right\|_2 \leq \frac{\mu_{(0)}}{2} \quad (8)$$

erfüllt wird. Nun setze $k = 0$, definiere eine Abbruchgenauigkeit $\varepsilon > 0$ und führe folgende Schritte durch:

(1) Bestimme $(x_{(k)}^N, y_{(k)}^N, s_{(k)}^N)$ als Lösung des linearen Gleichungssystems

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I_n \\ S_{(k)} & 0 & X_{(k)} \end{pmatrix} \cdot \begin{pmatrix} x_{(k)}^N \\ y_{(k)}^N \\ s_{(k)}^N \end{pmatrix} = \begin{pmatrix} b - A \cdot x_{(k)} \\ c - A^\top \cdot y_{(k)} - s_{(k)} \\ \mu_{(k)} \cdot e - X_{(k)} \cdot s_{(k)} \end{pmatrix}$$

und setze

$$(x_{(k+1)}, y_{(k+1)}, s_{(k+1)}) = (x_{(k)}, y_{(k)}, s_{(k)}) + (x_{(k)}^N, y_{(k)}^N, s_{(k)}^N).$$

(2) Falls $\mu_{(k)} \leq \varepsilon$, beende das Verfahren. Anderenfalls setze

$$\mu_{(k+1)} = \mu_{(k)} \cdot \left(1 - \frac{1}{6 \cdot \sqrt{n}} \right), \quad (9)$$

erhöhe k um 1 und gehe wieder zu Schritt (1).

Das Verfahren terminiert mit $(x_{(k+1)}, y_{(k+1)}, s_{(k+1)})$, wobei $x_{(k+1)}$ eine Approximation einer Optimallösung des linearen Programms ist.

Das Verfahren aus der Zusammenfassung zuvor wird **Kurz-Schritt-Verfahren** genannt, da sich $(x_{(k)}, y_{(k)}, s_{(k)})$ insbesondere für größere n nur mit einer kleinen Schrittweite entlang des zentralen Pfades bewegt. Weiterhin kann die rechte Seite des Gleichungssystems auch vereinfacht werden, da für alle k

$$b - A \cdot x_{(k)} = 0 \quad \text{und} \quad c - A^\top \cdot y_{(k)} - s_{(k)} = 0$$

gilt (s. auch Aufgabe 2.1). Das Kurz-Schritt-Verfahren wird daher auch als zulässiges Innere-Punkte-Verfahren bezeichnet, da wir in jedem Schritt stets zulässige Lösungen erhalten.

Darüber hinaus sei bemerkt, dass sowohl die Voraussetzung (8) als auch die Berechnung von $\mu_{(k)}$ gemäß der Vorschrift (9) nicht ohne Weiteres nachvollzogen werden können. Diese Bedingungen werden jedoch benötigt, um die Konvergenz des Verfahrens zu gewährleisten, s. Jarre und Stoer (2004).

Bevor wir uns ein erstes Beispiel ansehen, schließen wir noch folgende Aussage zur Komplexität bzw. Konvergenzgeschwindigkeit des Verfahrens an:



Satz 2.2 *Unter den Voraussetzungen aus Zusammenfassung 2.1 terminiert das Kurz-Schritt-Verfahren nach spätestens*

$$6 \cdot \log\left(\frac{\mu(0)}{\varepsilon}\right) \cdot \sqrt{n} \quad (10)$$

Iterationsschritten.

Das Besondere an der Aussage zuvor ist, dass das Kurz-Schritt-Verfahren damit eine in n polynomielle Komplexität besitzt, da jeder einzelne Iterationsschritt eine polynomielle Komplexität hat und die Anzahl der Schritte von der Größenordnung $\mathcal{O}(\sqrt{n})$ ist. Wir erinnern uns: Es gibt Beispiele, bei denen das Simplex-Verfahren eine in n exponentielle Komplexität besitzt.



Aufgabe 2.2 Bestimme die Komplexität eines einzelnen Iterationsschrittes des Kurz-Schritt-Verfahrens in Abhängigkeit von n und m .

Schließlich untersuchen wir ein Beispiel zum Kurz-Schritt-Verfahren:



Beispiel 2.1 Gegeben sei der Kostenvektor $c = (3, -5, 0, 2, 1) \in \mathbb{R}^5$ sowie die Matrix

$$A = \begin{pmatrix} 5 & -2 & 3 & -4 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ -2 & 5 & -3 & 1 & -1 \end{pmatrix} \in \mathbb{R}^{3 \times 5}$$

und der Vektor $b = (-6, 9, -3) \in \mathbb{R}^3$. Zur Anwendung des Kurz-Schritt-Verfahrens wählen wir die Startwerte

$$x_{(0)} = \begin{pmatrix} 1,12 \\ 2,44 \\ 4,62 \\ 6,56 \\ 5,66 \end{pmatrix}, \quad y_{(0)} = \begin{pmatrix} -3,49 \\ -14,03 \\ -1,52 \end{pmatrix}, \quad s_{(0)} = \begin{pmatrix} 17,41 \\ 9,65 \\ 5,91 \\ 3,59 \\ 2,97 \end{pmatrix}$$

sowie $\mu_{(0)} = 20$, welche sämtliche Voraussetzungen zur Anwendung des Kurz-Schritt-Verfahrens erfüllen.

Unter Verwendung von $\varepsilon = 10^{-6}$ terminiert das Verfahren nach 218 Iterationen mit einer Approximation einer Optimallösung, dessen Zielfunktionswert mit

$$-2,99999799$$

nur wenig vom Zielfunktionswert von -3 einer exakten Optimallösung abweicht.

Weiterhin sei erwähnt, dass die maximale Anzahl von Iterationen gemäß Gl. (10) durch 225 beschränkt ist. Dies ist nur unwesentlich mehr, als die in diesem Beispiel benötigten 218 Schritte.

Aufgabe 2.3 Prüfe, dass die Startwerte des Beispiels zuvor tatsächlich alle Voraussetzungen zur Anwendung des Kurz-Schritt-Verfahrens aus Zusammenfassung 2.1 erfüllen, d.h. prüfe insbesondere Voraussetzung (8).

Wie im Beispiel zuvor ist es im Allgemeinen leider alles andere als einfach, Startwerte zu finden, welche die Voraussetzungen zur Anwendung des Kurz-Schritt-Verfahrens erfüllen. Darüber hinaus ist die Konvergenzgeschwindigkeit des Verfahrens häufig vergleichsweise schlecht, da ähnlich dem Beispiel zuvor meist nahezu die maximale Anzahl von Iterationen gemäß Gl. (10) benötigt wird, bis das Verfahren terminiert.

3 Prädiktor-Korrektor-Verfahren

Um eines der Hauptprobleme des Kurz-Schritt-Verfahrens zu umgehen, nämlich das Finden von zulässigen Startwerten, werden wir den Algorithmus in diesem Abschnitt erweitern. Dazu sei bereits vorweggenommen, dass das folgende Verfahren zumindest theoretisch nicht immer konvergiert. Es handelt sich also um ein heuristisches Verfahren, welches in der Praxis häufig sehr gute Konvergenzeigenschaften besitzt, andererseits aus theoretischer Sicht keine sehr guten Eigenschaften hat, so dass beispielsweise die Konvergenz nicht für alle Eingabedaten sichergestellt werden kann.

In der folgenden Herleitung orientieren wir uns an Mehrotra (1992). Darin wurde das Verfahren erstmals vorgestellt, es kann jedoch in ähnlicher Art und Weise auch in Jarre und Stoer (2004) nachgeschlagen werden.

Herleitung Ähnlich zum Kurz-Schritt-Verfahren ist es das Ziel, eine Nullstelle von

$$f(x, y, s) = \begin{pmatrix} A \cdot x - b \\ A^\top \cdot y + s - c \\ X \cdot s - \mu \cdot e \end{pmatrix}$$

mit $x > 0$ und $s > 0$ für möglichst kleine Werte von μ unter Verwendung des Newton-Verfahrens zu approximieren. Dazu starten wir mit

$$(x_{(0)}, y_{(0)}, s_{(0)}) \in \mathbb{R}^{(2n+m)},$$

wobei $x_{(0)} > 0$ und $s_{(0)} > 0$ gelte. Anders als beim Kurz-Schritt-Verfahren müssen die Startwerte nicht zulässig sein, d.h. die Bedingungen

$$A \cdot x_{(0)} = b \quad \text{und} \quad A^\top \cdot y_{(0)} + s_{(0)} = c$$

müssen nicht erfüllt sein. Das Verfahren gliedert sich nun in zwei Schritte, nämlich in den **Prädiktor-Schritt** und den **Korrektor-Schritt**.

Unter Verwendung von $\mu = 0$ lösen wir im Prädiktor-Schritt analog zum Kurz-Schritt-Verfahren zunächst das lineare Gleichungssystem

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I_n \\ S_{(0)} & 0 & X_{(0)} \end{pmatrix} \cdot \begin{pmatrix} x_{(0)}^N \\ y_{(0)}^N \\ s_{(0)}^N \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -X_{(0)} \cdot s_{(0)} \end{pmatrix},$$

dessen Lösung $(x_{(0)}^N, y_{(0)}^N, s_{(0)}^N)$ der Newton-Richtung entspricht. Allerdings wollen wir die Bedingungen $x \geq 0$ und $s \geq 0$ weiterhin bewahren und setzen daher

$$(x_{(0)}^P, y_{(0)}^P, s_{(0)}^P) = (x_{(0)}, y_{(0)}, s_{(0)}) + (\alpha_{(0)}^x \cdot x_{(0)}^N, \alpha_{(0)}^s \cdot y_{(0)}^N, \alpha_{(0)}^s \cdot s_{(0)}^N)$$

unter Verwendung von

$$\alpha_{(0)}^x = \min \left\{ 1, \min \left\{ -\frac{x_{(0),j}}{x_{(0),j}^N} : j = 1, \dots, n \text{ mit } x_{(0),j}^N < 0 \right\} \right\} \in \mathbb{R},$$

$$\alpha_{(0)}^s = \min \left\{ 1, \min \left\{ -\frac{s_{(0),j}}{s_{(0),j}^N} : j = 1, \dots, n \text{ mit } s_{(0),j}^N < 0 \right\} \right\} \in \mathbb{R}.$$

Dabei sei $x_{(0),j}$ der j -te Eintrag des Vektors $x_{(0)} \in \mathbb{R}^n$ und analog für $s_{(0)} \in \mathbb{R}^n$. Da y im Vorzeichen unbeschränkt ist, wird hier keine eigene Schrittweite berechnet, sondern die Schrittweite von s verwendet, denn y und s werden zusammen in einer Gleichung behandelt. Dank dieser Wahl von $\alpha_{(0)}^x$ sowie $\alpha_{(0)}^s$ gilt schließlich

$$x_{(0)}^P \geq 0 \quad \text{und} \quad s_{(0)}^P \geq 0.$$

Mit $(x_{(0)}^P, y_{(0)}^P, s_{(0)}^P)$ haben wir nun allerdings nur einen Schätzer für den nächsten Schritt gefunden, welcher die Zulässigkeitsbedingungen keineswegs erfüllen muss.

Daher schließt sich nun der Korrektor-Schritt an: Zunächst setzen wir

$$\mu_{(0)} = \frac{1}{n} \cdot \frac{\left((x_{(0)}^P)^\top \cdot s_{(0)}^P \right)^2}{\left(x_{(0)}^\top \cdot s_{(0)} \right)^3} > 0.$$

Diese Gleichung zur Berechnung von $\mu_{(0)}$ lässt sich im Kern nicht herleiten. Es handelt sich hierbei um reine Erfahrungswerte, d.h. die Konvergenz des Verfahrens ist durch diese Wahl von $\mu_{(0)}$ in praktischen Anwendungen vergleichsweise gut. Nun lösen wir das lineare Gleichungssystem

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I_n \\ S_{(0)} & 0 & X_{(0)} \end{pmatrix} \cdot \begin{pmatrix} x_{(0)}^C \\ y_{(0)}^C \\ s_{(0)}^C \end{pmatrix} = \begin{pmatrix} b - A \cdot x_{(0)} \\ c - A^\top \cdot y_{(0)} - s_{(0)} \\ \mu_{(0)} \cdot e - X_{(0)} \cdot s_{(0)} - X_{(0)}^N \cdot s_{(0)}^N \end{pmatrix},$$

mit der Diagonalmatrix $X_{(0)}^N$ zum Vektor $x_{(0)}^N$. Durch den zusätzlichen Term

$$-X_{(0)}^N \cdot s_{(0)}^N$$

wird dabei versucht, den Fehler des Newton-Schrittes aus dem Prädiktor-Schritt zu kompensieren. Schließlich setzen wir

$$(x_{(1)}, y_{(1)}, s_{(1)}) = (x_{(0)}, y_{(0)}, s_{(0)}) + \left(\beta_{(0)}^x \cdot x_{(0)}^C, \beta_{(0)}^s \cdot y_{(0)}^C, \beta_{(0)}^s \cdot s_{(0)}^C \right)$$

unter Verwendung von

$$\beta_{(0)}^x = \min \left\{ 1, \gamma \cdot \min \left\{ -\frac{x_{(0),j}}{x_{(0),j}^C} : j = 1, \dots, n \text{ mit } x_{(0),j}^C < 0 \right\} \right\} \in \mathbb{R},$$

$$\beta_{(0)}^s = \min \left\{ 1, \gamma \cdot \min \left\{ -\frac{s_{(0),j}}{s_{(0),j}^C} : j = 1, \dots, n \text{ mit } s_{(0),j}^C < 0 \right\} \right\} \in \mathbb{R}.$$

Dabei ist $\gamma \in (0, 1)$ ein zuvor gewählter Sicherheitsfaktor, etwa $\gamma = 0,95$. Durch die Wahl von $\beta_{(0)}^x$ sowie $\beta_{(0)}^s$ werden schließlich wieder die Bedingungen

$$x_{(1)} > 0 \quad \text{und} \quad s_{(1)} > 0$$

bewahrt. Alle Schritte zuvor können nun iterativ wiederholt werden, bis geeignete Abbruchbedingungen erfüllt sind.

Wir fassen das Verfahren nun ausführlich zusammen. Dazu sei nochmals ausdrücklich darauf hingewiesen, dass das Verfahren in vielen praktischen Anwendungsfällen sehr gute Konvergenzeigenschaften besitzt. Allerdings gibt es auch Beispiele, in den das Verfahren nur sehr langsam gegen eine Optimallösung konvergiert oder auch gar nicht terminiert.

Zusammenfassung 3.1 (Prädiktor-Korrektor-Verfahren) Gegeben sei ein lineares Programm in Standardform, d.h.

$$\min c^\top \cdot x, \quad \text{sodass} \quad A \cdot x = b \quad \text{und} \quad x \geq 0$$

mit $c \in \mathbb{R}^n$, mit $b \in \mathbb{R}^m$ sowie mit $A \in \mathbb{R}^{m \times n}$. Dabei gelte $n \geq m$ und der Rang von A sei gleich m . Weiterhin nehmen wir an, dass eine beschränkte Optimallösung existiert. Darüber hinaus seien beliebige Startwerte

$$x_{(0)} > 0, \quad y_{(0)} \leq 0 \quad \text{und} \quad s_{(0)} > 0$$

gegeben. Nun setze $k = 0$, definiere eine Abbruchgenauigkeit $\varepsilon > 0$ sowie einen Sicherheitsfaktor $\gamma \in (0, 1)$ und führe folgende Schritte durch:

(1) Bestimme $(x_{(k)}^N, y_{(k)}^N, s_{(k)}^N)$ als Lösung des linearen Gleichungssystems

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I_n \\ S_{(k)} & 0 & X_{(k)} \end{pmatrix} \cdot \begin{pmatrix} x_{(k)}^N \\ y_{(k)}^N \\ s_{(k)}^N \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -X_{(k)} \cdot s_{(k)} \end{pmatrix}.$$

#

(2) Berechne die Werte

$$\alpha_{(k)}^x = \min \left\{ 1, \min \left\{ -\frac{x_{(k),j}^{(k)}}{x_{(k),j}^N} : j = 1, \dots, n \text{ mit } x_{(k),j}^N < 0 \right\} \right\},$$

$$\alpha_{(k)}^s = \min \left\{ 1, \min \left\{ -\frac{s_{(k),j}^{(k)}}{s_{(k),j}^N} : j = 1, \dots, n \text{ mit } s_{(k),j}^N < 0 \right\} \right\}$$

und setze

$$\left(x_{(k)}^P, y_{(k)}^P, s_{(k)}^P \right) = \left(x_{(k)}, y_{(k)}, s_{(k)} \right) + \left(\alpha_{(k)}^x \cdot x_{(k)}^N, \alpha_{(k)}^s \cdot y_{(k)}^N, \alpha_{(k)}^s \cdot s_{(k)}^N \right).$$

(3) Setze

$$\mu_{(k)} = \frac{1}{n} \cdot \frac{\left((x_{(k)}^P)^\top \cdot s_{(k)}^P \right)^2}{\left(x_{(k)}^\top \cdot s_{(k)} \right)}. \quad (11)$$

(4) Bestimme $(x_{(k)}^C, y_{(k)}^C, s_{(k)}^C)$ als Lösung des linearen Gleichungssystems

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I_n \\ S_{(k)} & 0 & X_{(k)} \end{pmatrix} \cdot \begin{pmatrix} x_{(k)}^C \\ y_{(k)}^C \\ s_{(k)}^C \end{pmatrix} = \begin{pmatrix} b - A \cdot x_{(k)} \\ c - A^\top \cdot y_{(k)} - s_{(k)} \\ r_{(k)}^x \end{pmatrix}$$

mit $r_{(k)}^x = \mu_{(k)} \cdot e - X_{(k)} \cdot s_{(k)} - X_{(k)}^N \cdot s_{(k)}^N$.

(5) Berechne die Werte

$$\beta_{(k)}^x = \min \left\{ 1, \gamma \cdot \min \left\{ -\frac{x_{(k),j}^{(k)}}{x_{(k),j}^C} : j = 1, \dots, n \text{ mit } x_{(k),j}^C < 0 \right\} \right\},$$

$$\beta_{(k)}^s = \min \left\{ 1, \gamma \cdot \min \left\{ -\frac{s_{(k),j}^{(k)}}{s_{(k),j}^C} : j = 1, \dots, n \text{ mit } s_{(k),j}^C < 0 \right\} \right\}$$

und setze

$$\begin{aligned} \left(x_{(k+1)}, y_{(k+1)}, s_{(k+1)} \right) &= \left(x_{(k)}, y_{(k)}, s_{(k)} \right) \\ &\quad + \left(\beta_{(k)}^x \cdot x_{(k)}^C, \beta_{(k)}^s \cdot y_{(k)}^C, \beta_{(k)}^s \cdot s_{(k)}^C \right). \end{aligned}$$

(6) Falls $\|X_{(k+1)} \cdot s_{(k+1)}\|_\infty < \varepsilon$ sowie

$$\|A \cdot x_{(k+1)} - b\|_\infty < \varepsilon \quad \text{und} \quad \|A^\top \cdot y_{(k+1)} + s_{(k+1)} - c\|_\infty < \varepsilon,$$

beende das Verfahren. Anderenfalls erhöhe k um 1 und gehe wieder zu Schritt (1).

Falls das Verfahren mit $(x_{(k+1)}, y_{(k+1)}, s_{(k+1)})$ terminiert, so ist $x_{(k+1)}$ eine Approximation einer Optimallösung des linearen Programms.

Geeignete Startwerte sind beispielsweise

$$x_{(0)} = e, \quad y_{(0)} = 0 \quad \text{und} \quad s_{(0)} = e.$$

Sind jedoch weitere Information über das lineare Programm vorhanden, können die Startwerte unter Umständen auch problemspezifisch sinnvoller gewählt werden. Weiterhin werden wir als Sicherheitsfaktor in allen Beispielen stets $\gamma = 0,95$ verwenden.

Falls das Verfahren nach einer vergleichsweise großen Anzahl an Iterationsschritten nicht terminiert oder falls $\|x_{(k)}\|_\infty$ bzw. $\|s_{(k)}\|_\infty$ übermäßig groß werden, so ist nicht mehr davon auszugehen, dass das Verfahren terminiert und es kann abgebrochen werden. Ist dies der Fall, dann kann das folgende Gründe haben:

- (1) Das lineare Programm ist unzulässig.
- (2) Das lineare Programm ist unbeschränkt.
- (3) Das Abbruchkriterium ε wurde ungeeignet (in der Regel zu klein) gewählt.

Es kann jedoch keine Aussage darüber getroffen werden, welcher dieser Gründe eingetreten ist.

Schließlich greifen wir das Beispiel aus dem Abschnitt zuvor wieder auf:

Beispiel 3.1 Gegeben sei der Kostenvektor $c = (3, -5, 0, 2, 1) \in \mathbb{R}^5$ sowie die Matrix

$$A = \begin{pmatrix} 5 & -2 & 3 & -4 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ -2 & 5 & -3 & 1 & -1 \end{pmatrix} \in \mathbb{R}^{3 \times 5}$$

und der Vektor $b = (-6, 9, -3) \in \mathbb{R}^3$. Zur Anwendung des Prädiktor-Korrektor-Verfahrens wählen wir die Startwerte $x_{(0)} = s_{(0)} = e$ und $y_{(0)} = 0$, d.h.

$$x_{(0)} = (1, 1, 1, 1, 1), \quad y_{(0)} = (0, 0, 0), \quad s_{(0)} = (1, 1, 1, 1, 1).$$

Unter Verwendung von $\varepsilon = 10^{-6}$ terminiert das Verfahren nach nur 8 Iterationen mit einer Approximation einer Optimallösung, dessen Zielfunktionswert mit

$$-2,999\,999\,79$$

nur wenig vom Zielfunktionswert von -3 einer exakten Optimallösung abweicht. Im Vergleich dazu: Das Kurz-Schritt-Verfahren benötigt im Beispiel aus dem Abschnitt zuvor bei vergleichbarer Genauigkeit 218 Iterationen.

4 Charakteristisches lineares Gleichungssystem

In den beiden Innere-Punkte-Verfahren aus den Abschnitten zuvor besteht in jedem Iterationsschritt die wesentliche Aufgabe darin, ein lineares Gleichungssystem der Form

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I_n \\ S & 0 & X \end{pmatrix} \cdot \begin{pmatrix} x^N \\ y^N \\ s^N \end{pmatrix} = \begin{pmatrix} r^b \\ r^c \\ r^x \end{pmatrix}$$

zu lösen. Mit $A \in \mathbb{R}^{m \times n}$ handelt es sich dabei um ein

$$(2 \cdot n + m) \times (2 \cdot n + m)$$

Gleichungssystem, sodass die Verfahren insbesondere für große n schnell ineffizient werden. Aufgrund der speziellen Struktur der Matrix kann das Lösen des Systems beispielsweise im Vergleich zum Gauß-Verfahren allerdings signifikant verbessert werden. Dies wollen wir in diesem Abschnitt vorstellen, s. auch Wright (1997):



Satz 4.1 Gegeben sei eine Matrix $A \in \mathbb{R}^{m \times n}$. Dabei gelte $n \geq m$ und der Rang von A sei gleich m . Weiter seien $x \in \mathbb{R}^n$ und $s \in \mathbb{R}^n$ zwei Vektoren mit $x > 0$ und $s > 0$ sowie mit den zugehörigen Diagonalmatrizen X und S . Zudem definieren wir

$$D = A \cdot S^{-1} \cdot X \cdot A^\top \in \mathbb{R}^{n \times n}.$$

Dann kann die Lösung des linearen Gleichungssystems

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I_n \\ S & 0 & X \end{pmatrix} \cdot \begin{pmatrix} x^N \\ y^N \\ s^N \end{pmatrix} = \begin{pmatrix} r^b \\ r^c \\ r^x \end{pmatrix}$$

für alle rechte Seiten $r^b \in \mathbb{R}^m$ sowie $r^c \in \mathbb{R}^n$ und $r^x \in \mathbb{R}^n$ bestimmt werden durch

$$\begin{aligned} D \cdot y^N &= r^b + A \cdot (S^{-1} \cdot X \cdot r^c - S^{-1} \cdot r^x), \\ s^N &= r^c - A^\top \cdot y^N, \\ x^N &= S^{-1} \cdot X \cdot s^N - S^{-1} \cdot r^x. \end{aligned}$$

Im Wesentlichen ist somit ein $n \times n$ Gleichungssystem zu lösen, wobei es sich bei D um eine stets symmetrische und positiv definite Matrix handelt.

Bemerkenswert dabei ist, dass aufgrund der Aussagen zuvor stets eine Cholesky-Zerlegung von D existiert, welche eine besonders effiziente Möglichkeit zur Lösung

des linearen Gleichungssystems bietet. Darüber hinaus sind in jedem Schritt des Prädiktor-Korrektor-Verfahrens zwei Gleichungssysteme mit identischer Matrix zu lösen, sodass die Cholesky-Zerlegung wiederverwendet werden kann. Erst dank dieser Ergebnisse können Inneren-Punkte-Verfahren effizient implementiert werden und bieten damit in der Praxis eine echte Alternative zum Simplex-Verfahren.

Zudem sei bemerkt, dass alle Multiplikationen mit X bzw. mit S^{-1} sehr einfach und mit einer Komplexität von nur $\mathcal{O}(n)$ durchgeführt werden können, da es sich dabei um einfache Diagonalmatrizen handelt.

Leider treten in der Praxis häufig numerische Rundungsfehler auf, welche die Inneren-Punkte-Verfahren durchaus instabil machen können: Viele Einträge der Vektoren x und s sind zwangsläufig sehr klein, sodass im Laufe der Verfahren sehr kleine oder sehr große Zahlen miteinander multipliziert und subtrahiert werden. Dies führt bei der Umsetzung der Verfahren oftmals zu Rundungsfehlern, sodass die Abbruchgenauigkeit ε nicht zu klein gewählt werden darf.

5 Anwendungsbeispiel Nullsummenspiele

In der Spieltheorie werden mathematische Modelle aufgestellt und analysiert, welche der Entscheidungsfindung dienen können. In diesem Abschnitt präsentieren wir als Spezialfall **Zwei-Personen Nullsummenspiele**. Derartige Modelle können als lineares Programm formuliert werden und zur Lösung bieten sich insbesondere Innere-Punkte-Verfahren an. Die Inhalte aus diesem Abschnitt sowie weiterführende Themen der Spieltheorie können beispielsweise in Owen (1982) oder Winter (2015) nachgelesen und vertieft werden.

Gegeben sei ein **Spiel** zwischen Spieler 1 und Spieler 2. Beide Spieler haben eine (möglicherweise unterschiedliche) Menge an Strategien, aus denen gleichzeitig jeweils eine Strategie gewählt wird. Dabei ist der Ausgang des Spiels festgelegt, sobald beide Spieler ihre Strategie gewählt haben. Weiterhin gilt bei Nullsummenspielen, dass der Gewinn von Spieler 1 genau dem Verlust von Spieler 2 entspricht und umgekehrt. Den Gewinn bzw. Verlust bezeichnen wir entsprechend als positive bzw. negative **Auszahlung**, jeweils aus Sicht von Spieler 1 gesehen:

Beispiel 5.1 *Das bekannte Spiel **Stein-Schere-Papier** ist ein sehr einfaches Zwei-Personen Nullsummenspiel. Beide Spieler wählen aus der gleichen Strategiemenge*

{Stein, Schere, Papier}

und als Auszahlung bedeutet eine 1, dass Spieler 1 gewinnt. Das Spiel kann damit vollständig beschrieben werden durch die **Spielmatrix**

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



Dabei entspricht die Strategie von Spieler 1 einer Zeile der Matrix und die Strategie von Spieler 2 einer Spalte der Matrix. Der Ausgang der Spiels entspricht den Auszahlungen gemäß der Spielmatrix: Wählt Spieler 1 das Papier (3. Zeile) und Spieler 2 die Schere (2. Spalte), so erhalten wir eine Auszahlung von -1 , d.h. Spieler 1 verliert und Spieler 2 gewinnt.



Aufgabe 5.1 Erweitere die Spielmatrix des Spiels Stein-Schere-Papier, falls beide Spieler zusätzlich noch den Brunnen als weitere Strategie erhalten. Dabei gelte: Stein und Schere fallen in den Brunnen (und verlieren), Papier deckt den Brunnen ab (und gewinnt).

Dass beide Spieler durchaus unterschiedliche Strategiemengen haben können, zeigt das folgende Beispiel:



Beispiel 5.2 Ein sehr stark vereinfachtes Roulette könnte folgendermaßen aussehen: Spieler 2 wählt eine Zahl aus der Menge

$$\{1, 2, 3, 4, 5, 6\}$$

und Spieler 1 wählt aus der Strategiemenge

$$\{\text{gerade, ungerade, kleiner 3, größer 4}\}.$$

Pro Spiel zahlt Spieler 1 einen Einsatz von 3. Falls die Zahl gemäß der Strategie von Spieler 2 die Eigenschaft gemäß der Strategie von Spieler 1 besitzt, so erhält Spieler 1 einen Gewinn von 7. Die Auszahlung setzt sich zusammen aus Gewinn abzüglich Einsatz, sodass wir

$$A = \begin{pmatrix} -3 & 4 & -3 & 4 & -3 & 4 \\ 4 & -3 & 4 & -3 & 4 & -3 \\ 4 & 4 & -3 & -3 & -3 & -3 \\ -3 & -3 & -3 & -3 & 4 & 4 \end{pmatrix}$$

als Spielmatrix erhalten.

Angenommen, $A = (a_{i,j}) \in \mathbb{R}^{m \times n}$ ist eine Spielmatrix, welche die Auszahlung eines Spiels definiert (aus Sicht von Spieler 1), und

$$x = e_r \in \mathbb{R}^m \quad \text{sowie} \quad y = e_s \in \mathbb{R}^n$$

seien zwei Einheitsvektoren. Dann ist

$$x^\top \cdot A \cdot y = e_r^\top \cdot A \cdot e_s = a_{r,s}$$

die Auszahlung, falls Spieler 1 die Strategie r und Spieler 2 die Strategie s wählt.

Zur Wahl der Strategie stellen wir folgende Grundannahme: Beide Spieler versuchen jeweils ihren schlimmsten Fall so gut wie möglich zu gestalten, d.h. Spieler 1 versucht die Funktion

$$v(x) = \min\{x^\top \cdot A \cdot y : y \in \mathbb{R}^n\}$$

zu maximieren, wobei für $x \in \mathbb{R}^m$ und $y \in \mathbb{R}^n$ zunächst nur Einheitsvektoren zulässig sind. Da die Spielmatrix jeweils die Auszahlung aus Sicht von Spieler 1 definiert, versucht Spieler 2 seine Strategie derart zu wählen, dass die Auszahlung an Spieler 1 minimiert wird. Dies bedeutet, dass Spieler 2 gewillt ist, die Funktion

$$w(y) = \max\{x^\top \cdot A \cdot y : x \in \mathbb{R}^m\}$$

zu minimieren. Angenommen, Spieler 1 und Spieler 2 können ihre Strategien derart wählen, sodass

$$\max\{v(x) : x \in \mathbb{R}^m\} = \min\{w(y) : y \in \mathbb{R}^n\} \quad (12)$$

gilt (wobei für x und y weiterhin nur Einheitsvektoren zulässig sind), dann werden Spieler 1 und Spieler 2 stets eine feste Strategie wählen und sich niemals von dieser abwenden. Wir erhalten einen stabilen Zustand des Spiels.

Ausblick Da für x und y bislang nur Einheitsvektoren zulässig sind, bieten sich auch folgende Formulierung an: Wir definieren

$$v^* = \max_{i=1,\dots,m} \min_{j=1,\dots,n} a_{i,j} \quad \text{und} \quad w^* = \min_{j=1,\dots,n} \max_{i=1,\dots,m} a_{i,j}.$$

Damit gilt stets $v^* \leq w^*$ und die Bedingung (12) ist äquivalent zu

$$v^* = w^*.$$

Die Schreibweise zuvor wird jedoch in der folgenden Herleitung hilfreich sein.

Aufgabe 5.2 Untersuche die Bedingung (12) anhand der Spielmatrix

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

sowie anhand der Spielmatrix des Stein-Schere-Papier Spiels.

Um auch dann einen stabilen Zustand zu erreichen, falls die Bedingung (12) nicht erfüllt ist, werden in der Spieltheorie **gemischte Strategien** zugelassen. Genauer sei

$$X = \left\{ x = (x_1, \dots, x_m) \in \mathbb{R}^m : x \geq 0 \text{ und } x_1 + \dots + x_m = 1 \right\}$$

die Menge aller gemischten Strategien von Spieler 1 mit folgender Bedeutung: Die gemischte Strategie $x = (x_1, \dots, x_m)$ bedeutet, dass sich Spieler 1 mit einer Wahrscheinlichkeit von x_i für die Strategie i entscheiden wird. Die Menge der gemischten Strategien von Spieler 2 sei analog

$$Y = \left\{ y = (y_1, \dots, y_n) \in \mathbb{R}^n : y \geq 0 \text{ und } y_1 + \dots + y_n = 1 \right\}$$

mit entsprechender Bedeutung. Zusammenfassend ist Spieler 1 auf der Suche nach einer gemischten Strategie $x \in X$, welche die Funktion

$$v(x) = \min\{x^\top \cdot A \cdot y : y \in Y\} \quad (13)$$

maximiert. Entsprechend ist Spieler 2 auf der Suche nach einer gemischten Strategie $y \in Y$, welche die Funktion

$$w(y) = \max\{x^\top \cdot A \cdot y : x \in X\} \quad (14)$$

minimiert.

Die Besonderheit dabei ist, dass die Optimierungsprobleme (13) und (14) auch als lineare Programme formuliert werden können. Ohne genaue Herleitungen geben wir hier nur das Ergebnis an: Gemäß der Grundannahme hat Spieler 1 eine gemischte Strategie x zu wählen, welche sich aus einer Optimallösung des folgenden linearen Programmes ergibt:

$$\begin{array}{ll} \max t, & \text{sodass} \\ & A^\top \cdot x - t \cdot e \geq 0, \\ & x_1 + \dots + x_m = 1, \\ & x_i \geq 0 \quad \text{für alle } i \in \{1, \dots, m\}, \\ & t \leq 0. \end{array}$$

Dabei gelte $t \in \mathbb{R}$ sowie $e = (1, \dots, 1) \in \mathbb{R}^n$. Analog erhalten wir für Spieler 2 das lineare Programm

$$\begin{array}{ll} \min t, & \text{sodass} \\ & A \cdot y - t \cdot e \leq 0, \\ & y_1 + \dots + y_n = 1, \\ & y_j \geq 0 \quad \text{für alle } j \in \{1, \dots, n\}, \\ & t \leq 0. \end{array}$$

Durch Substitution mit negativen Variablen kann eingesehen werden, dass die beiden linearen Programme dual zueinander sein. Wir erhalten daher aufgrund der starken Dualität das folgende Ergebnis:

Satz 5.1 (Min-Max-Satz) *In jedem Zwei-Personen Nullsummenspiel existieren gemischte Strategien, sodass der Gewinn von Spieler 1 gleich dem Verlust von Spieler 2 ist.*

Unter Verwendung von gemischten Strategien besitzt also jedes Zwei-Personen Nullsummenspiel einen stabilen Zustand, welchen wir als **Gleichgewicht** bezeichnen. Da wir sowohl an einer Optimallösung des Primales als auch an einer des Dualen interessiert sind, bieten sich insbesondere Innere-Punkte-Verfahren an.

Beispiel 5.3 Dank der Formulierung als lineares Programm sind wir nun in der Lage, ein Gleichgewicht unter Verwendung von gemischten Strategien im Steinschere-Papier Spiel aus Beispiel 5.1 zu berechnen.

Die Optimallösungen der linearen Programme liefern die gemischten Strategien

$$x = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \quad \text{und} \quad y = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right),$$

welche zu einem Gleichgewicht mit Auszahlung bzw. Zielfunktionswert

$$x^\top \cdot A \cdot y = 0$$

führen. Erwartungsgemäß erhalten wir für beide Spiele identische gemischte Strategien. Dies ist stets dann der Fall, wenn die Spielmatrix schiefssymmetrisch ist, d.h. falls $m = n$ und

$$A = -A^\top$$

gilt.

Aufgabe 5.3 Löse auch das Zwei-Personen Nullsummenspiel zur Spielmatrix aus Beispiel 5.2. Angenommen, das Spiel wird sehr oft durchgeführt, welcher Spieler wird dann auf lange Sicht seinen Gewinn steigern?

Zur weiteren numerischen Analyse wurden diverse quadratische Spielmatrizen mit zufälligen Einträgen aus dem Intervall $[-10, 10]$ generiert. Anschließend wurden die beiden linearen Programme zur Bestimmung eines Gleichgewichts des Spiels gelöst. Angewandt wurde dabei das Prädiktor-Korrektor-Verfahren wie im Abschnitt zuvor beschrieben, wobei als Abbruchgenauigkeit $\varepsilon = 10^{-6}$ verwendet wurde.

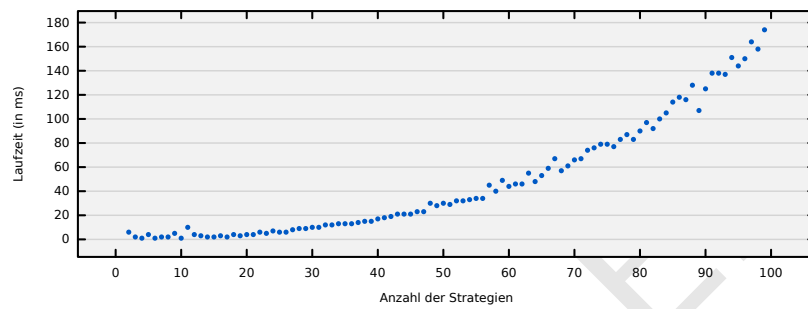
Abb. 1 zeigt die Laufzeiten sowie die Anzahl der benötigten Iterationen zur Lösung der Probleme, jeweils aufgetragen gegen die Anzahl der Strategien. Offenbar scheint die Anzahl der Iterationen mit weniger als 15 weitestgehend konstant und somit unabhängig von der Größe des Problems zu sein. Die numerisch ermittelte Laufzeit ist daher von der Größenordnung

$$\mathcal{O}(n^3),$$

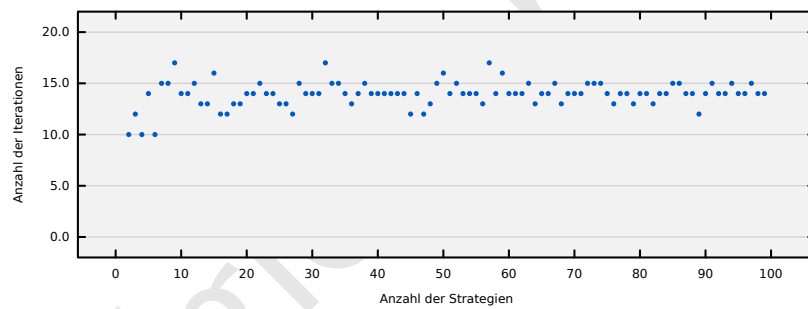
falls n die Anzahl der Strategien von Spieler 1 und Spieler 2 ist.

Schließlich vergleichen wir noch das Prädiktor-Korrektor-Verfahren am Beispiel von Zwei-Personen Nullsummenspielen zum Simplex-Verfahren. Dabei muss zunächst



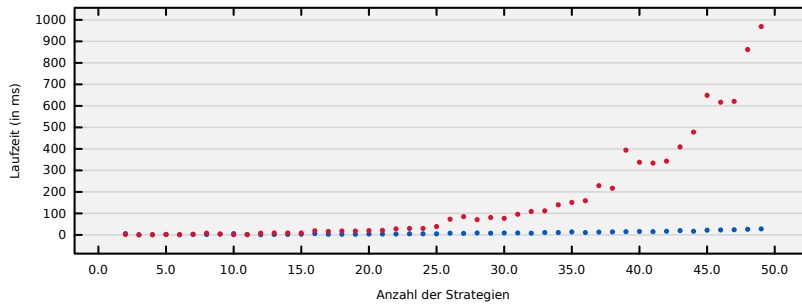


a Laufzeit in Abhängigkeit der Anzahl der Strategien

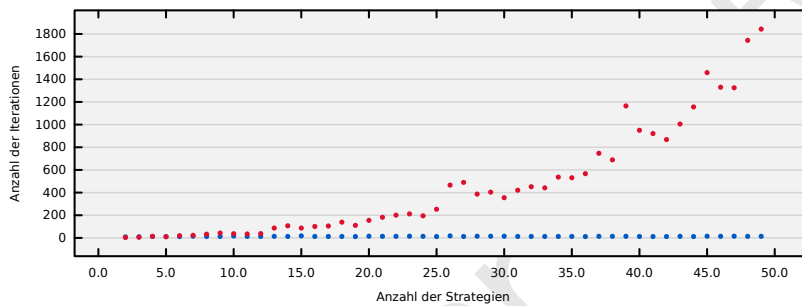


b Anzahl der Iterationen in Abhängigkeit der Anzahl der Strategien

Abb. 1 Numerische Analyse zum Prädiktor-Korrektor-Verfahren am Beispiel von Zwei-Personen Nullsummenspielen mit zufällig generierten Spielmatrizen. Aufgetragen sind die Laufzeiten sowie die Anzahl der benötigten Iterationen zur Lösung der Probleme, jeweils gegen die Anzahl der Strategien



a Laufzeit in Abhängigkeit der Anzahl der Strategien



b Anzahl der Iterationen in Abhängigkeit der Anzahl der Strategien

Abb. 2 Numerische Analyse zum Vergleich zwischen Prädiktor-Korrektor-Verfahren (blau) und Simplex-Verfahren (rot) am Beispiel von Zwei-Personen Nullsummenspielen mit zufällig generierten Spielmatrizen. Aufgetragen sind die Laufzeiten sowie die Anzahl der benötigten Iterationen zur Lösung der Probleme, jeweils gegen die Anzahl der Strategien

erwähnt werden, dass das Simplex-Verfahren insbesondere in Phase 1 je nach Pivotregel häufig nicht terminiert bzw. kreist. Daher wurde in der folgenden Analyse stets Blands Pivotregel verwendet, um die Endlichkeit des Simplex-Verfahrens zu gewährleisten. Alle sonstigen Eingabedaten wurden wie zuvor zufällig generiert.

Abb. 2 zeigt die numerischen Ergebnisse. Offenbar steigt die Anzahl der Iterationen unter Verwendung des Simplex-Verfahrens in Abhängigkeit der Größe des Problems sehr stark an, was sich entsprechend auch auf die Laufzeit zur Lösung des Problems auswirkt.

Zusammenfassend zeigt die Analyse, dass zur Lösung von Zwei-Personen Nullsummenspielen das Prädiktor-Korrektor-Verfahren dem Simplex-Verfahren ganz klar vorzuziehen ist. Trotzdem sei nochmals wiederholt, dass daraus keine generelle Schlussfolgerung gezogen werden darf: Es hängt stets vom jeweiligen Problem ab, welche Lösungsverfahren zu bevorzugen sind.

Literatur

- W. Domschke, A. Drexl, R. Klein, A. Scholl, 2015. *Einführung in Operations Research*. Springer, Berlin, 9. Auflage.
- H.W. Hamacher, K. Klamroth, 2006. *Lineare Optimierung und Netzwerkoptimierung*. Vieweg, Wiesbaden, 2. Auflage.
- F. Jarre, J. Stoer, 2004. *Optimierung*. Springer, Berlin, 1. Auflage.
- N. Karmarkar. 1984. A new polynomial-time algorithm for linear programming. *Combinatorica*, **4**: 373–395.
- S. Mehrotra. 1992. On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization*, **2**: 575–601.
- G. Owen, 1982. *Game Theory*. Academic Press Inc, San Diego, 2. Auflage.
- S. Winter, 2015. *Grundzüge der Spieltheorie: Ein Lehr- und Arbeitsbuch für das (Selbst-)Studium*. Springer Gabler, Heidelberg, 1. Auflage.
- S.J. Wright, 1997. *Primal-Dual Interior-Point Methods*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1. Auflage.