Iterative Lösung linearer Ungleichungssysteme

Ulrich Eckhardt

Von der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
genehmigte Dissertation
Referent: Prof. Dr. W. Krabs
Korreferent: Prof. Dr. F. Reutter
D 81 (Diss. T.H. Aachen)

Berichte der Kernforschungsanlage Jülich, Zentralinstitut für Angewandte Mathematik,
Jül–880–MA, August 1972.

Vorwort 2010

Die vorliegende Form des Berichts stammt aus dem Jahre 2010. Es sind gegenüber dem Text von 1972 einige wenige Dinge korrigiert und einige aktualisiert worden. Einige dieser Aktualisierungen sind als Fußnoten beigefügt. Diese Fußnoten betreffen neuere Literatur und Bezüge zu Algorithmen für Neuronale Netzwerke, insbesondere den *Greedy-Algorithmus* [20, Chapter 25]. Das hier vorgestellte Verfahren aus dem Jahre 1972 ist dieser Greedy-Algorithmus. Man findet hier insbesondere eine eingehende Konvergenzuntersuchung, deren Resultate über das in dem zitierten Buch Bewiesene hinausgehen.

Inhaltsverzeichnis

Kapitel 1

Einleitung

Viele Probleme der numerischen Mathematik lassen sich zurückführen auf die Aufgabe, eine Lösung $x_1, \ldots, x_d$ eines linearen Ungleichungssystems

$$\sum_{j=1}^{d} a_{ij} x_j \geq b_i \qquad i = 1, \ldots, n \tag{U}$$

zu finden. Insbesondere läßt sich eine lineare Optimierungsaufgabe (LO) mit Hilfe des Dualitätssatzes der linearen Optimierung (vgl. [23, §5.1] oder [109, Kap. II.3]) auf eine Aufgabe der Form (U) zurückführen. Umgekehrt kann man ein lineares Ungleichungssystem auch als lineare Optimierungsaufgabe betrachten [23, §3.4].

Ein Zweipersonen-Nullsummenspiel (Sp) läßt sich als lineare Optimierungsaufgabe – und damit als lineares Ungleichungssytem – formulieren. Im wesentlichen gilt auch die Umkehrung (zur Definition eines Zweipersonen-Nullsummenspiels und wegen einer präzisen Formulierung der Äquivalenz zu einem linearen Optimierungsproblem vgl. [109, Kap. IV.2]).

Ein lineares Gleichungssystem (LG) kann als lineares Ungleichungssystem geschrieben werden, wenn man von der Tatsache Gebrauch macht, daß eine Gleichung $x = y$ äquivalent ist zu den beiden Ungleichungen $x \geq y$ und $y \geq x$.

Ein besonders interessantes Problem ist das lineare Komplementärproblem (LC) (vgl. [70] oder [33]). Es läßt sich zeigen, daß man ein lineares Optimierungsproblem als ein spezielles lineares Komplementärproblem auffassen kann [70].

Drücken wir die Aussage „Die Aufgabe (A) ist auf die Aufgabe (B) zurückführbar“ durch (B) $\longrightarrow$ (A) aus, dann läßt sich das folgende Schema aufstellen:

$$\text{(LG)} \quad \longleftarrow \quad \text{(U)} \quad \longleftrightarrow \quad \text{(LO)} \quad \longleftrightarrow \quad \text{(Sp)} \quad \longleftarrow \quad \text{(LC)}$$

Zur Lösung der Aufgabe (LO) (und damit auch der Aufgaben (U) und (Sp)) verwendet man vorzugsweise Eliminationsverfahren, insbesondere die Simplexmethode [23, 106, 109]. Ähnliche Verfahren existieren auch zur Lösung von (LC) [70]. Derartige Methoden haben jedoch einige Nachteile. Insbesondere sind Eliminationsverfahren numerisch instabil und es ist oft sehr schwer, die Rundungsfehler unter Kontrolle zu bringen. Zudem wird durch die Pivotisierung eine etwa

vorhandene numerisch vorteilhafte spezielle Struktur des Problems zerstört. Weiterhin gibt es bei den Eliminationsverfahren zur Lösung linearer Ungleichungssysteme keine akzeptablen Abschätzungen für die Anzahl der zur Lösung erforderlichen Pivotschritte beziehungsweise die Güte der jeweiligen Zwischenlösungen. Klee und Minty [66] haben sogar gezeigt, daß schon bei sehr einfachen Problemen der Rechenaufwand bei der Simplexmethode außerordentlich groß werden kann. Schließlich ist es kaum möglich, schon bekannte gute Näherungslösungen numerisch zu nutzen.

Es ist bekannt, daß man sich bei der Lösung der Aufgabe (LG) mit Erfolg iterativer Verfahren bedient. Es liegt also nahe, sich zu fragen, inwieweit iterative Verfahren auch für die Probleme (U), (LO), (Sp) beziehungsweise (LC) verwendet werden können. Deshalb soll jetzt ein Überblick über bereits vorhandene derartige Methoden und deren Anwendung gegeben werden. Besonders interessant sind dabei Methoden, die die Lösung eines linearen Ungleichungssystems mit unendlich vielen Ungleichungen ermöglichen. Bei solchen Problemen sind natürlich iterative Verfahren sehr vorteilhaft, weil dann Eliminationsverfahren sicher nicht mehr anwendbar sind.

In der Zeit um 1950 standen – unter dem Einfluß John von Neumanns – die Zweipersonen-Nullsummenspiele im Vordergrund des Interesses. Brown und von Neumann [13] gaben ein System linearer Differentialgleichungen an, dessen Lösung für große Argumentwerte t die Lösung eines Spiels approximiert. Die Güte der Approximation ist dabei von der Ordnung $O(1/t)$. Davon ausgehend entwickelte Brown [14] die Methode des „fictitious play“, eine iterative Methode zur Bestimmung des Spielwerts. Julia Robinson [95] bewies die Konvergenz des Verfahrens (vgl. auch [109, Kap. IV.3]). Shapiro [99] gab als Konvergenzordnung $O\left(1/\sqrt[k]{r}\right)$ and (k ganzzahlig, positiv, fest, r Iterationsindex). Unlängst ermittelte Čaĭkovskiĭ [16] für zwei verschiedene Versionen der Brownschen Methode Konvergenzordnungen von $O(1/r)$ beziehungsweise $O(\alpha^r)$ $(0 < \alpha < 1)$.

Andere Verfahren stammen von v. Neumann [84], Braithwaite [8] und Danskin [29]. Die Methode von Braithwaite liefert für eine gewissen Klasse von Spielen bereits nach endlich vielen Iterationsschritten eine Lösung und die von Danskin ist auch auf stetige Spiele anwendbar, was dem Fall unendlich vieler linearer Ungleichungen entspricht.

Nachteilig bei allen diesen Methoden ist die schlechte Konvergenz. Als ein Vorteil ist allerdings anzusehen, daß solche Verfahren leicht auf komplizierte Spiele verallgemeinert werden können (vgl. auch Shapley [100]). So lösten Matrose und Mukundan [74] ein sogenanntes rekursives Spiel mit der Methode von Brown.

Eine große Klasse von Verfahren zur iterativen Lösung linearer Ungleichungssysteme geht auf eine einfache Idee von Fejér aus dem Jahre 1922 zurück [46]: Die Menge $P \subseteq \mathbb{R}^d$ aller Lösungen von (U) ist eine konvexe Menge. Es sei $x_0 \in \mathbb{R}^d$ keine Lösung. Dann gibt es (vgl. etwa [23, Anhang]) eine Hyperebene, die P von x_0 trennt. Sei y_0 die Projektion von x_0 auf diese Hyperebene. Sämtliche Punkte $y = x_0 + t \cdot (y_0 - x_0)$, $0 < t < 2$, haben zu jedem Punkt aus P einen geringeren Abstand als x_0 selbst (Abbildung 1.1).

Basierend auf dieser Idee formulierten Agmon [2] und Motzkin und Schoenberg [83] ein Verfahren, das eine Folge von Punkten liefert, die linear gegen eine Lösung von (U) konvergiert, falls eine solche Lösung existiert. Lineare Konvergenz einer Folge ist dabei folgendermaßen definiert:

Definition 1.1 *Die Folge $\{a_r\}_{r=1}^{\infty}$ eines normierten Vektorraumes mit Norm $\|\cdot\|$ heißt* linear konvergent *gegen ein Element a, wenn es eine Zahl α gibt mit $0 < \alpha < 1$, so daß*

$$\frac{\|a_{r+1} - a_r\|}{\|a_r - a\|} \leq \alpha \qquad \textit{für alle } r.$$

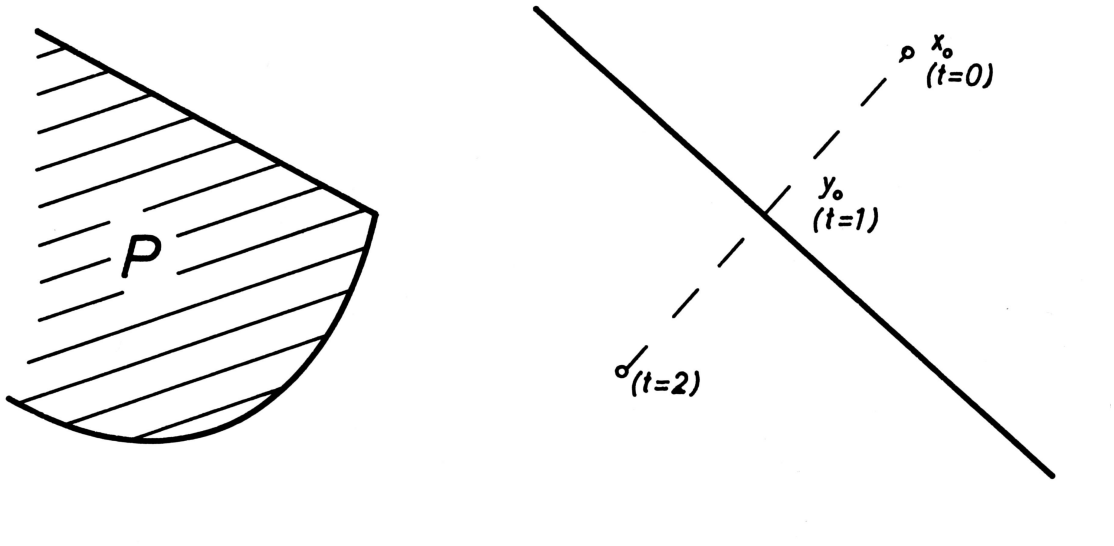

Abbildung 1.1: Die Idee von Fejér.

Ist die Aufgabe (U) jedoch nicht lösbar, dann kann man wenig über das Verhalten des Agmonschen Verfahrens aussagen (vgl. hierzu auch Abschnitt 11). Das ist ein schwerwiegender Nachteil der Methode. Denselben Nachteil besitzen auch andere solche Methoden, die man als *Fejérsche Methoden* oder *Projektionsmethoden* bezeichnet. Andererseits lassen sie sich ohne große Schwierigkeiten verallgemeinern auf endlich oder unendlich viele Ungleichungen in Hilberträumen, wobei dann allerdings nicht mehr lineare Konvergenz gezeigt werden kann. Man kennt inzwischen eine ganze Reihe von Varianten der Fejérschen Idee [39, 40, 41, 59, 60, 76, 79, 87, 89, 93][1]. Man kann mit solchen Methoden auch die etwas allgemeinere Aufgabe lösen, ein Element des Durchschnitts konvexer Mengen zu finden [10, 12, 51]. Naheliegend ist die Anwendung auf (lineare oder konvexe) Optimierungsaufgaben [12, 15, 41, 42, 43]. Vaisbord [107] beschreibt die Lösung eines Kontrollproblems mit einer Projektionsmethode. Verfahren derselben Bauart verwendet man auch zur Minimierung konvexer Funktionale [90, 98, 72] oder zur genäherten Lösung linearer Gleichungssysteme [48], [44, Kap. III]. In diesem Kontext bezeichnet man sie auch als *Relaxationsverfahren.*

Bei den Verfahren vom Fejérschen Typ zur Lösung linearer Ungleichungssysteme erhält man im Falle endlich vieler Ungleichungen lineare Konvergenz. Ein gewisser Nachteil ist, daß man die Konstante α in Definition 1.1 nicht einfach charakterisieren kann [61]. Für unendlich viele lineare Ungleichungen hat man keine Aussagen über die Konvergenzgüte.

Ein besonders interessantes Anwendungsgebiet linearer Ungleichungssysteme ist die Zeichenerkennung (pattern recognition) [5, 85, 97]. Eine fundamentale Aufgabe dieser Theorie läßt sich etwa wie folgt formulieren: Gegeben seien zwei (endliche) Punktmengen A und B im $\mathbb{R}^d$. Gesucht

[1] In einem 1974 erschienenen Buch von Belenkij et al. [6] wurden Iterationsverfahren zur Lösung von Aufgaben der Spieltheorie und der Optimierung sowohl theoretisch als auch algorithmisch dargestellt.

ist ein Vektor $a \in \mathbb{R}^d$, so daß

$$\begin{aligned} \langle a, x\rangle &> 0 \qquad \text{für } x \in A, \\ \langle a, x\rangle &< 0 \qquad \text{für } x \in B \end{aligned}$$

($\langle \cdot, \cdot \rangle$ ist dabei das Skalarprodukt im $\mathbb{R}^d$). Das heißt, a trennt A und B. Als Beispiel betrachten wir die Wettervorhersage. Es sei $x_t \in \mathbb{R}^d$ ein Vektor von d Meßwerten am Tage t, etwa Luftdruck, Temperatur, Luftfeuchtigkeit usw. Wir definieren $x_t \in A$, wenn es am Tage $t+1$ regnet und $x_t \in B$, wenn es am Tage $t+1$ nicht regnet. Haben wir die Meßreihe für einen längeren Zeitraum durchgeführt und gibt es einen solchen Vektor a, dann können wir mit ziemlicher Sicherheit an einem Tage t das Wetter am Tage $t+1$ prophezeien, indem wir prüfen, ob $\langle a, x_t\rangle$ positiv oder negativ ausfällt.[2]

Bei dieser Aufgabenstellung haben sich iterative Verfahren gut bewährt. Von Vorteil ist hier, daß man Fehlklassifikationen (im Beispiel: falsche Wettervorhersagen) dazu benutzen kann, die bereits berechnete Hyperebene durch einige Iterationsschritte zu verbessern. Man hat also ein „lernendes System“. Sind die beiden Mengen nicht trennbar, dann führen Verfahren vom Fejérschen Typ zu Schwierigkeiten (vgl. jedoch [59]). Das ist der Hauptgrund dafür, daß man bei der praktischen Berechnung der Hyperebenen häufig auf die aufwendigere Simplexmethode zurückgreift. Trotzdem wurden für diesen Problemkreis eine Fülle von iterativen Verfahren vorgeschlagen [3, 7, 9, 32, 49, 50, 58, 59, 64, 76, 86, 110]. Ein verwandtes Anwendungsgebiet ist die Darstellung von logischen Funktionen durch Schwellwertfunktionen [37] sowie die Taxonomie in den Biowissenschaften [102].

Viele spezielle Aufgabenstellungen, die auf Optimierungsprobleme führen, sind durch iterative Verfahren gelöst worden. Hier sind insbesondere Transportprobleme zu nennen, die sich von dem klassischen Transportproblem [23, §4.8] durch zusätzliche Komplikationen (Nichtlinearitäten usw.) unterscheiden [11, 31, 91].

Zur Lösung des diskreten linearen Tschebyscheff-Approximationsproblems gab Krabs [68] ein Iterationsverfahren an, das sich in der Praxis gut bewährt hat. Das diskrete lineare Tschebyscheff-Approximationsproblem kann als lineares Optimierungsproblem formuliert werden [23, §16]. Ein lineares Optimierungsproblem kann man lösen, indem man gleichzeitig ein Eliminationsverfahren und ein Iterationsverfahren anwendet, wobei sich beide Verfahren gegenseitig steuern [35].

Habetler und Price [53] kündigten ein Iterationsverfahren zur Lösung des linearen Komplementärproblems an. Torsti und Aurela [105] führten ein Problem aus der Physik auf ein Komplementärproblem zurück und lösten dies iterativ. Eine besonders interessante Aufgabenstellung wird von Fridman und Chernina [47] mitgeteilt. Es handelt sich darum, die Kräfteverteilung an zwei elastischen Körpern, die sich in Kontakt befinden, zu berechnen[3]. Dieses physikalische Modell liefert eine sehr schöne anschauliche Interpretation des linearen Komplementärproblems. Auch dieses Problem wurde iterativ behandelt. Ein besonders überzeugendes Beispiel für die iterative Lösung eines Komplementärproblems stammt von Cryer [26, 27, 28]. Es handelt sich um ein Problem, das bei der Diskretisierung einer freien Randwertaufgabe für Zapfenlager entsteht. Die Koeffizientenmatrix ist dabei sehr groß (etwa 10 000 Zeilen und Spalten), so daß die Anwendung eines Eliminationsverfahrens nicht empfehlenswert ist. Außerdem hat die Matrix als Diskretisierungsmatrix eine besondere Struktur, die bei der Pivotisierung verloren gehen würde.

[2] Man bezeichnet die Realisierung einer solchen Maschine, die imstande ist, zu „lernen“ auch als ein *Perzeptron*. Derartige Maschinen wurden als Modelle für neuronale Vorgänge angesehen [7, 86, 97]. In neuerer Zeit erlebten sie als sogenannte *neuronale Netzwerke* [85] oder *Support Vector Machines* [24] eine Renaissance.

[3] Dieses sogenannte *Kontaktproblem* geht auf Heinrich Hertz zurück [57].

Cryer löste dieses Problem durch eine spezielle Variante des *successive overrelaxation*-Verfahrens. Die mitgeteilten numerischen Ergebnisse sind durchaus überzeugend,

Leider liegen bislang nur wenige numerische Vergleiche iterativer Methoden mit der Simplexmethode vor. Hershkovitz [56] verglich die Methode von Brown [14] mit der von v. Neumann [84]. Es stellte sich durchweg eine schlechtere Konvergenz des v. Nemannschen Verfahrens heraus. Hoffman et al. [61] verglichen die Methoden von Brown [14] und von Agmon [2] mit der Simplexmethode. Die gewonnenen Ergebnisse kann man heute nur mit großer Vorsicht verwerten, denn sowohl die behandelten Probleme als auch die verwendete Rechenanlage kann man jetzt kaum noch als typisch bezeichnen. Fedorenko [45] gibt Vergleichszsahlen für die Simplexmethode und eine iterative Methode an. Mays [75] untersuchte verschiedene iterative Verfahren und Smith [101] verglich ein iteratives Verfahren mit der Simplexmethode. Tanabe [104] gibt numerische Erfahrungen mit einer Projektionsmethode zur Lösung linearer Gleichungssysteme an. Generell kann man sagen, daß sich iterative Verfahren mit Vorteil verwenden lassen, wenn die Koeffizientenmatrix des Ungleichungssystems besonders groß ist, wenn man den Einfluß von Rundungsfehlern in Grenzen halten will oder wenn die Koeffizientenmatrix eine numerisch besonders günstige Struktur hat, die durch ein Eliminationsverfahren zerstört werden würde.

In der vorliegenden Arbeit soll ein Verfahren vorgestellt werden, das die Vorzüge der genannten Verfahren in sich vereinigt, jedoch viele ihrer Nachteile nicht besitzt. So liefert es in jedem Falle eine sinnvolle Aussage, auch wenn die Aufgabe (U) keine Lösung besitzt. Weiterhin läßt es sich im Falle unendlich vieler linearer Ungleichungen in einem Hilbertraum formulieren. Die Konvergenz kann in großer Allgemeinheit abgeschätzt werden, in vielen Fällen läßt es sich sogar zeigen, daß die Lösung bereits nach endlich vielen Schritten gefunden wird. Die Konvergenz ist bestenfalls linear, im ungünstigsten Falle von der Ordnung $O\left(\frac{1}{\sqrt{r}}\right)$.

Die Methode hat eine gewisse Verwandtschaft mit einem kürzlich von Mitchell, Dem'janov und Malozemov [81, 82] für eine spezielle Aufgabenstellung vorgeschlagenen Verfahrens und beruht nicht auf der Fejérschen Idee.

Kapitel 2

Konvexe Mengen

In diesem Abschnitt sollen die wichtigsten Eigenschaften konvexer Mengen für den späteren Gebrauch zusammengestellt werden.

Es sei $\mathcal{E}$ ein reeller linearer Raum mit Nullelement $\Theta_{\mathcal{E}}$.

Definition 2.1 *Die Menge $K \subseteq \mathcal{E}$ heißt* konvex*, wenn gilt:*

$$x \in K, y \in K, 0 \leq p \leq 1 \implies p \cdot x + (1-p) \cdot y \in K.$$

Eine gewisse Bedeutung haben spezielle konvexe Mengen:

Definition 2.2 *Die Menge $C \subseteq \mathcal{E}$ heißt* konvexer Kegel*, (mit Scheitel $\Theta_{\mathcal{E}}$), wenn gilt:*

$$x \in C, y \in C, p \geq 0 \implies p \cdot x \in C \text{ und } x + y \in C.$$

Offenbar ist ein konvexer Kegel auch eine konvexe Menge.

Es sei jetzt M irgendeine Teilmenge von $\mathcal{E}$.

Definition 2.3 *Die* konvexe Hülle *von M ist definiert als*

$$\text{conv } M = \left\{ x \in \mathcal{E} \;\middle|\; x = \sum_{j=1}^{k} p_j a_j, p_j \geq 0, \sum_{j=1}^{k} p_j = 1, a_j \in M \right\},$$

wobei k eine endliche Zahl ist.

Man nennt M auch ein Erzeugendensystem *für K, wenn $K = \text{conv } M$.*

Es ist klar, daß conv M eine konvexe Menge ist.

Definition 2.4 *Der von M erzeugte konvexe Kegel ist definiert als*

$$\text{cone } M = \left\{ x \in \mathcal{E} \;\middle|\; x = \sum_{j=1}^{k} p_j a_j, p_j \geq 0, a_j \in M \right\},$$

Auch hier sei die Summe als endliche Summe verstanden.

cone M ist natürlich ein konvexer Kegel im Sinne der Definition 2.2.

Wichtig ist der folgende Satz, der besagt, daß man eine gewisse Trennungseigenschaft nur auf einem Erzeugendensystem nachzuweisen braucht, damit sie für die ganze konvexe Menge gilt:

Satz 2.1 *Es sei $\mathcal{E}$ ein unitärer Raum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, $K \subseteq \mathcal{E}$ eine konvexe Meenge, M ein Erzeugendensystem von K, $b \in \mathcal{E}$ fest. Es sei weiterhin $\widehat{x} \in \mathcal{E}$ ein festes Element und γ eine Zahl.*

Die Ungleichung

$$\langle \widehat{x}, y - b \rangle \geq \gamma \tag{2.1}$$

gilt genau dann für alle $y \in \operatorname{cl} K$, wenn sie für alle $y \in M$ gilt.
($\operatorname{cl} K$ *= abgeschlossene Hülle von K).*

Beweis 1. Die Ungleichung gelte für alle $y \in \operatorname{cl} K$. Wegen $M \subseteq K \subseteq \operatorname{cl} K$ gilt sie dann auch für alle $y \in M$.

2. Die Ungleichung (2.1) gelte für alle Elemente aus M. Sei $y \in K$, also $y = \sum p_j \cdot y_j$, $p_j \geq 0$, $\sum p_j = 1$, $y_j \in M$. Dann ist

$$\langle \widehat{x}, y - b \rangle = \left\langle \widehat{x}, \sum p_j \cdot y_j - b \right\rangle = \sum p_j \cdot \langle \widehat{x}, y_j - b \rangle \geq \sum p_j \cdot \gamma = \gamma,$$

also gilt (2.1) für alle $y \in K$, damit natürlich auch für alle $y \in \operatorname{cl} K$. □

Es sei jetzt $\mathcal{E}$ mit einer Topologie versehen, M irgendeine Teilmenge von $\mathcal{E}$. Wir bezeichnen mit

$\operatorname{cl} M$	die abgeschlossene Hülle von M,
$\operatorname{int} M$	das (topologische) Innere von M,
$\operatorname{bd} M$	den Rand von M.

Für eine konvexe Menge $K \subseteq \mathcal{E}$ sei $L \subseteq E$ die kleinste lineare Mannigfaltigkeit[1], die K enthält. Dann nennen wir

$\operatorname{relint} K$	das (relative) Innere von K bezüglich der in L induzierten Relativtopologie,
$\operatorname{relbd} K$	$= K - \operatorname{relint} K$ den relativen Rand von K.

Wir werden später noch zwei wichtige Aussagen benötigen, die im $\mathbb{R}^d$, dem d-dimensionalen euklidischen Vektorraum gelten. Die Beweise sind etwas lang, so daß auf die einschlägige Literatur verwiesen wird:

Satz 2.2 (Satz von Carathéodory) [2] *Sei $M \subseteq \mathbb{R}^d$, $x \in \operatorname{conv} M$.*
Dann gibt es eine Darstellung von x in der Form

$$x = \sum_{j=1}^{d+1} p_j y_j, p_j \geq 0, \sum p_j = 1, y_j \in M,$$

das heißt, x ist als Konvexkombination von höchstens $d+1$ Elementen aus M darstellbar.

[1] Eine *lineare Mannigfaltigkeit* ist eine Menge der Form $x_0 + V$, wo x_0 ein festes Element von $\mathcal{E}$ ist und V ein linearer Teilraum von $\mathcal{E}$.

[2] Constantin Carathéodory (1873 –1930), [17].

Beweis Siehe [103, Chap. 2.2]. □

Wir definieren:

Definition 2.5 *Die Vektoren $a_1, \ldots, a_k \in \mathbb{R}^d$ heißen* affin abhängig*, wenn es Zahlen $p_1, \ldots, p_k$ gibt, so daß $\sum p_j = 1$ und $\sum_{j=1}^{k} p_j a_j = \Theta_d$.*
Die konvexe Hülle von $d+1$ affin unabhängigen Vektoren des $\mathbb{R}^d$ bezeichnet man als (nichtentartetes) Simplex.

Es gilt der

Satz 2.3 *Es sei $A \subseteq \mathbb{R}^d$ eine endliche Menge, die $d+1$ affin unabhängige Vektoren enthalte. Dann gilt:*
$y \in$ int conv A genau dann, wenn x als Konvexkombination aller Elemente aus A mit positiven Koeffizienten darstellbar ist.

Beweis Siehe [96, Part I, Sec. 6, Theorem 6.9]. □

Korollar 2.1 *Ein nichtentartetes Simplex hat innere Punkte.*

Kapitel 3

Ein Verfahren im Hilbertaum

Es sei $\mathcal{H}$ ein Hilbertraum, $K \subseteq \mathcal{H}$ eine konvexe Menge, M ein Erzeugendensystem von K und $b \in \mathcal{H}$ ein festes Element.

Zentrale Bedeutung in der Theorie konvexer Mengen hat der sogenannte (strikte) Trennungssatz. Wir wollen ihn für die vorliegende Situation formulieren, für weitergehende Einzelheiten sei auf die Arbeit von Hurwicz [63, Theorem II.2] oder auf das Buch von Valentine [108, Kapitel II] verwiesen:

Satz 3.1 (Trennungssatz) *Ist $b \notin \operatorname{cl} K$, dann gibt es ein Element $\widehat{x} \in \mathcal{H}$, so daß*

$$\delta := \inf_{x \in \operatorname{cl} K} \langle \widehat{x}, x - b \rangle > 0, \tag{3.1}$$

wobei $\langle \cdot, \cdot \rangle$ das Skalarprodukt in $\mathcal{H}$ bezeichnet.

Wegen Satz 2.1 ist die Aussage (3.1) äquivalent zu

$$\langle \widehat{x}, x - b \rangle \geq \delta > 0 \qquad \text{für alle } x \in M.$$

Geometrisch bedeutet Satz 3.1, daß im Falle $b \notin \operatorname{cl} K$ die Hyperebene $\{x \in \mathcal{H} \mid \langle \widehat{x}, x \rangle \geq \delta + \langle \widehat{x}, b \rangle\}$ das Element b strikt von $\operatorname{cl} K$ trennt:

$$\langle \widehat{x}, x \rangle \geq \delta + \langle \widehat{x}, b \rangle > \langle \widehat{x}, b \rangle \qquad \text{für alle } x \in \operatorname{cl} K.$$

In diesem Abschnitt soll ein Verfahren vorgestellt werden, das es erlaubt, konstruktiv zu entscheiden, ob $b \in \operatorname{cl} \operatorname{conv} K$ beziehungsweise ob es eine solche trennende Hyperebene gibt. Dazu muß allerdings vorausgesetzt werden, daß die erzeugende Menge M beschränkt ist. Am Ende des Paragraphen wird dann untersucht werden, inwieweit man sich von dieser Voraussetzung lösen kann.

Wir betrachten demgemäß die folgende Aufgabenstellung:

$$\text{Gesucht ist eine Folge } \{x_r\} \text{ mit } x_r \in K \text{ für alle } r \text{ und } \lim_{r \to \infty} x_r = b. \tag{A}$$

Die folgende Aufgabe nennen wir die zu (A) *duale Aufgabe*:

$$\text{Gesucht ist ein } \widehat{x} \in \mathcal{H} \text{ so daß } \langle \widehat{x}, y - b \rangle > 0 \quad \text{für alle } y \in M. \tag{D}$$

Ist (D) lösbar, braucht (3.1) nicht lösbar zu sein. Wir kommen auf die Lösbarkeit von (3.1) noch zurück (starke Lösbarkeit von (D), Definition 3.1).

Zur Abkürzung setzen wir für $x, y \in \mathcal{H}$:

$$\Phi(x, y) = \langle x - b, y - b \rangle$$

und

$$\phi(x) = \|x - b\|^2.$$

Weiterhin sei vorausgesetzt:

$$\text{Es gebe eine Zahl } m, \text{ so daß gilt } \phi(x) \leq M \text{ für alle } x \in M. \tag{V}$$

Für jedes $x \in \mathcal{H}$ sei eine noch zu spezifizierende Auswahlmenge $A(x)$ gegeben. Damit definieren wir das Verfahren:

0. Sei $x_1 \in K$, $r := 1$.

1.1. $A(x_r) = \emptyset$. Wir brechen ab. (VF)

1.2. Andernfalls sei $y_r \in A(x_r)$ gewählt.

2. Setze

$$\mu_r = \frac{\phi(x_r) - \Phi(x_r, y_r)}{\phi(x_r) + \phi(y_r) - 2 \cdot \Phi(x_r, y_r)} \tag{3.2}$$

und

$$x_{r+1} = \mu_r \cdot x_r + (1 - \mu_r) \cdot y_r. \tag{3.3}$$

3. $r := r + 1$, gehe zu 1.

Der zweite Schritt von (VF) hat einen einfachen geometrischen Hintergrund (Abbildung 3.1). (3.3) besagt, daß x_{r+1} auf der Geraden durch x_r und y_r liegt. Dabei fordert man zweckmäßigerweise, daß $x_r \neq y_r$ ist. Die einfachste Möglichkeit, dies sicherzustellen, ist die Forderung

$$x \notin A(x) \qquad \text{für alle } x \in \mathcal{H}. \tag{3.4}$$

Wählt man μ_r nach (3.2), dann ist x_{r+1} gerade derjenige Punkt auf der Geraden durch x_r und y_r, für den $\phi(x)$ minimal ist. Stellt man noch sicher, daß $0 \leq \mu_r \leq 1$ ist, dann liegt x_{r+1} auf der Verbindungsstrecke von x_r und y_r. Ist $x_r \in K$, dann impliziert dies wegen $y_r \in M$, daß auch $x_{r+1} \in K$ ist. Wegen (3.2) ist $0 \leq \mu_r \leq 1$ gleichwertig zu

$$\phi(x_r) \geq \Phi(x_r, y_r) \quad \text{und} \quad \phi(y_r) \geq \Phi(x_r, y_r).$$

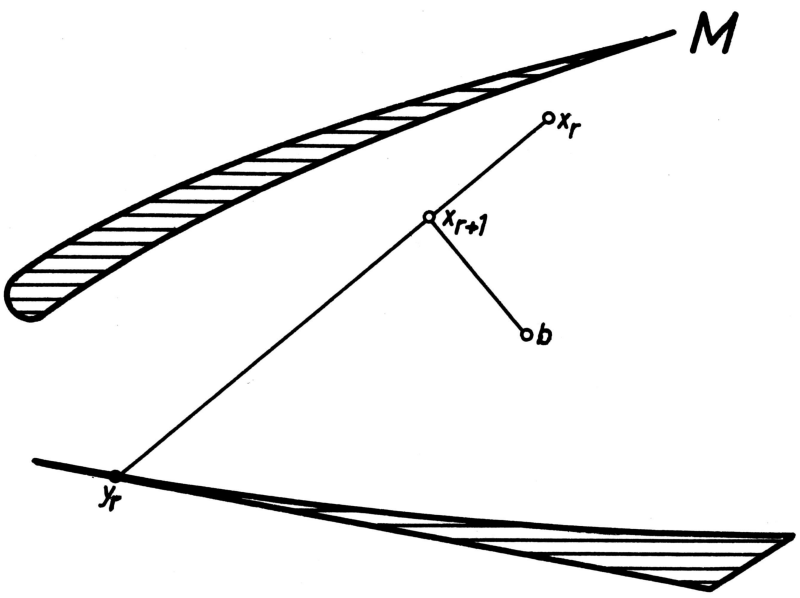

Abbildung 3.1: Illustration des Verfahrens (VF).

Die Fälle $\mu_r = 0$ (das heißt $x_{r+1} = y_r$ und $\phi(y_r) = \Phi(x_r, y_r)$) und $\mu_r = 1$ ($x_{r+1} = x_r$ und $\phi(x_r) = \Phi(x_r, y_r)$) sind uninteressant. Es ist daher sinnvoll, zu fordern

$$\begin{array}{l}\phi(y) > \Phi(x,y) \\ \phi(x) > \Phi(x,y)\end{array} \qquad \text{für alle } x \in \mathcal{H} \text{ und alle } y \in A(x). \tag{3.5}$$

Zunächst betrachten wir die Auswahlmenge

$$A_0(x) = \{y \in M \mid \Phi(x,y) \leq 0\}. \tag{3.6}$$

Hier sind (3.4) und (3.5) erfüllt, falls $x \neq b$ und $y \neq b$. Es gilt der Konvergenzsatz:

Satz 3.2 *Es gelte (V) und es werde jeweils $y_r \in A_0(x_r)$ ausgewählt. Dann können folgende Fälle eintreten:*

1. *Für ein r ist $x_r = b$ oder $y_r = b$, also (A) (trivial) gelöst.*
2. *Für ein r ist $A_0(x_r) = \emptyset$, also (D) gelöst.*
3. *Für alle r ist $A_0(x_r) \neq \emptyset$, $x_r \neq b$, $y_r \neq b$. Dann gilt*

$$\phi(x_r) = \|x_r - b\|^2 \leq \frac{m}{r}, \tag{3.7}$$

also $\lim x_r = b$.

Beweis Zunächst folgt aus $y_r \in A_0(x_r)$ und $x_r \neq b$, $y_r \neq b$, daß $0 < \mu_r < 1$ ist, also $x_{r+1} \in K$, falls $x_r \in K$.

Zu zeigen ist noch, daß (3.7) erfüllt ist, wenn die Fälle 1. und 2. nicht eintreten. Aus (3.2) und (3.3) errechnet man

$$\phi\left(x_{r+1}\right) = \frac{\phi(x_r) \cdot \phi(y_r) - \Phi(x_r, y_r)^2}{\phi(x_r) + \phi(y_r) - 2 \cdot \Phi(x_r, y_r)} \leq \frac{\phi(x_r) \cdot \phi(y_r)}{\phi(x_r) + \phi(y_r)}. \tag{3.8}$$

Weil die Funktion $F(x) = \frac{x}{a+x}$ für $y > 0$ monoton wächst, kann man $\phi(x_r)$ in Zähler und Nenner durch die in (V) definierte Zahl m nach oben abschätzen:

$$\phi(x_{r+1}) \leq \phi(x_r) \cdot \frac{m}{\phi(x_r) + m}.$$

Durch Induktion beweist man daraus (3.7). □

Es sei darauf hingewiesen, daß Satz 3.2 ein spezieller Trennungssatz ist. Man kann tatsächlich Satz 3.1 aus Satz 3.2 herleiten. Hier soll jedoch das Verfahren (VF) zur Lösung des Aufgabenpaares (A), (D) behandelt werden, ohne weiter auf diesen Punkt einzugehen.

Wir nehmen jetzt an, daß (A) nicht lösbar ist, das heißt, $b \notin \text{cl } K$. Dann gibt es eine Zahl $\gamma > 0$, so daß $\phi(x) \geq \gamma$ für alle $x \in K$. Andernfalls nämlich gäbe es eine Folge $\{z_r\}$ mit $z_r \in K$ für alle r und $\phi(z_r) \longrightarrow 0$. Das hieße aber $\lim z_r = b$, im Widerspruch zu $b \notin \text{cl } K$.

Für ein festes δ mit $0 < \delta < \frac{1}{2}$ definieren wir jetzt die Auswahlmenge

$$A_\delta(x) = \{y \in M \mid \Phi(x, y) \leq \delta \cdot \phi(x) \text{ und } \Phi(x, y) < \phi(y)\}. \tag{3.9}$$

Auch hier sind die Bedingungen (3.4) und (3.5) erfüllt.

Es gilt der

Satz 3.3 *Es gelte (V), (A) sei nicht lösbar. Sei γ wie oben gewählt und $y_r \in A_\delta(x_r)$, $0 < \delta < \frac{1}{2}$. Das Verfahren (VF) bricht dann nach spätestens*

$$\frac{\log \gamma - \log \phi(x_1)}{\log m - \log(m + \gamma \cdot (1 - 2 \cdot \delta))} \tag{3.10}$$

Schritten mit einer Lösung $\widehat{x} = x_r - b$ von (D) ab, so daß gilt

$$\Phi(x_r, y) > \delta \cdot \gamma > 0 \quad \textit{für alle } y \in M. \tag{3.11}$$

Beweis Aus $y_r \in A_\delta(x_r)$ folgt wieder $0 < \mu_r < 1$, also $x_r \in K$ für alle r bis zum Abbruch.

Mit (3.8) wird

$$\frac{\phi(x_{r+1})}{\phi(x_r)} \leq \frac{\phi(y_r)}{\phi(y_r) + \phi(x_r) \cdot (1 - 2 \cdot \delta)} \leq \frac{m}{m + \gamma \cdot (1 - 2 \cdot \delta)} < 1, \tag{3.12}$$

es liegt also bis zum Abbruch lineare Konvergenz vor.

Aus (3.12) folgt (3.10) wegen $\phi(x_r) \geq \gamma$ für alle r, und (3.11) ergibt sich aus der Abbruchbedingung $A_\delta(x_r) = \emptyset$. □

Die Bedeutung des Satzes 3.3 liegt darin, daß man – falls (A) nicht lösbar ist – stets eine Lösung $\widehat{x}$ von (D) ermitteln kann, so daß

$$\langle \widehat{x}, y - b \rangle \geq \delta \cdot \gamma > 0 \quad \text{für alle } y \in M.$$

Dabei hat $\widehat{x}$ eine Darstellung $\widehat{x} = x - b$, $x \in K$.

Wir definieren:

Definition 3.1 *$\widehat{x}$ heißt* starke Lösung *von (D), wenn es eine Zahl $\delta > 0$ gibt, so daß*

$$\langle \widehat{x}, y - b \rangle \geq \delta \quad \textit{für alle } y \in M.$$

Es gilt der

Satz 3.4 *Ist (D) stark lösbar, dann ist (A) nicht lösbar.*

Beweis Es sei $x_r \in K$ für $r = 1, \ldots$, $\lim x_r = b$ und $\widehat{x}$ eine starke Lösung von (D). Dann gibt es ein $\delta > 0$ mit

$$\delta \leq \langle \widehat{x}, x_r - b \rangle \leq \|\widehat{x}\| \cdot \|x_r - b\| \longrightarrow 0,$$

ein Widerspruch. □

Satz 3.1 ist die Umkehrung von Satz 3.4.

Für praktische Zwecke hat die Aufgabe (D) eine gewisse Bedeutung. Wir wollen jetzt untersuchen, wie man (D) mit (VF) lösen kann, wenn (V) nicht erfüllt ist. Eine gewisse Rolle spielt dabei die Forderung

$$b \notin \text{cl } M,$$

die besagt, daß (A) nicht bereits auf M lösbar ist. Bei einigen Beweisen müssen wir voraussetzen, daß diese Forderung erfüllt ist, um gewisse Ausartungsfälle auszuschließen.

Es sei

$$P = \{x \in \mathcal{H} \mid \langle x - b, y - b \rangle > 0 \quad \text{für alle } y \in M\}.$$

Dann gilt der

Satz 3.5 *Es sei $b \notin$ cl M und P enthalte einen inneren Punkt. Dann ist (D) stark lösbar.*

Beweis Ist $b \notin \text{cl } M$, dann gibt es ein $\gamma > 0$, so daß $\phi(x) \geq \gamma$ für alle $x \in M$.

Sei z^0 innerer Punkt von P. Dann gibt es ein $\varepsilon > 0$, so daß $z^0 + \varepsilon \cdot x \in P$ für alle x mit $\|x\| = 1$, das heißt

$$0 < \left\langle z^0 + \varepsilon \cdot x - b, y - b \right\rangle = \left\langle z^0 - b, y - b \right\rangle + \varepsilon \cdot \langle x, y - b \rangle$$

für alle $y \in M$. Für festes $y \in M$ wählen wir $x = -\dfrac{y - b}{\|y - b\|}$, dann wird

$$\left\langle z^0 - b, y - b \right\rangle \geq \varepsilon \cdot \|y - b\| \geq \varepsilon \cdot \gamma > 0,$$

unabhängig von y. □

Gilt (V), dann ist Satz 3.5 umkehrbar:

Satz 3.6 *Es sei (V) erfüllt. Dann gilt:*

$$\text{int } P \neq \emptyset \text{ und } b \notin \text{cl } M \quad \Longleftrightarrow \quad \text{(D) ist stark lösbar.}$$

Beweis Wegen Satz 3.5 ist nur noch die Richtung „$\Longleftarrow$“ zu zeigen. Sei also (D) stark lösbar. Zunächst ist wegen Satz 3.4 (A) nicht lösbar, also $b \notin \text{cl } M$.

Sei $\widehat{x}$ starke Lösung von (D), das heißt $\langle \widehat{x}, y - b\rangle \geq \delta > 0$ für alle $y \in M$. Dann ist für irgendein x mit $\|x\| = 1$ und eine Zahl $\mu > 0$:

$$\begin{aligned} \langle \widehat{x} + \mu \cdot x, y - b\rangle &= \langle \widehat{x}, y - b\rangle + \mu \cdot \langle x, y - b\rangle \geq \delta - \mu \cdot \|x\| \cdot \|y - b\| \geq \\ &\geq \delta - \mu \cdot m \geq \delta - \frac{\delta}{2} = \frac{\delta}{2} > 0, \end{aligned}$$

falls $\mu \leq \dfrac{\delta}{2 \cdot m}$. □

Satz 3.5 ist sicher nicht allgemein umkehrbar:

Beispiel 3.1 *Es sei* $M = K = \left\{ x \in \mathbb{R}^2 \;\middle|\; x = \begin{pmatrix} 1 \\ \eta \end{pmatrix}, \eta \in \mathbb{R} \right\}$, $b = \Theta_2$.
Dann ist $P = \left\{ x \in \mathbb{R}^2 \;\middle|\; x = \begin{pmatrix} \xi \\ 0 \end{pmatrix}, \xi > 0 \right\}$, *also* int $P = \emptyset$.
Jedoch ist für $x = \begin{pmatrix} \xi \\ 0 \end{pmatrix} \in P$ *und* $y = \begin{pmatrix} 1 \\ \eta \end{pmatrix} \in M$*:*

$$\langle x, y\rangle = \xi > 0,$$

unabhängig von y*, also (D) stark lösbar, und es ist schließlich* $b \notin \text{cl } K$.

Wir definieren jetzt, falls $b \notin M$

$$\Pi M = \left\{ y \in \mathcal{H} \;\middle|\; y = \frac{x - b}{\|x - b\|} + b, x \in M \right\}, \tag{3.13}$$

die Projektion von M auf eine Sphäre vom Radius 1 mit dem Mittelpunkt b.

Unter Umständen kann man die Lösbarkeit von (D) bezüglich M auf die Lösbarkeit bezüglich ΠM zurückführen. ΠM ist beschränkt, so daß hier keine Schwierigkeiten auftreten.

Es gilt zunächst der einfache

Satz 3.7 *Es sei* $b \notin M$ *und*

$$P' = \{ x \in \mathcal{H} \mid \langle x - b, y - b\rangle > 0 \quad \textit{für alle } y \in \Pi M \}.$$

Dann ist $P' = P$.

Beweis $x \in P \Longleftrightarrow \langle x - b, y - b\rangle > 0$ für alle $y \in M$.
$\Longleftrightarrow \dfrac{\langle x - b, y - b\rangle}{\|y - b\|} > 0$ für alle $y \in M$
$\Longleftrightarrow \left\langle x - b, \dfrac{y - b}{\|y - b\|} + b - b \right\rangle > 0$ für alle $y \in M$
$\Longleftrightarrow \langle x - b, z - b\rangle > 0$ für alle $y \in \Pi M$. □

Satz 3.8 *Sei $b \notin \text{cl } M$. Dann gilt:*
(D) stark lösbar über $\Pi M \Longrightarrow$ (D) stark lösbar über M und (A) nicht lösbar.

Beweis Wegen $b \notin \text{cl } M$ gibt es ein $\gamma > 0$ mit $\phi(x) \geq y$ für alle $x \in M$. Weil (D) stark lösbar ist über ΠM, gibt es ein $\delta > 0$ und ein $\widehat{x} \in M$, so daß

$$\frac{\langle \widehat{x}, y - b \rangle}{\|y - b\|} \geq \delta \quad \text{für alle } y \in M,$$

$\Longrightarrow \langle \widehat{x}, y - b \rangle \geq \delta \cdot \gamma > 0 \quad$ für alle $y \in M$
$\Longrightarrow$ (D) stark lösbar über M.

Mit Satz 3.4 gilt: (A) nicht lösbar. □

Ist $b \notin \text{cl } M$ und ist $\widehat{x}$ eine starke Lösung von (D) bezüglich ΠM, dann ist natürlich $\widehat{x}$ auch eine starke Lösung von (D) bezüglich M. Es gilt sogar

Satz 3.9 *Sei $b \notin \text{cl } M$ und $\text{int } P \neq \emptyset$. Dann ist (D) stark lösbar durch ein $\widehat{x}$ mit $\widehat{x} = x - b$, $x \in K$.*
x kann durch (VF) ermittelt werden.

Beweis Wie wir im Anschluß an Satz 3.3 gesehen haben, ist die Aussage des Satzes wahr, wenn (V) gilt.

Sei (V) nicht erfüllt. Mit Satz 3.6 und Satz 3.7 folgt, daß (D) stark lösbar ist bezüglich ΠM. Die durch (VF) ermittelte Lösung von (D) hat die Form $\widehat{y} = y - b$, wo $y \in \text{conv } \Pi M$, also

$$y = \sum \frac{p_j}{\|x_j - b\|} \cdot x_j - b \cdot \sum \frac{p_j}{\|x_j - b\|} \cdot x_j + b,$$

wo $x_j \in M$, $p_j \geq 0$ und $\sum p_j = 1$. Wir setzen

$$p = \sum \frac{p_j}{\|x_j - b\|}$$

und

$$q_j = \frac{1}{p} \cdot \frac{p_j}{\|x_j - b\|},$$

dann ist $q_j \geq 0$ und $\sum q_j = 1$ und $x = \sum q_j \cdot x_j \in K$, also

$$\widehat{x} = \frac{1}{p} \cdot \widehat{y} = \frac{1}{p} \cdot (y - b) = x - b, \quad x \in K.$$

□

Es soll jetzt and einem Beispiel gezeigt werden, daß die Konvergenzabschätzung (3.7) streng ist:

Beispiel 3.2 *Es sei $\{e_k\}_{k=1}^{\infty}$ ein Orthonormalsystem in $\mathcal{H}$. Wir setzen $M = \{e_k\}$ und $b = \Theta_{\mathcal{H}}$. Dann wird in (V) $m = 1$. Mit $x_1 = e_1$ wird $\Phi(x_1, e_j) = 0$ für $j > 1$. Wählen wir $y_1 = e_2$, wird $x_2 = \frac{1}{2} \cdot (e_1 + e_2)$ und allgemein, wie man durch Induktion nachweist, mit $y_r = e_{r+1}$:*

$$x_r = \frac{1}{r} \cdot \sum_{j=1}^{r} e_j,$$

also

$$\phi(x_r) = \frac{1}{r}.$$

Schärfere Konvergenzasussagen erhält man, wenn man die Lage von b bezüglich M geeignet charakterisiert. Vorher soll jedoch an einigen Beispielen gezeigt werden, wie die Methode anwendbar ist. In den meisten Anwendungen ist K als lineares Bild einer konvexen Menge gegeben. Man versucht also, Näherungslösungen einer linearen Operatorgleichung

$$Au = b, \qquad u \text{ aus einer konvexen Menge}$$

zu finden. Ist diese konvexe Menge zu einfach gebaut, etwa eine lineare Mannigfaltigkeit, dann gibt es bessere Verfahren als (VF). Allerdings sei angemerkt, daß (VF) ein besonders einfaches Verfahren ist und daß es in vielen Fällen eine außerordentlich große numerische Stabilität besitzt.

Kapitel 4

Lineare Integralgleichungen mit positivem Kern

Es sei I das Intervall $[0,1]$, $\mathcal{H} = L^2(I)$, $A : \mathcal{H} \longrightarrow \mathcal{H}$ ein linearer Integraloperator mit Kern $K(s,t) \in L^2(I \times I)$, so daß

$$K(s,t) > 0 \text{ fast überall in } I \times I. \tag{4.1}$$

Gesucht ist eine Lösung u, λ von

$$B_\lambda u := Au - \lambda \cdot u = \int_0^1 K(s,t) \cdot u(t)\, dt - \lambda \cdot u(s) = \Theta_{\mathcal{H}} \tag{4.2}$$

mit $u \geq 0$ fast überall in I und $\int_0^1 u(t)\, dt = 1$.

Die folgende Aussage ist aus der Perron-Frobenius-Theorie für positive Operatoren bekannt:

Satz 4.1 *Es gelte (4.1). Dann ist (4.2) eindeutig lösbar durch ein $u^* \in \mathcal{H}$ und ein λ^*. Zusätzlich gilt $\lambda^* > 0$.*

Beweis Vgl. [25]. □

Wir beweisen einige Hilfssätze:

Hilfssatz 4.1 *Es sei $\lambda > 0$, $0 < K(s,t) \leq k$ fast überall in $I \times I$, $u(t) \in \mathcal{H}$ mit $u(t) \geq 0$ fast überall in I und $\int_0^1 u(t)\, dt = 1$. Dann gilt*

1. *$u^*(t) \leq k/\lambda^*$ fast überall in I,*
2. *$\|B_\lambda u\|^2 \leq C \Longrightarrow \|u\|^2 \leq \dfrac{C + 2 \cdot \lambda \cdot k}{\lambda^2}$.*

Beweis Es ist $B_{\lambda^*} u^* = \Theta_{\mathcal{H}}$, also

$$B_{\lambda^*} u^* \leq k - \lambda^* \cdot u^* \text{ fast überall in } I,$$

daraus folgt die Behauptung 1.

Aus $C \geq \|B_\lambda u\|^2 = \|Au - \lambda \cdot u\|^2 = \|Au\|^2 + \lambda^2 \cdot \|u\|^2 - 2 \cdot \lambda \cdot \langle u, Au\rangle \geq \lambda^2 \cdot \|u\|^2 - 2 \cdot \lambda \cdot k$ folgt die Behauptung 2. □

Wir definieren:

$$C = \left\{ u \in \mathcal{H} \;\middle|\; u(t) \geq 0 \;\text{ f. ü. in } I, \int_0^1 u(t)\, dt = 1 \right\}$$

und für eine feste Zahl m:

$$C_m = C \cap \{u \in \mathcal{H} \mid u(t) \leq m \;\text{ f. ü. in } I\}.$$

Hilfssatz 4.2 *Sei $0 < \lambda \neq \lambda^*$. Dann gibt es eine positive Zahl γ, so daß $\|B_\lambda u\| \geq \gamma$ für alle $u \in C$.*

Beweis Andernfalls gäbe es eine Folge $\{u_r\}$ mit $\|B_\lambda u_r\| \longrightarrow 0$. Nach Hilfssatz 4.1 ist dann die Folge $\{\|u_r\|\}$ von oben beschränkt, etwa durch m. Damit gibt es [1, §24] einen schwachen Limes u einer Teilfolge $\{u_{r_k}\}$. Nach dem Satz von Banach und Saks [94, §38] ist $u \in \text{cl conv } (\{u_{r_k}\})$. Weil C_m abgeschlossen und konvex ist, liegt u in C_m.

Behauptung: $\|B_\lambda u\| = 0$, das heißt, $u = u^*$ und $\lambda = \lambda^*$, im Widerspruch zur Voraussetzung.

Um dieses zu beweisen, setzen wir $v = B_\lambda^* B_\lambda u$, wo B_λ^* der zu B_λ adjungierte Operator ist. Wegen der schwachen Konvergenz der u_{r_k} gegen u gilt

$$\lim \langle v, u_{r_k}\rangle = \langle v, u\rangle = \langle B_\lambda^* B_\lambda u, u\rangle = \|B_\lambda u\|^2$$

und

$$|\langle v, u_{r_k}\rangle| = |\langle B_\lambda u, B_\lambda u_{r_k}\rangle| \leq \|B_\lambda u\| \cdot \|B_\lambda u_{r_k}\| \longrightarrow 0,$$

also $\|B_\lambda u\| = 0$. □

Der folgende Satz motiviert die Anwendung von (VF):

Satz 4.2 *Es sei $m \geq u^*$ fast überall in I.*
Folgende Aussagen sind äquivalent:

1. *$0 \leq \lambda \neq \lambda^*$,*
2. *Es gibt ein $u \in C_m$, so daß $\langle B_\lambda u, B_\lambda v\rangle > 0$ für alle $v \in C_m$,*
3. *Es gibt ein $u \in C_m$ und eine positive Zahl δ, so daß $\langle B_\lambda u, B_\lambda v\rangle \geq \delta$ für alle $v \in C_m$.*

Beweis Klar ist die Implikationskette $3 \Longrightarrow 2 \Longrightarrow 1$. Wir zeigen jetzt $1 \Longrightarrow 3$. Dazu haben wir nur zu zeigen, daß die durch $M = B_\lambda C_m$ und $b = \Theta_\mathcal{H}$ definierte Aufgabe (A) nicht lösbar ist. Die Aussage 3 folgt dann aus Satz 3.3.

Hilfssatz 4.2 besagt aber gerade, daß (A) nicht lösbar ist, wenn $\lambda \neq \lambda^*$. □

Satz 4.2 gibt uns eine Möglichkeit, λ^* abzuschätzen. Es sei $m \geq u^*$ fast überall in I, $u_0 \in C_m$ und R_0 irgendeine Menge positiver Zahlen, die λ^* enthält. Schließlich sei $\lambda_0 \in R_0$ irgendein Startwert. Auf das durch $M = B_{\lambda_0} C_m$ und $b = \Theta_\mathcal{H}$ definierte Aufgabenpaar (A), (D) wenden wir (VF) an. Folgende Fälle können eintreten:

1. Ist zufällig $\lambda_0 = \lambda^*$, dann liefert (VF) eine Folge $\{u_r\}$, $u_r \in C_m$, $B_{\lambda_0} u_r \longrightarrow \Theta_{\mathcal{H}}$. Wie wir beim Beweis von Hilfssatz 4.2 gesehen haben, hat jede schwach konvergente Teilfolge von $\{u_r\}$ u als Grenzwert.

2. Andernfalls liefert (VF) nach endlich vielen Schritten ein $\overline{u} \in C_m$, so daß $\langle B_{\lambda_0}\overline{u}, B_{\lambda_0} v\rangle > 0$ für alle $v \in C_m$. Mit Satz 4.2 und Hilfssatz 4.2 folgt, daß (A) nicht lösbar ist. Iterieren wir gemäß Satz 3.3 weiter, dann finden wir nach endlich vielen Schritten ein $\widetilde{u} \in C_m$ und ein $\tau > 0$, so daß

$$\langle B_{\lambda_0}\widetilde{u}, B_{\lambda_0} v\rangle \geq \tau \quad \text{für alle } v \in C_m. \tag{4.3}$$

   Setzen wir für irgendein $u \in \mathcal{H}$

$$\overline{\Lambda}_u = \complement\{\lambda \in R \mid \langle B_\lambda u, B_\lambda v\rangle > 0 \quad \text{für alle } v \in C_m\} \tag{4.4}$$

   $\complement$ bezeichnet das mengentheoretische Komplement), dann ist $\lambda^* \in \overline{\Lambda}_{\widetilde{u}}$, $\lambda_0 \notin \overline{\Lambda}_{\widetilde{u}}$, folglich ist $\lambda^* \in R_0 \cap \overline{\Lambda}_{\widetilde{u}}$, und diese Menge ist echt in R_0 enthalten.

Diese Vorgehensweise hat einige Nachteile. Zunächst ist es keineswegs einfach, $\overline{\Lambda}_{\widetilde{u}}$ zu bestimmen. Wir werden deshalb in Abschnitt 4.2 untersuchen, ob man einfache Obermengen von $\overline{\Lambda}_u$ bestimmen kann.

Weiterhin kann die Anzahl der Iterationsschritte, die bis zum Abbruch erforderlich sind, sehr groß werden, insbesondere wenn λ_0 in der Nähe von λ^* liegt. Mann kan nun versuchen, die Tatsache auszunutzen, daß bei jedem Schritt der Verfahrens (VF) $\|B_{\lambda_0} u_r\|$ kleiner wird. Schließt man λ^* vermittels eines einfachen Einschließungssatzes unter Benutzung von u_r und Au_r ein – diese beiden Funktionen stehen bei jedem Iterationsschritt ohnehin zur Verfügung – dann kann man erwarten, daß die aus u_{r+1} und Au_{r+1} gewonnene Einschließung besser ist. Die Abschätzung von λ^* vermittels $\overline{\Lambda}_u$ (oder einer Obermenge von $\overline{\Lambda}_u$) bleibt dann als „ultima ratio“. Zwei Beispiele in Abschnitt 4.3 werden zeigen, daß diese Vorgehensweise vorteilhaft ist.

Es sei angemerkt, daß keine Konvergenzaussagen gegeben werden. Es wird lediglich gezeigt, daß man, ausgehend von einer Menge R_0, die λ^* enthält, stets eine echte Teilmenge von R_0 finden kann, die ebenfalls λ_0 enthält.

4.1 Ein Einschließungssatz

Der folgende Einschließungssatz geht auf Collatz [21] zurück:

Satz 4.3 *Es sei $u \in \mathcal{H}$.*

$$\alpha = \operatorname{ess.}\inf_{t\in I} \frac{A^*u}{u}, \qquad \beta = \operatorname{ess.}\sup_{t\in I} \frac{A^*u}{u}.$$

Setzen wir

$$\Lambda_u = [\alpha, \beta]\,,$$

dann ist $\lambda^* \in \Lambda_u$.

Beweis Sei etwa $-\infty < \lambda < \alpha$. Dann ist $\lambda u < A^* u$ für fast alle $t \in I$. Daraus folgt $B_\lambda^* u > 0$ für fast alle $t \in I$, also $\langle B_\lambda^* u, v\rangle > 0$ für alle $v \in C$, das heißt, $\langle u, B_\lambda v\rangle > 0$ für alle $v \in C$, folglich $\lambda \neq \lambda^*$. Ebenso argumentiert man für $\lambda > \beta$. □

Die Menge λ_u ist ein (möglicherweise unendliches) Intervall. Sie hat eine interessante Eigenschaft:

Satz 4.4 *Es sei für ein $u \in C$*

$$\lambda_u = \frac{\langle u, Au\rangle}{\langle u, u\rangle}.$$

Dann ist $\lambda_u \in \Lambda_u$ und $\|B_{\lambda_u} u\| = \min_\lambda \|B_\lambda u\|$.

Beweis Wäre $\lambda_u \notin \Lambda_u$, also etwa $\lambda_u < A^* u/u$ für fast alle $t \in I$, dann wäre $B_{\lambda_u}^* u > 0$ für fast alle $t \in I$. Wegen $u \in C$ folgt $\left\langle B_{\lambda_u}^* u, u\right\rangle = \langle u, B_{\lambda_u} u\rangle > 0$, jedoch ist

$$\langle u, B_{\lambda_u} u\rangle = \langle u, Au\rangle - \frac{\langle u, Au\rangle}{\langle u, u\rangle} \cdot \langle u, u\rangle = 0,$$

ein Widerspruch.

Analog verfährt man für $\lambda_u > A^* u/u$ für fast alle $t \in I$. Die letzte Aussage des Satzes folgt durch Differenzieren von $\|B_\lambda u\|^2$ nach λ und Nullsetzen der ersten Ableitung. □

Wir betrachten zwei Beispiele für die Awendung der Sätze 4.3 und 4.4:

Beispiel 4.1

$$K(s,t) = \begin{cases} s \cdot (1-t) & \text{für } 0 \le s \le t \le 1, \\ t \cdot (1-s) & \text{für } 0 \le t \le s \le 1 \end{cases}$$

[25, Beispiel 6]. Hier ist $\lambda^* = \dfrac{1}{\pi^2} = 0.101\ 321$.

Wählen wir die Ansatzfunktion $u = 1$, so wird die Einschließung nach Satz 4.3:

$$0 \le \lambda^* \le \frac{1}{8}.$$

Es ist $\lambda_u = \dfrac{1}{12}$ und $\|B_{\lambda_u} u\|^2 = \dfrac{1}{720} = 0.001\ 389$.

Mit der Ansatzfunktion $u = 6 \cdot t \cdot (1-t)$ erhalten wir

$$0.0833 \le \frac{1}{12} \le \lambda^* \le \frac{5}{48} = 0.1042$$

und $\lambda_u = \dfrac{17}{168} = 0.101\ 190$ sowie $\|B_{\lambda_u} u\|^2 = 0.000\ 014$.

Beispiel 4.2 *$K(s,t) = s + t$, $s, t \in I$ [25, Beispiel 15]. Es ist $\lambda^* = 1.077\ 35$.*

Wir geben die Einschließung von λ^* für einige Ansatzfunktionen in einer Tabelle:

$u(t)$	α	β	λ_u	$\|B_{\lambda_u}u\|^2$
1	0.5	1.5	1	0.083 333
t	0.833	∞	1	0.027 778
$1+t$	0.833	1.167	1.071 428	0.015 873
$1+t+t^2$	0.972	1.176	1.073 574	0.015 620
e^t	1.0	1.132	1.075 766	0.005 507

4.2 Bestimmung einer Obermenge von $\overline{\Lambda}_u$

Es sei $u \in \mathcal{H}$, $\lambda \in R$ und $\tau > 0$ gegeben, so daß $\langle B_\lambda u, B_\lambda v\rangle \geq \tau$ für alle $v \in C_m$ ist, das heißt, es liege die Situation (4.3) vor.

Satz 4.5 *Es sei*

$$\begin{aligned} \alpha &\leq \min_{v\in C_m} \langle B_\lambda u, v\rangle\,, \\ \beta &\geq \max_{v\in C_m} \langle B_\lambda u, v\rangle\,. \end{aligned} \tag{4.5}$$

Weiterhin sei die Menge J gemäß

$$J = \begin{cases} \left(\lambda + \dfrac{\tau}{2\cdot\alpha}, \lambda + \dfrac{\tau}{2\cdot\beta}\right) & \text{für } \alpha < 0 < \beta, \\ \left(-\infty, \lambda + \dfrac{\tau}{2\cdot\beta}\right) & \text{für } \alpha \geq 0, \beta > 0, \\ \left(\lambda + \dfrac{\tau}{2\cdot\alpha}, +\infty\right) & \text{für } \beta \leq 0, \alpha < 0, \\ (-\infty, +\infty) & \text{für } \alpha = \beta = 0 \end{cases}$$

definiert und $\Lambda_u^{\alpha,\beta} = \complement J$.

Dann ist $\lambda^* \in \overline{\Lambda}_u \subseteq \Lambda_u^{\alpha,\beta}$.

Beweis Für $v \in C_m$ ist

$$\begin{aligned} \langle B_{\lambda+\mu}u, B_{\lambda+\mu}v\rangle &= \langle B_\lambda u, B_\lambda v\rangle - 2\cdot\mu\cdot\langle B_\lambda u, v\rangle + \mu^2\cdot\langle u, v\rangle \geq \\ &\geq \tau - 2\cdot\mu\cdot\langle B_\lambda u, v\rangle \geq \tau - 2\cdot\mu\cdot\beta > 0, \end{aligned}$$

falls $\langle B_\lambda u, v\rangle > 0$ und $\mu < \dfrac{\tau}{2\cdot\beta}$

Analog beweist man die anderen Fälle. □

Die folgenden beiden Sätze zeigen, wie man auf einfache Weise Schranken α und β finden kann:

Satz 4.6 *Die Zahlen*

$$\alpha = -m \cdot \|B_\lambda u\| \qquad \textit{und} \qquad \beta = +m \cdot \|B_\lambda u\|$$

genügen den Bedingungen (4.5).

Beweis Es ist wegen $v \in C_m$

$$|\langle B_\lambda u, v\rangle| \le \|B_\lambda u\| \cdot \|v\| \le m \cdot \|B_\lambda u\|.$$

□

Satz 4.7 *Setzt man*

$$N(u) = \{t \in I \mid (B_\lambda u)(t) < 0\}$$

und

$$P(u) = \{t \in I \mid (B_\lambda u)(t) > 0\},$$

dann genügen die Zahlen

$$\alpha = m \cdot \int_{N(u)} B_\lambda u \; dt \qquad \textit{und} \qquad \beta = m \cdot \int_{P(u)} B_\lambda u \; dt$$

den Bedingungen (4.5).

Beweis Sei $v \in C_m$, $f(t) = (B_\lambda u)(t)$. Dann ist

$$\langle f, v\rangle = \int_0^1 f(t) \cdot v(t) \; dt = \int_{P(u)} f \cdot v \; dt + \int_{N(u)} f \cdot v \; dt \le m \cdot \int_{P(u)} f(t) \; dt.$$

Analog zeigt man die Abschätzung für α. □

Beispiel 4.3 *Wir betrachten den Kern des Beispiels 4.2:* $K(s,t) = s + t$, $s, t \in I$.

Wählen wir $u \equiv 1$ *in* I, *dann wird* $B_\lambda u = t + \frac{1}{2} - \lambda$ *und*

$$\langle B_\lambda u, B_\lambda v\rangle = \frac{7}{12} - \lambda + \lambda^2 + \int_0^1 t \cdot v(t) \; dt - \lambda \cdot \int_0^1 t \cdot v(t) \; dt \ge \frac{7}{12} - 2 \cdot \lambda + \lambda^2.$$

Für $\lambda = 0$ *kann man somit* $\tau = \frac{7}{12}$ *wählen. Aus Satz 4.6 erhält man*

$$\begin{aligned} -\frac{m}{\sqrt{12}} \ge -0.289 \cdot m &=: \alpha, \\ 0.289 \cdot m &=: \beta. \end{aligned}$$

Eine untere Schranke für λ^* *ergibt sich aus* β *zu*

$$1.010 \cdot \frac{1}{m} \le \lambda^*.$$

Man benötigt also eine Abschätzung für m.

Ähnlich geht es mit Satz 4.7. Es ist hier $P(u) = I$ und $N(u) = \emptyset$, folglich

$$\beta = m \cdot \int_0^1 B_0 u \; dt = m \cdot \int_0^1 \left(t + \frac{1}{2} \right) \; dt = m,$$

daraus wird

$$0.291 \cdot \frac{1}{m} < \frac{1}{24 \cdot m} \leq \lambda^*.$$

Diese Abschätzung ist wesentlich schlechter als die nach Satz 4.6.

Bei dem vorliegenden Beispiel kann man $\overline{\Lambda}_u$ errechnen, es ist (für $m = \infty$)

$$\overline{\Lambda}_u = [0.683, 2.317]\,.$$

4.3 Beispiele

An zwei Beispielen soll die Wirksamkeit der beschriebenen Methode demonstriert werden. Es empfiehlt sich, nach Möglichkeit die einfachere Einschließung nach Satz 4.3 zu verwenden. Damit bietet sich die folgende Vorgehensweise an: Sei $u_0 \in C_m$ gewählt, R_0 eine Menge von positiven Zahlen, die λ^* enthält, etwa $R_0 = \Lambda_{u_0}$ gemäß Satz 4.3. Weiterhin sei $\lambda_0 \in R_0$ gewählt, etwa $\lambda_0 = \lambda_{u_0}$ gemäß Satz 4.4. Gibt es ein $v \in C_m$, so daß $\langle B_{\lambda_0} u_0, B_{\lambda_0} v \rangle = \left\langle B^*_{\lambda_0} B_{\lambda_0} u_0, v \right\rangle \leq 0$, dann ist $\lambda_0 \in \overline{\Lambda}_{u_0}$, wir können also einen Iterationsschritt mit (VF) ausführen und erhalten u_1. Gibt es kein solches v, dann ist $\lambda_0 \notin \overline{\Lambda}_{u_0}$. In diesem Falle iterieren wir gemäß Satz 3.3 und erhalten nach endlich vielen Schritten ein $u^{(r)} =: u_1$, so daß (4.3) gilt. Mit Satz 4.5 (und Satz 4.6 beziehungsweise Satz 4.7) berechnen wir ein Intervall $\Lambda^{\alpha,\beta}_{u_1}$, so daß $\lambda_0 \notin \Lambda^{\alpha,\beta}_{u_1}$. Wir setzen dann $R'_0 = R_0 \cap \Lambda^{\alpha,\beta}_{u_1}$. Mit u_1 und $\lambda_1 \in R_0$ (beziehungsweise R'_0) können wir den beschriebenen Vorgang wiederholen.

Beispiel 4.4 *Wir untersuchen hier wieder den Kern des Beispiels 4.2: $K(s,t) = s + t$, $s, t \in I$.*

Mit $u_0 \equiv 1$ wird nach Satz 4.3:

$$\frac{1}{2} \leq \lambda^* \leq \frac{3}{2},$$

$\lambda_0 := \lambda_{u_0} = 1$ *und* $\|B_{\lambda_0} u_0\|^2 = \dfrac{1}{12}$.

Wählen wir $v_0 = 2 \cdot t$ und iterieren einmal mit (VF), wird

$$u_1 = \frac{7}{13} + \frac{12}{13} \cdot t$$

mit

$$\|B_{\lambda_0} u_1\|^2 = \frac{1}{156}.$$

Mit u_1 erhalten wir die Einschließung nach Satz 4.3:

$$1.0714 \leq \frac{15}{14} \leq \lambda^* \leq \frac{41}{38} \leq 1.0790,$$

sowie $\lambda_1 = \lambda_{u_1} = \dfrac{195}{181} = 1.077\ 348$ *und* $\|B_{\lambda_1} u_1\|^2 = \dfrac{1}{367\ 068}$.

Beispiel 4.5 $K(s,t) = |s-t|$, $s,t \in I$ *[25, Beispiel 14].* $\lambda^* = 0.347\ 408$.

Wir beginnen wieder mit $u_0 \equiv 1$. *Nach Satz 4.3 wird*

$$\frac{1}{4} \leq \lambda^* \leq \frac{1}{2}.$$

$\lambda_0 = \lambda_{u_0} = \dfrac{1}{3}$ *und* $\|B_{\lambda_0} u_0\|^2 = \dfrac{1}{180}$.

Setzen wir $v_0 = (2 \cdot t - 1)^2$, *dann ist*

$$u_1 = 0.551\ 948 + 0.448\ 052 \cdot (2 \cdot t - 1)^2$$

und $\|B_{\lambda_0} u_1\|^2 = 2.8 \cdot 10^{-6}$. *Satz 4.3 liefert mit* u_1 *die Abschätzung*

$$0.344\ 848 \leq \lambda^* \leq 0.351\ 458$$

und $\lambda_1 := \lambda_{u_1} = 0.347\ 399$ *mit* $\|B_{\lambda_1} u_1\|^2 < 10^{-6}$.

Kapitel 5

Anwendung auf Quadraturformeln

Es sei $\mathcal{H} = \mathcal{H}_2$ der Hardysche Hilbertraum aller im Inneren des Einheitskreises der komplexen Zahlenebene holomorphen Funktionen $f(z) = \sum_{j=0}^{\infty} a_j \cdot z^j$, so daß

$$\|f\|^2 = \sum_{j=0}^{\infty} |a_j|^2 < \infty.$$

(Zu diesem Raum vgl. [92, 80]). Sind $f = \sum a_j \cdot z^j$ und $g = \sum b_j \cdot z^j$ zwei Elemente aus $\mathcal{H}_2$, dann ist

$$\langle f, g \rangle = \sum a_j \cdot \overline{b}_j.$$

Der Einfachheit halber beschränken wir uns auf Funktionen mit reellen Entwicklungksoeffizienten a_j.

$\mathcal{H}_2$ ist isomorph zu dem (reellen) Hilbertschen Folgenraum ℓ^2 vermöge der Identifizierung

$$\sum_{j=1}^{\infty} a_j \cdot z^j \in \mathcal{H}_2 \qquad \longleftrightarrow \qquad (a_j)_{j=0}^{\infty} \in \ell^2.$$

In $\mathcal{H}_2$ betrachten wir das stetige lineare Funktional

$$If = \int_0^1 f(t)\ dt.$$

Nach dem Satz von Riesz [1, Kap. II, Nr. 16] entspricht diesem Funktional ein eindeutig bestimmtes Element aus $\mathcal{H}_2$, das wir auch mit I bezeichnen wollen, so daß

$$If = \langle I, f \rangle \qquad \text{für alle } f \in \mathcal{H}_2.$$

Es ist

$$I = \frac{1}{z} \cdot \log \frac{1}{1-z} \in H_2 \qquad \longleftrightarrow \qquad I = \left(\frac{1}{k+1} \right)_{k=0}^{\infty} \in \ell^2.$$

Weiterhin betrachten wir die Punktfunktionale

$$\phi_t(f) = f(t), \qquad t \in [0, 1).$$

Die diesen Funktionalen entsprechenden Elemente aus $\mathcal{H}_2$ wollen wir mit $I(t)$ bezeichnen:

$$I(t) = \frac{1}{1 - z \cdot t} \in \mathcal{H}_2 \qquad \longleftrightarrow \qquad I(t) = (t^k) \in \ell^2.$$

Setzen wir

$$J_r = \sum_{j=1}^{r} p_j^{(r)} \cdot I\left(t_j^{(r)}\right),$$

dann wird eine Folge $\{J_r\}$ gesucht, so daß

$$I = \lim_{r \to \infty} J_r. \tag{5.1}$$

Die $t_j^{(r)}$ nennen wir die *Knoten* und die $p_j^{(r)}$ die *Gewichte* der *Quadraturformel* J_r. Sind hierbei alle $p_j^{(r)} \geq 0$, dann haben wir die besonders interessante Klasse der Quadraturformeln mit positiven Gewichten (vgl. [69, Kap. 12]).

Eine detaillierte Beschreibung der hier skizzierten Verhältnisse findet sich in [34].

Wilf [111] gab eine Folge $\left\{J_r^{(W)}\right\}$ von Quadraturformeln an mit

$$\left\|E_r^{(W)}\right\|^2 = \left\|I - J_r^{(W)}\right\|^2 = O\left(\frac{\log r}{r}\right).$$

Engels und Mangoldt [38] konstruierten eine Folge $\left\{J_r^{(E,M)}\right\}$ von Quadraturformeln mit

$$\left\|E_r^{(E,M)}\right\|^2 = \left\|I - J_r^{(E,M)}\right\|^2 = O\left(\frac{1}{r}\right). \tag{5.2}$$

Haber [52] und Johnson und Riess [65] erhielten die gleiche Konvergenzordnung für spezielle Folgen von Quadraturformeln. Die Abschätzung (5.2) erinnert an (3.7). Wir wollen jetzt mit (VF) eine Folge von Quadraturformeln konstruieren, die nach (5.2) konvergiert. Für $t \in [0, 1)$ sind die $I(t)$ nicht beschränkt. Um Satz 3.2 anwenden zu können setzen wir daher

$$M = \left\{ \frac{\pi}{2} \cdot I(t) \cdot \sqrt{1 - t^2} \;\middle|\; t \in [0, 1) \right\}$$

Dann ist für jedes $t \in [0, 1)$

$$\left\| \frac{\pi}{2} \cdot I(t) \cdot \sqrt{1 - t^2} \right\|^2 = \frac{\pi^2}{4} \cdot (1 - t)^2 \cdot \sum_{k=0}^{\infty} t^{2k} = \frac{\pi^2}{4},$$

also (V) erfüllt.

Die durch M und $b = I$ definierte Aufgabe (D) ist nicht lösbar. Wäre nämlich $a = (a_j)_{j=1}^{\infty} \in \ell^2$ eine Lösung von (D), dann wäre für alle $t \in [0, 1)$

$$\left\langle a, \frac{\pi}{2} \cdot \sqrt{1 - t^2} \cdot I(t) - I \right\rangle > 0,$$

und wenn wir $f(z) = \sum a_j \cdot z^j$ setzen:

$$\frac{\pi}{2} \cdot \sqrt{1 - t^2} \cdot f(t) - \int_0^1 f(s)\, ds > 0 \qquad \text{für alle } t \in [0, 1).$$

Daraus folgt

$$0 < \frac{\pi}{2} \cdot \int_0^1 f(t)\, dt - \int_0^1 \frac{dt}{\sqrt{1-t^2}} \cdot \int_0^1 f(s)\, ds = 0,$$

ein Widerspruch.

Damit bricht (VF), angewandt auf M mit $b = I$ nicht ab, es gilt also der

Satz 5.1 *Es gibt eine Folge $\{J_r\}$ von Quadraturformeln mit positiven Gewichten, so daß*

$$\|I - J_r\|^2 \leq \frac{\pi^2}{4r}. \tag{5.3}$$

Das Verfahren (VF) gibt uns die Möglichkeit, Quadraturformeln zu berechnen, die gemäß (5.3) konvergieren.

Diese Quadraturformeln haben die zusätzliche Eigenschaft, daß sie die Funktion

$$g(t) = \frac{1}{\sqrt{1-t^2}}$$

exakt quadrieren. Es sei darauf hingewiesen, daß $g \notin \mathcal{H}_2$.

Ist f irgendein Element aus $\mathcal{H}_2$, dann kann man also den Fehler $E_r f$, der bei der numerischen Quadratur von f mit der Quadraturformel J_r entsteht, folgendermaßen abschätzen:

$$|E_r f| \leq \|E_r\| \cdot \|f\| \leq \frac{\pi}{2\sqrt{r}} \cdot \|f\|,$$

die Abschätzung hängt also nur von r und $\|f\|$, insbesondere nicht von Schranken für irgendwelche Ableitungen von f ab.

Kapitel 6

Die Öffnung der erzeugenden Menge

Es sei $M \subseteq \mathcal{H}$ eine Menge, $K = \operatorname{conv} M$ und $b \in \mathcal{H}$.

Definition 6.1 *Die* Öffnung *der erzeugenden Menge M bezüglich b sei definiert als*

$$\alpha_b(M) = \sup_{\substack{x \in K \\ x \neq b}} \inf_{\substack{y \in M \\ y \neq b}} \frac{\langle x - b, y - b \rangle}{\|x - b\| \cdot \|y - b\|}.$$

Wir veranschaulichen uns die geometrische Bedeutung von $\alpha_b(M)$ im $\mathbb{R}^2$. Ist $b \in \operatorname{cl} M$, dann ist auch $b \in \operatorname{cl} K$ und somit die Aufgabe (A) lösbar. Wir nehmen an, daß dieser (triviale) Fall nicht vorliege, es sei also $b \notin \operatorname{cl} M$. Für festes x und y ($x \neq b$, $y \neq b$) ist

$$c(x, y) = \frac{\langle x - b, y - b \rangle}{\|x - b\| \cdot \|y - b\|}$$

der Cosinus des Winkels zwischen den Vektoren $x - b$ und $y - b$.

Für festes $x \neq b$ ist die Menge

$$C(x) = \left\{ x \in \mathbb{R}^2 \mid y \neq b, \quad c(x, y) \geq \inf_{z \in M} c(z, y) \right\} \cup \{b\}$$

ein (nicht notwendig konvexer) Kegel mit dem Scheitel b und der Symmetrieachse $x - b$, der M enthält. $\inf_{z \in M} c(x, z)$ ist der Cosinus des halben Öffnungswinkels des Kegels $C(x)$. Dann ist $\alpha_b(M)$ gerade der Cosinus des halben Öffnungswinkels $w_b(M)$ des kleinsten solchen Kegels, dessen Achse zusätzlich noch $\operatorname{cl} K$ schneidet.

Ist der Winkel $w_b(M)$ spitz (das heißt, $\alpha_b(M) > 0$), dann ist $b \notin \operatorname{cl} K$ (Abbildung 6.1), ist er stumpf (das heißt, $\alpha_b(M) < 0$), dann ist $b \in \operatorname{relint} K$ (Abbildung 6.2). Im Falle $\alpha_b(M) = 0$ liegt b gerade auf dem Rande von $\operatorname{cl} K$, und $w_b(M)$ ist ein rechter Winkel (Abbildung 6.3).

Die Zahl $\alpha_b(M)$ werden wir dazu benutzen, die Konvergenzgeschwindigkeit von (VF) zu bestimmen. Zunächst jedoch wollen wir den eben skizzierten geometrischen Sachverhalt präzisieren.

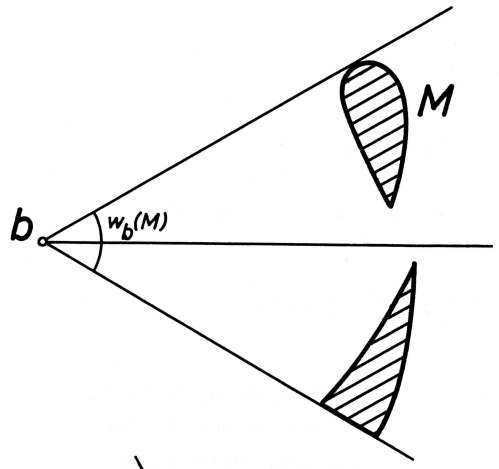

Abb. 6.1: $\alpha_b(M) > 0$

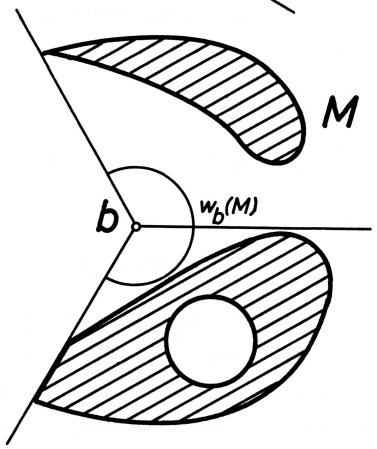

Abb. 6.2: $\alpha_b(M) < 0$

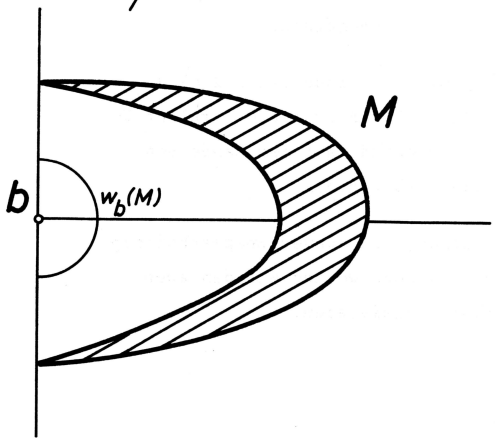

Abb. 6.3: $\alpha_b(M) = 0$

Zuerst beweisen wir den

Hilfssatz 6.1 *Es gelte für ein $\widehat{x} \in \mathcal{H}$ und eine Zahl $\delta > 0$:*

$$\langle \widehat{x}, y - b \rangle \geq \delta \cdot \|y - b\| \qquad \text{für alle } y \in M.$$

Dann gilt diese Ungleichung auch für alle $y \in K$.

Beweis Sie $y \in K$, das heißt, $y = \sum p_j \cdot y_j$, $p_j \geq 0$, $\sum p_j = 1$, $y_j \in M$. Dann ist

$$\|y - b\| = \Big\| \sum p_j \cdot (y_j - b) \Big\| \leq \sum p_j \cdot \|y_j - b\|,$$

damit

$$\langle \widehat{x}, y - b \rangle = \sum p_j \cdot \langle \widehat{x}, y_j - b \rangle \geq \delta \cdot \sum p_j \cdot \|y_j - b\| \geq \delta \cdot \|y - b\|.$$

□

Damit gilt der

Satz 6.1 *Sei $b \notin \text{cl } M$. Dann ist*

1. $\alpha_b(\text{cl } M) = \alpha_b(M)$,
2. $\alpha_b(K) \leq \alpha_b(M)$,
3. $\alpha_b(M) > 0 \quad \Longrightarrow \quad \alpha_b(M) = \alpha_b(K)$.

Beweis 1. Klar!
2. folgt direkt aus der Definition.
3. Zu jedem $\varepsilon > 0$, $\varepsilon < \alpha_b(M)$ gibt es ein $x \in K$, $x \neq b$, so daß

$$\frac{\langle x - b, y - b \rangle}{\|x - b\| \cdot \|y - b\|} \geq \alpha_b(M) - \varepsilon \qquad \text{für alle } y \in M \ (y \neq b).$$

Aus dem Satz 2.1 folgt, daß diese Ungleichung auch für alle $y \in K$ gilt, das heißt also

$$\alpha_b(K) \geq \alpha_b(M) - \varepsilon.$$

Mit 1. folgt die Behauptung. □

Der Fall $\alpha_b(M) = 0$ und $\alpha_b(K) < 0$ kann tatsächlich eintreten:

Beispiel 6.1 *Es sei $\mathcal{H} = \ell^2$ der Hilbertsche Folgenraum, e_k das Element aus ℓ^2, dessen k-te Komponente 1 ist, alle anderen Komponenten mögen verschwinden. Setzen wir $b = \Theta_{\mathcal{H}}$ und $M = \{\mu \cdot e_k \mid \mu \in \mathbb{R},\ k = 1, 2, \dots\}$, dann ist $K = \text{conv } M$ die Menge aller derjenigen Elemente aus ℓ^2, bei denen höchstens endlich viele Komponenten nicht verschwinden.*

Es ist int $K = \emptyset$, cl $K = \ell^2$ *und* relint $K = K$.

Sei $x = (x_k) \in K$, $x \neq \Theta_{\mathcal{H}}$. *Wir wählen* $y_r = \left(y_k^{(r)}\right) \in K$ *gemäß*

$$y_k^{(r)} = \begin{cases} -x_k & \text{falls } k \leq r, \\ 0 & \text{sonst .} \end{cases}$$

Dann ist

$$\lim_{r\to\infty} \frac{\langle x, y_r\rangle}{\|x\| \cdot \|y_r\|} = -1,$$

also $\alpha_b(K) \leq -1$. *Nach Definition ist aber stets* $\alpha_b(K) \geq -1$, *also* $\alpha_b(K) = -1$.

Sei andererseits $y = \mu \cdot e_k \in M$, $\mu \neq 0$. *Dann ist* $\dfrac{y}{\|y\|} = \text{sign } (\mu) \cdot e_k$. *Wählen wir* $x_r = \left(x_k^{(r)}\right)$ *gemäß*

$$x_k^{(r)} = \begin{cases} \dfrac{1}{\sqrt{r}} & \text{falls } k \leq r, \\ 0 & \text{sonst,} \end{cases}$$

dann ist $\|x_r\| = 1$ *und* $x_r \in K$ *sowie*

$$\inf_{y\in M} \frac{\langle x_r, y\rangle}{\|x_r\| \cdot \|y\|} = -\frac{1}{\sqrt{r}},$$

also $\alpha_b(M) \geq 0$. *Mit Satz 6.1, Teil 3. wird* $\alpha_b(M) = 0$.

Das Vorzeichen von $\alpha_b(M)$ charakterisiert die Lage von b bezüglich K:

Satz 6.2 *Es gelte (V), das heißt, M sei beschränkt. Weiterhin sei $b \notin \text{cl } M$. Dann ist mit den Bezeichnungen des Abschnitts 2:*

1. $\alpha_b(M) > 0 \iff b \notin \text{cl } K$,
2. $b \in \text{relint } K \implies \alpha_b(M) < 0$,
   $\alpha_b(M) < 0 \implies b \in \text{relint cl } K$.
3. $\alpha_b(M) = 0 \implies b \in \text{relbd cl } K$,
   $b \in \text{relbd } K \implies \alpha_b(M) = 0$.

Beweis Zu 1. a) $\alpha_b(M) > 0 \Longrightarrow$ (D) ist stark lösbar über ΠM (vgl Abschnitt 3). Mit $b \notin \text{cl } M$ folgt, daß $b \notin \text{cl } K$ (Satz 3.8).
b) Ist umgekehrt $b \notin \text{cl } K$, so gilt wegen (V), daß (D) stark lösbar ist über M durch ein $\widehat{x} \in K$ (Satz 3.3):

$$\langle \widehat{x} - b, y - b\rangle \geq \gamma > 0 \qquad \text{für alle } y \in M.$$

Wegen (V) ist folglich $\alpha_b(M) > 0$.

Zu 2. a) Sei $b \in \text{relint } K$, L sei die kleinste lineare Mannigfaltigkeit, die K enthält. Dann gibt es ein $\varepsilon_0 > 0$, so daß

$$\|b - x\| \leq \varepsilon < \varepsilon_0 \text{ und } x \in L \implies x \in K.$$

Sei $x \in K$, $x \neq b$, gewählt, $z = \dfrac{x-b}{\|x-b\|}$. Dann ist für $0 < \varepsilon < \varepsilon_0$:
$y = b - \varepsilon \cdot z \in L$, $\|b-y\| < \varepsilon \Longrightarrow y \in K$, also

$$b - \varepsilon \cdot z = \sum p_j \cdot y_j,\ p_j \geq 0,\ \sum p_j = 1,\ y_j \in M,$$

sowie

$$\langle z, y-b\rangle = -\varepsilon = \sum p_j \cdot \langle z, y_j - b\rangle\,.$$

Daraus folgt, daß $\langle z, y_j - b\rangle \leq -\varepsilon$ für mindestens ein j, insbesondere

$$\inf_{y\in M} \langle z, y-b\rangle \leq -\varepsilon.$$

Wegen (V) ist $\|y-b\| \leq m$ für alle $y \in M$, also

$$\inf_{y\in M} \frac{\langle x-b, y-b\rangle}{\|x-b\|\cdot\|y-b\|} \leq -\frac{\varepsilon}{m} < 0,$$

somit $\alpha_b(M) \leq -\dfrac{\varepsilon}{m} < 0$.
b) Sei jetzt umgekehrt $\alpha := \alpha_b(M) < 0$. Wegen $b \notin \text{cl } M$ gibt es eine Zahl $\gamma > 0$, so daß $\|y-b\| \geq \gamma$ für alle $y \in M$. Es sei δ und ε so gewählt, daß

$$0 < \delta < \gamma \qquad \text{und} \qquad (\alpha+\varepsilon)\cdot m^2 + 2\delta m + \delta^2 < 0.$$

Es sei L die kleinste lineare Mannigfaltigkeit, die cl K enthält und $x \in L$ so gewählt, daß $\|x-b\| \leq \delta$.

Behauptung $x \in \text{cl } K$.
Um dieses zu beweisen, zeigen wir, daß $\alpha_x(\text{cl } K) < 0$ ist. Aus der bereits bewiesenen Aussage 1. des Satzes folgt dann $x \in \text{cl (conv (cl } K)) = \text{cl } K$.
Sei also $z \in \text{cl } K$, $z \neq b$, $z \neq x$, fest vorgegeben. Dann gibt es ein $y \in \text{cl } K$, so daß

$$\begin{aligned} & \langle z-b, y-b\rangle &<& \ (\alpha+\varepsilon)\cdot\|z-b\|\cdot\|y-b\| \leq (\alpha+\varepsilon)\cdot m^2, \\ \Longrightarrow \qquad & \langle z-x, y-x\rangle &=& \ \langle z-b, y-b\rangle + \langle b-x, y-b\rangle + \langle z-x, b-x\rangle \leq \\ & &\leq& \ (\alpha+\varepsilon)\cdot m^2 + 2\delta\cdot m + \delta^2 < 0. \end{aligned}$$

Weiterhin ist

$$\begin{aligned} \|y-x\| &\leq \|y-b\| + \|b-x\| \leq m+\delta, \\ \|z-x\| &\leq \|z-b\| + \|b-x\| \leq m+\delta. \end{aligned}$$

folglich

$$\frac{\langle z-x, y-x\rangle}{\|z-x\|\cdot\|y-x\|} \leq \frac{(\alpha+\varepsilon)\cdot m^2 + 2\delta\cdot m + \delta^2}{(m+\delta)^2} < 0,$$

also $\alpha_x(\text{cl } K) < 0$.

Zu 3. Die Aussage 3. ergibt sich aus den beiden Aussagen 1. und 2. □

Eine Ergänzung von Satz 6.1 ist das

Korollar 6.1 *Es gelte (V) und es sei $b \notin \text{cl } M$.*
Dann gilt: $\alpha_b(M) < 0 \implies \alpha_b(K) = -1$.

Beweis Wegen $\alpha_b(M) < 0$ ist $b \in \text{relint cl } K$. Es sei L die kleinste lineare Mannigfaltigkeit, die cl K enthält. Dann gibt es eine Kugel B um b, so daß $B \cap L$ ganz in K liegt. Ist $x \in \text{cl } K$, $x \neq b$ fest gewählt, dann gibt es ein $\lambda > 0$, so daß $b + \lambda \cdot (b - x)$ in $B \cap L \subseteq \text{cl } K$ liegt, also wird

$$\frac{\langle x - b, b + \lambda \cdot (b - x) - b \rangle}{\|x - b\| \cdot \|b + \lambda \cdot (b - x) - b\|} = -1,$$

das heißt, $\alpha_b(\text{cl } K) = -1$, damit auch $\alpha_b(K) = -1$ (Satz 6.1, Aussage 1). □

Ist $\mathcal{H}$ endlichdimensional, dann gilt die folgende symmetrische Version des Satzes 6.2:

Korollar 6.2 *Es sei $\mathcal{H} = \mathbb{R}^d$ endlichdimensional, es gelte (V) und es sei $b \notin \text{cl } M$. Dann ist*

1. $\alpha_b(M) > 0 \iff b \notin \text{cl } K$,
2. $\alpha_b(M) < 0 \iff b \in \text{relint } K$,
3. $\alpha_b(M) = 0 \iff b \in \text{relbd } K$.

Beweis Folgt direkt aus der Tatsache, daß im $\mathbb{R}^d$ gilt: relint cl K = relint K [103, Satz 3.2.13, S. 90]. □

Auf die Voraussetzung (V) in Satz 6.2 kann nicht völlig verzichtet werden:

Beispiel 6.2 *Es sei $\mathcal{H} = \mathbb{R}^2$, $b = \Theta_2$ und*

$$M = K = \left\{ \begin{pmatrix} 1 \\ \eta \end{pmatrix} \;\middle|\; \eta \in \mathbb{R} \right\}.$$

Dann ist für jedes $x = (1, \eta_1)^\top$ und $y = (1, \eta_2)^\top$:

$$\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} = \frac{1 + \eta_1 \cdot \eta_2}{\sqrt{1 + \eta_1^2} \cdot \sqrt{1 + \eta_2^2}},$$

und man folgert leicht, daß $\alpha_b(M) = 0$ ist. Offenbar ist aber $b \notin \text{cl } K$.

Zur Veranschaulichung rechnen wir für ein einfaches Erzeugendensystem M im $\mathbb{R}^d$ einmal $\alpha_b(M)$ aus:

Beispiel 6.3 *Es sei $b = \Theta_d$, e_k sei der k-te Einheitsvektor des $\mathbb{R}^d$ und*

$$M = \{\pm e_k \mid k = 1, \ldots, d\}.$$

Dann ist $K = \text{cl } K$ ein der Einheitsspäre einbeschriebener Würfel mit den Ecken $\pm e_k$ und der Kantenlänge $\sqrt{2}$.

Sei $x \in K$, also $x = \sum p_j \cdot e_j - \sum q_j \cdot e_j$, $\sum (p_j + q_j) = 1$, $p_j \geq 0$, $q_j \geq 0$. Dann ist

$$\frac{\langle x, \pm e_k \rangle}{\|x\|} = \frac{p_k - q_k}{\sqrt{\sum (p_j + q_j)^2}}.$$

Wie man sich leicht überzeugt, ist hier $\alpha = \alpha_b(M) = -1/\sqrt{d}$.

Die Öffnung α von M bezüglich b hat eine besondere Bedeutung für die Konvergenz von (VF). Insbesondere folgt für $\alpha < 0$ mit einer speziellen Auswahlmenge lineare Konvergenz. Dazu definieren wir für $0 < C < 1$ die Menge

$$A^{(C)}(x) = \left\{ y \in M \mid \Phi(x,y) \le C \cdot \alpha \cdot \sqrt{\phi(x) \cdot \phi(y)} \right\}.$$

Damit gilt der

Satz 6.3 *Sei $\alpha < 0$, $0 < C < 1$, und bei jedem Schritt des Verfahrens (VF) werde $y_r \in A^{(C)}(x_r)$ gewählt. Dann konvergiert die Folge der x_r gemäß*

$$\phi(x_r) \le \phi(x_1) \cdot \left(1 - C^2 \cdot \alpha^2\right)^{r-1}. \tag{6.1}$$

Beweis Aus (3.8) erhält man wegen $\phi(x_r) - 2 \cdot \Phi(x_r, y_r) \ge 0$:

$$\frac{\phi(x_{r+1})}{\phi(x_r)} \le 1 - \frac{\Phi(x_r, y_r)^2}{\phi(x_r) \cdot \phi(y_r)} \le 1 - C^2 \cdot \alpha^2,$$

woraus die Behauptung folgt. □

Zur Vereinfachung der Argumentation nehmen wir jetzt ohne Beschränkung der Allgemeinheit an, daß $b = \Theta_{\mathcal{H}}$ ist. Ist dies nicht der Fall, ersetzen wir die Menge M durch

$$M' = \{x \in \mathcal{H} \mid x = y - b,\ y \in M\}.$$

Dann ist $b \in \text{cl conv } M \quad \Longleftrightarrow \quad \Theta_{\mathcal{H}} \in \text{cl conv } M'$.

Der Fall $\alpha < 0$ mit der Auswahlmenge $A^{(C)}(x)$ hat noch andere interessante Konvergenzeigenschaften.

Satz 6.4 *Sei $\alpha < 0$, $b = \Theta_{\mathcal{H}} \notin \text{cl } M$, es werde nach $A^{(C)}(x_r)$ ausgewählt. Dann gilt für die nach (3.2) definierten Größen μ_r:*

$$\prod_{r=1}^{\infty} \mu_r > 0 \qquad \textit{und} \qquad \sum_{r=1}^{\infty} (1 - \mu_r) < \infty.$$

Beweis Es ist für alle $r = 1, 2, \ldots$

$$\phi(x_{r+1}) = \mu_r \cdot \Phi(x_r, y_r) + (1 - \mu_r) \cdot \phi(x_r),$$

wie man aus (3.2) und (3.3) herleitet. Wegen $\phi(x) = \|x\|^2$ wird

$$\sum_{r=1}^{\infty} (1 - \mu_r) = \sum_{r=1}^{\infty} \frac{\|x_{r+1}\|^2}{\|y_r\|^2} - \sum_{r=1}^{\infty} \mu_r \cdot \frac{\langle x_r, y_r \rangle}{\|y_r\|^2},$$

wobei

$$0 \le -\mu_r \cdot \langle x_r, y_r \rangle \le -\langle x_r, y_r \rangle \le \|x_r\| \cdot \|y_r\|.$$

Wegen $b \notin \text{cl } M$ gibt es ein $\gamma > 0$, so daß $\|y\| \ge \gamma$ für alle $y \in M$, und mit (6.1) wird dann

$$\sum_{r=1}^{\infty} (1 - \mu_r) \le \frac{\|x_1\|^2}{\gamma^2} \cdot \sum_{r=1}^{\infty} \left(1 - C^2 \cdot \alpha^2\right)^r + \frac{\|x_1\|}{\gamma} \cdot \sum_{r=1}^{\infty} \left(1 - C^2 \cdot \alpha^2\right)^{\frac{r+1}{2}}.$$

Beide Reihen auf der rechten Seite sind geometrische Reihen, also konvergent. Nach einem Satz aus der Theorie der unendlichen Produkte ist aber die Konvergenz der unendlichen Reihe $\sum(1-\mu_r)$ äquivalent zur Konvergenz des unendlichen Proukts $\prod \mu_r$ gegen einen endlichen, von Null verschiedenen Grenzwert [67, Seite 227, Satz 4], womit alles bewiesen ist. □

Aus Satz 6.4 folgern wir

Satz 6.5 *Es sei $\alpha < 0$, $\Theta_{\mathcal{H}} \notin \mathrm{cl}\, M$ und es werde aus $A^{(C)}(x_r)$ ausgewählt. Es sei weiterhin y ein festes Element aus M. Gibt es dann für irgendein r eine Darstellung*

$$x_r = p_r \cdot y + (1 + p_r) \cdot z_r \quad z_r \in K, \quad 0 < p_r \leq 1,$$

dann gibt es für jedes r eine solche Darstellung, so daß die p_r durch eine feste positive Zahl gleichmäßig von unten beschränkt sind.

Beweis Es ist

$$x_{r+1} = \mu_r \cdot p_r \cdot y + \mu_r \cdot (1 - p_r) \cdot z_r + (1 - \mu_r) \cdot y_r = p_{r+1} \cdot y + (1 - p_{r+1}) \cdot z_{r+1},$$

wobei $z_{r+1} \in K$ und $p_{r+1} = \mu_r \cdot p_r$. Aus der Beschränktheit des unendlichen Produkts in Satz 6.4 folgt die Beschränktheit der p_r. □

Satz 6.6 *Es sei $\alpha < 0$, $\Theta_{\mathcal{H}} \notin \mathrm{cl}\, M$ und es werde aus $A^{(C)}(x_r)$ ausgewählt. Weiterhin gelte (V). Dann liefert das Verfahren (VF) eine Lösung von (A) der Form*

$$\Theta_{\mathcal{H}} = \sum_{j=0}^{\infty} p_j \cdot z_j, \quad p_j > 0, \quad \sum_{j=0}^{\infty} p_j = 1, \quad z_j \in M. \tag{6.2}$$

Beweis Es ist $x_r = \sum_{j=0}^{r-1} p_j^{(r)} \cdot y_j$, wobei $y_0 = x_1 \in K$, $y_j \in M$ für $j \geq 1$, $p_j^{(r)} \geq 0$ und $\sum p_j^{(r)} = 1$. Aus (3.3) erhält man die Rekursionsbeziehung für $p_j^{(r)}$:

$$p_j^{(r+1)} = \begin{cases} \mu_r \cdot p_j^{(r)} & \text{für } j < r \\ 1 - \mu_r & \text{für } j = r. \end{cases}$$

Für festes j existiert $\lim_{r\to\infty} p_j^{(r)} := p_j$ und es ist $p_j > 0$ (Satz 6.5) für alle j. Außerdem ist für jedes $j \geq 1$ und für $k > k$:

$$0 < p_j < p_j^k < 1 - \mu_j.$$

Sei jetzt $\varepsilon > 0$ beliebig vorgegeben, r und k fest gewählt, $k < r$. Dann ist

$$\begin{aligned} \left\| x_r - \sum_{j=0}^{\infty} p_j \cdot y_j \right\| &= \left\| \sum_{j=0}^{k} p_j^{(r)} \cdot y_j - \sum_{j=0}^{\infty} p_j \cdot y_j \right\| \leq \\ &\leq \sum_{j=0}^{k} \left(p_j^{(r)} - p_j \right) \cdot \|y_j\| + \sum_{j=k+1}^{r-1} \left(p_j^{(r)} - p_j \right) \cdot \|y_j\| + \sum_{j=r}^{\infty} p_j \cdot \|y_j\| \leq \\ &\leq m \cdot \left[\sum_{j=0}^{k} \left(p_j^{(r)} - p_j \right) + \sum_{j=k+1}^{\infty} (1 - \mu_j) \right]. \end{aligned}$$

Die zweite Summe der letzten Zeile ist Reihenrest einer konvergenten Reihe (Satz 6.4), wird also kleiner als $\frac{\varepsilon}{2 \cdot m}$, wenn k hinreichend groß ist. Jeder der Summanden der ersten Summe ist für hinreichend großes r $(> k)$ kleiner als $\frac{\varepsilon}{2 \cdot m \cdot (k+1)}$, damit wird der gesamte Ausdruck kleiner als ε. Wegen $\lim x_r = \Theta_{\mathcal{H}}$ ist auch $\sum_{j=0}^{\infty} p_j \cdot y_j = \Theta_{\mathcal{H}}$.

Ebenso zeigt man, daß $\sum_{j=0}^{\infty} p_j = 1$. □

Satz 6.6 kann als unendlichdimensionale Verallgemeinerung des Satzes von Carathéodory (Satz 2.2) aufgefaßt werden. Er ist für $\alpha = 0$ sicher nicht richtig, wie Beispiel 3.2 zeigt.

Noch eine Anmerkung für den späteren Gebrauch: Offenbar gelten die Sätze 6.3 und 6.6 auch dann, wenn man die Bedingung $\alpha < 0$ ersetzt durch die schwächere Bedingung:

Bei jedem Iterationsschritt läßt sich ein $y_r \in M$ finden, so daß

$$\langle x_r, y_r \rangle \leq C \cdot \|x_r\| \cdot \|y_r\| < 0, \tag{6.3}$$

wo C eine feste Zahl ist.

Es sei noch darauf hingewiesen, daß in den Sätzen 6.3 bis 6.5 kein Gebrauch von der Voraussetzung (V) gemacht wurde.

Kapitel 7

Konvexe Hüllen in endlichdimensionalen Räumen

Es sei jetzt $\mathcal{H} = \mathbb{R}^d$ und wieder $b = \Theta_d$.

Hier können wir natürlich mehr aussagen als im Hilbertraum. Insbesondere ist hier der Satz von Carathéodory (Satz 2.2) sehr nützlich.

Bei der Lösung des Aufgabenpaares (A), (D) mit dem Verfahren (VF) sei x_r die r-te Iterierte. Wegen des Satzes von Carathéodory können wir x_r als Konvexkombination von höchstens $d+1$ Elementen aus M darstellen. Numerisch ist die Reduktion einer Darstellung eines Elements aus conv M nicht schwierig, man kann ein einfaches Eliminationsverfahren angeben [106, Kap. V, §5].

Es sei jetzt M abgeschlossen und beschränkt, also kompakt. Ist dann $\Theta_d \in \text{cl conv } M$, also (A) lösbar, dann gibt es eine Folge $\{x_r\}$ mit $x_r \in K = \text{conv } M$ und $\lim x_r = \Theta_d$. Jedes x_r hat nach dem Satz von Carathéodory eine Darstellung

$$x_r = \sum_{j=1}^{d+1} p_j^{(r)} \cdot y_j^{(r)} \quad \text{mit} \quad y_j^{(r)} \in M, \quad p_j^{(r)} \geq 0, \quad \sum_{j=1}^{d+1} p_j^{(r)} = 1.$$

Die Folge der $y_1^{(r)}$ besitzt eine gegen y_1 konvergente Teilfolge $\left\{y_1^{(r_k)}\right\}$ wegen der Kompaktheit von M. Die Folge der $p_1^{(r_k)}$ wiederum besitzt eine gegen p_1 konvergente Teilfolge und die entsprechende Teilfolge der $y_2^{(r)}$ besitzt eine gegen y_2 konvergente Teilfolge usw. Es gibt also p_j mit $p_j \geq 0$, $\sum p_j = 1$ und $y_j \in M$, so daß

$$\sum_{j=1}^{d+1} p_j \cdot y_j = \Theta_d.$$

Wir haben also den bekannten

Satz 7.1 *Ist M kompakt, dann ist $K = \text{conv } M$ abgeschlossen. Ist (A) lösbar, dann kann man die Lösung darstellen in der Form*

$$\Theta_d = \sum_{j=1}^{d+1} p_j \cdot y_j, \quad p_j \geq 0, \quad \sum_{j=1}^{d+1} p_j = 1, \quad y_j \in M. \tag{7.1}$$

Beweis der Abgeschlossenheit von conv M: Ist $z \in \text{cl conv } M$, dann setzen wir wieder $M' = \{x \in \mathbb{R}^d \mid x = y - z, y \in M\}$. Wie gezeigt, ist dann $\Theta_d \in \text{conv } M'$, und dies ist äquivalent zu $z \in \text{conv } M$. □

Es gilt der Alternativsatz

Satz 7.2 *Sei M kompakt. Genau einer der beiden folgenden Fälle tritt auf:*

1. $\Theta_d \in \text{conv } M$,
2. *(D) ist lösbar.*

Beweis Wegen Satz 3.2 bleibt nur noch zu zeigen, daß nicht beide Aussagen gleichzeitig wahr sein können. Ist $\Theta_d \in \text{conv } M$, gilt (7.1). Wäre auch noch (D) lösbar durch ein $\widehat{x} \in \mathbb{R}^d$, würde gelten:

$$0 = \left\langle \widehat{x}, \sum p_j \cdot y_j \right\rangle = \sum p_j \cdot \langle \widehat{x}, y_j \rangle > 0,$$

ein Widerspruch. □

Ist $\Theta_d \in \text{conv } M$, kann trotzdem die folgende Aufgabe lösbar sein:

Gesucht ist ein $\widehat{x} \in \mathbb{R}^d$ (beziehungsweise $\widehat{x} \in \text{conv } M$), $\widehat{x} \neq \Theta_d$, so daß

$$\langle \widehat{x}, y \rangle \geq 0 \qquad \text{für alle } y \in M.$$

In diesem Falle wäre die Öffnung von M bezüglich Θ_d Null.

Definition 7.1 *Der Vektor $\widehat{x} \in \mathbb{R}^d$ heißt* schwache Lösung *von (D) (mit $b = \Theta_d$), wenn $\widehat{x} \neq \Theta_d$ und*

$$\langle \widehat{x}, y \rangle \geq 0 \qquad \textit{für alle } y \in M.$$

Es gilt der einfache

Satz 7.3 *Ist $\sum_{j \in J} p_j \cdot y_j = \Theta_d$, $p_j > 0$, $\sum_{j \in J} p_j = 1$, $y_j \in M$, J eine endliche Indexmenge, dann gilt die Implikation:*

$$x \textit{ schwache Lösung von (D)} \quad \Longrightarrow \quad \langle x, y_j \rangle = 0 \;\; \textit{für alle } j \in J.$$

Beweis Trivial. □

Die schwache Lösung von (D) liegt also in dem linearen Teilraum

$$U = \left\{ x \in \mathbb{R}^d \mid \langle x, y_j \rangle = 0 \quad \text{für alle } j \in J \right\}.$$

Haben wir eine Lösung von (A) – etwa mit (VF) – gefunden, dann können wir zur Bestimmung der schwachen Lösung von (D) die Dimension des Problems auf mindestens $d - 1$ reduzieren. Problematisch ist hier jedoch wegen $\alpha = \alpha_b(M) = 0$ (vgl. Definition 6.1) die langsame Konvergenz des Verfahrens.

Wir setzen nun voraus, daß die folgende naheliegende Forderung auch noch erfüllt ist:

M enthalte $d+1$ affin unabhängige Vektoren.
Bezeichnungsweise: dim $M = d$.

Ist diese Forderung nicht erfüllt, dann können wir die Dimension des Problems reduzieren.

Ist $\alpha = 0$, M kompakt und dim $M = d$, dann löst das Verfahren (VF) die beiden Aufgaben (A) und (D) (schwach) gleichzeitig. Um dieses zu zeigen, formulieren wir zunächst den

Satz 7.4 *Sei $\alpha = 0$, $\Theta_d \notin \operatorname{cl} M$ und M kompakt. (VF) werde mit der Auswahlmenge $A_0(x)$ (3.6) angewandt auf einen Starvektor*

$$x_0 = \sum_{j=0}^{d} p_j^{(0)} \cdot z_j \qquad \textit{mit} \quad p_j^{(0)} > 0, \quad \sum p_j^{(0)} = 1, \quad z_j \in M,$$

wobei die z_j affin unabhängig seien.
Dann ist

$$\lim_{r\to\infty} \frac{\|x_{r+1}\|}{\|x_r\|} = 1.$$

Beweis Es ist $0 < \|x_{r+1}\|/\|x_r\| < 1$. Wäre $\|x_{r+1}\|/\|x_r\| \le \delta < 1$ für alle r, dann würde wegen (3.8) gelten:

$$\frac{\|x_{r+1}\|^2}{\|x_r\|^2} = \frac{1 - \frac{\langle x_r, y_r\rangle^2}{\|x_r\|^2 \cdot \|y_r\|^2}}{1 + \frac{\|x_r\|^2}{\|y_r\|^2} - 2 \cdot \frac{\langle x_r, y_r\rangle}{\|y_r\|^2}} \le \delta^2$$

somit

$$1 - \frac{\langle x_r, y_r\rangle^2}{\|x_r\|^2 \cdot \|y_r\|^2} \le \delta^2 \cdot \left(1 + \frac{\|x_r\|^2}{\|y_r\|^2} - \frac{2}{\|y_r\|^2} \cdot \langle x_r, y_r\rangle\right).$$

Der Ausdruck auf der rechten Seite konvergiert wegen $b\ (= \Theta_d) \notin \operatorname{cl} M$ für $r \to \infty$ gegen δ^2 (weil M kompakt ist, ist (V) sutomatisch erfüllt), also gibt es zu jedem $\varepsilon > 0$ eine Zahl r_0, so daß

$$1 - \frac{\langle x_r, y_r\rangle^2}{\|x_r\|^2 \cdot \|y_r\|^2} \le \delta^2 + \varepsilon \qquad \text{für alle } r \ge r_0.$$

Wählt man ε so klein, daß $\delta^2 + \varepsilon < 1$, dann folgt, daß für alle hinreichend großen r gelten muß

$$\langle x_r, y_r\rangle \le -\sqrt{1 - \varepsilon - \delta^2} \cdot \|x_r\| \cdot \|y_r\| < 0.$$

Damit ist die Bedingung (6.3) erfüllt, es gelten also insbesondere Satz 6.5 und 6.6. Wurde der Startvektor wie oben gewählt, dann bleiben die Koeffizienten z_j, $j = 1, \ldots, d$, von unten beschränkt. Nach Satz 6.6 gibt es eine Darstellung

$$\Theta_d = \sum_{j=0}^{d} p_j \cdot z_j + \sum_{j=d+1}^{\infty} p_j \cdot y_j, \quad p_j \ge 0, \quad \sum_{j=0}^{\infty} p_j = 1,$$

wobei $p_j > 0$ für $j = 0, \ldots, d$. Mit $\sigma = \sum_{j=0}^{d} p_j$ ist

$$\Theta_d = \sigma \cdot \sum_{j=0}^{d} \frac{p_j}{\sigma} \cdot z_j + (1-\sigma) \cdot \sum_{j=d+1}^{\infty} \frac{p_j}{1-\sigma} \cdot y_j = \sigma \cdot \sum_{j=0}^{d} \mu_j \cdot z_j + (1-\sigma) \cdot y,$$

wo $0 < \sigma \leq 1$, $\mu_j > 0$ für $j = 0, \ldots, d$, $\sum_{j=0}^{d} \mu_j = 1$ und $y \in \operatorname{conv} M$. Damit ist Satz 2.3 anwendbar, es ist $\Theta_d \in \operatorname{int} \operatorname{conv} \{z_0, \ldots, z_d, y\} \subseteq \operatorname{int} \operatorname{conv} M$.

Wegen Satz 6.1 und 6.2 wäre dann $\alpha < 0$, ein Widerspruch. □

Satz 7.4 beinhaltet eine ähnliche negative Konvergenzaussage wie das Beispiel 6.1, nämlich daß im Falle $\alpha = 0$ mit Sicherheit die Konvergenz des Verfahrens nicht mehr linear sein kann, daß also die Konvergenzschranke (3.7) wenn nicht streng, so doch nicht sehr weit davon entfernt sein wird, streng zu sein. Dies ist auch der Grund dafür, daß es sich nicht empfiehlt, Optimierungsaufgaben mit (VF) zu lösen. Führt man nämlich eine Optimierungsaufgabe auf das Aufgabenpaar (A), (D) zurück, dann ist die Öffnung dieses Aufgabenpaares Null. Darauf kann man wohl auch die unbefriedigenden Ergebnisse zurückführen, die Hoffman u.a. [61] bei der Lösung von Optimierungsaufgaben mit dem Agmonschen Verfahren erzielten.

Nun kommen wir zur schwachen Lösbarkeit von (D) zurück. Wegen der vorausgesetzten Kompaktheit von M können wir aus der folgenden Auswahlmenge auswählen:

$$A_{\min}(x) = \left\{ y \in M \mid \langle x, y \rangle = \min_{z \in M} \langle x, z \rangle \right\}.$$

Dann gilt der

Satz 7.5 *Unter den Voraussetzungen des Satzes 7.4 werde (VF) mit der Auswahlmenge $A_{\min}(x)$ durchgeführt. Dann ist jeder Häufungspunkt der Folge $\left\{ \dfrac{x_r}{\|x_r\|} \right\}$ eine schwache Lösung von (D).*

Beweis Sei $\alpha_r = \langle x_r, y_r \rangle$. Es ist $\alpha_r \leq 0$ für alle r und wegen Satz 7.4 ist $\lim_{r \to \infty} \alpha_r = 0$.

Sei $\left\{ \dfrac{x_{r_k}}{\|x_{r_k}\|} \right\}$ eine konvergente Teilfolge mit Grenzwert $\widehat{x}$, $\|\widehat{x}\| = 1$. Für jedes r ist $\langle x_r, y \rangle \geq \alpha_r$ für alle $y \in M$ nach Definition von $A_{\min}(x)$, folglich $\langle \widehat{x}, y \rangle \geq 0$ für alle $y \in M$. □

Kapitel 8

Kegel im $\mathbb{R}^d$

Es sei $M \subseteq \mathbb{R}^d$ eine kompakte Menge, $b \in \mathbb{R}^d$.

Wir untersuchen jetzt, ob $b \in \operatorname{cone} M$, das heißt, gesucht sind p_j mit

$$\sum_{j=1}^{k} p_j \cdot y_j = b, \quad p_j \geq 0, \quad y_j \in M. \tag{AC}$$

Man kann versuchen, die Aufgabe

$$\begin{aligned} &\sum_{j=1}^{k} p_j \cdot y_j - p \cdot b = \Theta_d, \quad p_j \geq 0, \quad p \geq 0, \\ &\sum_{j=1}^{k} p_j + p = 1, \qquad y_j \in M \end{aligned} \tag{8.1}$$

mit (VF) zu lösen und hoffen, daß dabei $p > 0$ bleibt.

Unter gewissen Bedingungen kann man $p > 0$ garantieren, falls (AC) überhaupt lösbar ist:

Satz 8.1 *Es gelte für jede Lösung von (AC), daß*

$$\sum_{j=1}^{k} p_j \leq m. \tag{8.2}$$

Genau dann ist $p > 0$ für jede Lösung von (8.1), wenn (AC) lösbar ist.

Beweis Ist (AC) lösbar, dann auch (8.1) mit $p > 0$.

Sei jetzt (8.1) lösbar mit $p = 0$, das heißt, es gibt $q_j \geq 0$, $\sum q_j = 1$, $y_j \in M$, so daß $\sum_{j=1}^{k} q_j \cdot y_j = \Theta_d$.

Wäre auch (AC) lösbar, würde gelten

$$\sum_{j=1}^{k} p_j \cdot y_j + \gamma \cdot \sum_{j=1}^{k} q_j \cdot y_j = b$$

und

$$\sum_{j=1}^{k} p_j + \gamma \cdot \sum_{j=1}^{k} q_j > m,$$

falls γ hinreichend groß ist, im Widerspruch zu (8.2). □

Kann man also eine Abschätzung (8.2) finden, dann kann man (VF) unbesorgt auf die Aufgabe (8.1) anwenden. Man kann (8.2) erzwingen, indem man zu (AC) die Gleichung

$$\sum_{j=1}^{k} p_j + p = m$$

mit $p \geq 0$ hinzunimmt. In vielen Fällen ist es aus praktischen Gründen leicht möglich, den Wertebereich der p_j einzugrenzen. Das folgende Beispiel zeigt jedoch, daß der Ansatz (8.1) durchaus nicht immer zum Ziele führt:

Beispiel 8.1 *Es sei* $M = \{a_1, \ldots, a_5\}$ *mit*

$$\begin{aligned} a_1 &= (0,1,0)^\top, \\ a_2 &= (0,-1,0)^\top, \\ a_3 &= \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)^\top \\ a_4 &= \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)^\top \\ a_5 &= \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^\top \\ b &= (1,0,0)^\top. \end{aligned}$$

Die Aufgabe (AC) ist lösbar durch $b = \frac{1}{\sqrt{2}} \cdot (a_3 + a_4)$.

Wir wählen als Startvektor

$$x_0 = \frac{1}{5+\sqrt{2}} \cdot \left(a_1 + a_2 + a_3 + a_4 + a_4 - \sqrt{2} \cdot b\right) = \frac{1}{5+\sqrt{2}} \cdot a_5.$$

Für jedes r *ist, wie man leicht durch Induktion verifiziert,* $y_r \in \{a_1, a_2, a_5\}$ *und damit auch* $x_r \in \operatorname{conv}(\{a_1, a_2, a_5\})$*, folglich muß der Faktor von* b *gegen Null konvergieren.*

Wir beweisen nun eine Reihe von Sätzen, die zu einem Verfahren führen, das eine systematische iterative Behandlung von (AC) ohne weitere Zusatzbedingungen erlaubt.

Satz 8.2 *Die Aufgabe (AC) ist genau dann lösbar, wenn es eine Zahl* $\gamma > 0$ *gibt, so daß die Aufgabe*

$$\sum q_j \cdot (y_j - \gamma \cdot b) = \Theta_d, \quad q_j \geq 0, \quad \sum q_j = 1 \tag{8.3}$$

lösbar ist.

Beweis Trivial. □

Die Menge M ist kompakt, also beschränkt. Es sei $\|y\| \le m$ für alle $y \in M$. Dann gilt der

Satz 8.3 *Ist* $\gamma > \dfrac{m}{\|b\|}$, *dann ist (8.3) nicht lösbar.*

Beweis Es wäre sonst

$$m < \gamma \cdot \|b\| = \left\| \sum q_j \cdot y_j \right\| \le \sum q_j \cdot \|y_j\| \le m \cdot \sum q_j = m,$$

ein Widerspruch. □

Satz 8.4 *Es gebe einen Vektor* $x \in \mathbb{R}^d$ *und eine Zahl* γ_0, *so daß*

$$\langle x, y - \gamma_0 \cdot b \rangle > 0 \qquad \textit{für alle } y \in M. \tag{8.4}$$

Dann gilt

1. $\langle x, b \rangle < 0 \quad \Longrightarrow \quad$ *(8.3) ist nicht lösbar für* $\gamma \ge \gamma_0$,
2. $\langle x, b \rangle > 0 \quad \Longrightarrow \quad$ *(8.3) ist nicht lösbar für* $\gamma \le \gamma_0$,
3. $\langle x, b \rangle = 0 \quad \Longrightarrow \quad$ *(8.3) ist überhaupt nicht lösbar.*

Beweis Es sei für festes $y \in M$

$$f(\gamma) = \langle x, y - \gamma \cdot b \rangle .$$

Dann ist $f(\gamma_0) > 0$ und $\dfrac{df}{d\gamma} = -\langle x, b \rangle$, woraus die Behauptungen 1. bis 3. folgen. □

Gestützt auf die Sätze 8.2 bis 8.4 bietet sich das folgende Verfahren an

0. Wähle $\gamma_0 > 0$ (etwa nach Satz 8.3). Setze $n := 0$.
1. Wende (VF) auf die Aufgabe (8.3) an. Entweder konvergiert es gegen eine Lösung von (8.3), also auch von (AC), oder es liefert nach endlich vielen Schritten einen Vektor x_r mit

   $$\langle x_r - \gamma_n \cdot b, y - \gamma_n \cdot b \rangle > 0 \qquad \text{für alle } y \in M.$$

   Setze $x = x_r - \gamma_n \cdot b$.

   1.1. $\langle x, b \rangle > 0$. Die Aufgabe (8.3) ist nicht lösbar für $\gamma \le \gamma_n$, insbesondere für $\gamma = 0$. Wir können daher unbesorgt (8.1) mit (VF) behandeln.

   1.2. $\langle x, b \rangle = 0$. (8.3) und damit (AC) ist nicht lösbar.

   1.3. $\langle x, b \rangle < 0$. (8.3) ist nicht lösbar für $\gamma \ge \gamma_n$.
   Gehe zu 2.

2. Setzen wir

$$g(\gamma) = \|yx_r - \gamma \cdot b\|^2,$$

dann ist $g'(\gamma_n) = -2 \cdot \langle x, b\rangle > 0$, g wird durch Verkleinerung von γ kleiner. Das Minimum von g wird angenommen für

$$\gamma' = \frac{\langle x_r, b\rangle}{\|b\|^2}.$$

2.1. $\gamma' > 0$. Setze $\gamma_{n+1} := \gamma'$ und gehe zu 3.

2.2. $\gamma' \leq 0$. Dann setzen wir etwa $\gamma_{n+1} = \dfrac{\gamma_n}{2}$ und gehen zu 3.

3. Setze $n := n + 1$ und gehe zu 1.

Die Konvergenz des Verfahrens ergibt sich aus der Tatsache, daß im Schritt 1. $\|x_r - \gamma_n \cdot b\|$ gemäß (3.7) konvergiert und in Schritt 2. $\|x_r - \gamma_n \cdot b\|$ auf jeden Fall verkleinert wird.

Es sei jetzt α die Öffnung der Menge $M \cup \{-b\}$ bezüglich Θ_d. Mit Satz 6.5 kommt man für $\alpha < 0$ bei direkter Iteration mit (8.1) zu einer interessanten Endlichkeitsaussage. Dazu nehmen wir an, M enthalte d linear unabhängige Vektoren. Es ist keine Einschränkung der Allgemeinheit, anzunehmen, daß diese linear unabhängigen Vektoren die Einheitsvektoren $e_1, \ldots, e_d$ sind. Es gilt der

Satz 8.5 *Sei $\alpha < 0$. $\Theta_d \notin \operatorname{cl} M$, M enthalte d Einheitsvektoren, es werde nach $A^{(C)}(x)$ ausgewählt.*

Es sei der Startvektor

$$\begin{aligned} x_0 \quad = \quad & \sum p_j^{(0)} \cdot e_j - p^{(0)} \cdot b + \sum q_j^{(0)} \cdot y_j \quad p_j^{(0)} > 0, \quad p^{(0)} > 0, \\ & q_j^{(0)} > 0, \quad \sum p_j^{(0)} + \sum q_j^{(0)} = 1, \quad y_j \in M. \end{aligned}$$

Dann ist nach endlich vielen Schritten des Verfahrens (VF)

$$p_j(r) \geq x_j^{(r)}, \qquad j = 1, \ldots, d, \tag{8.5}$$

wobei $x_j^{(r)}$ die j-te Komponente von x_r ist. (AC) ist also lösbar durch

$$\frac{1}{p^{(r)}} \cdot \left[\sum \left(p_j^{(r)} - x_j^{(r)}\right) \cdot e_j + \sum q_j^{(r)} \cdot y_j\right] = b.$$

Beweis Nach Satz 6.5 sind die $p_j^{(r)}$ für alle r durch eine feste positive Zahl von unten beschränkt. Die x_r konvergieren gegen Θ_d, also muß nach endlich vielen Schritten (8.5) gelten, woraus die Behauptung folgt. □

Das Verfahren wird also nur im Falle $\alpha = 0$ nicht nach endlich vielen Schritten abbrechen.

Kapitel 9

Anwendung auf abgeschlossene konvexe Mengen

Als Anwendung der Überlegungen der Abschnitte 7 und 8 untersuchen wir jetzt die folgende Aufgabenstellung

> Sei $P \subseteq \mathbb{R}^d$ konvex und abgeschlossen. Gesucht ist ein Element $x \in P$.

Diese Aufgabenstellung ist wichtig im Zusammenhang mit konvexen Optimierungsproblemen [23]. Die Anwendung des Verfahrens (VF) stützt sich auf den

Satz 9.1 *Es sei $P \subseteq \mathbb{R}^d$ konvex und abgeschlossen. Dann ist P darstellbar in der Form*

$$P = \left\{ x \in \mathbb{R}^d \mid \langle a_t, x \rangle \geq b_t \quad t \in T \right\} \tag{9.1}$$

wo T eine Indexmenge ist.

Der Beweis dieses Satzes beruht auf dem Trennungssatz, er ist etwa bei Stoer und Witzgall [103, Chap. 3.3] zu finden.

Mit abgeschlossenen Mengen, die in der Form (9.1) gegeben sind (im allgemeinen mit unendlichem T), beschäftigt sich die „semi-infinite programming theory“ von Charnes, Cooper und Kortanek [18].

Wir setzen

$$M = \left\{ \left(a_t^\top, b_t\right)^\top \in \mathbb{R}^{d+1} \mid t \in T \right\}$$

und zeigen, daß es keine Beschränkung der Allgemeinheit ist, vorauszusetzen, daß M eine kompakte Teilmenge der Einheitssphäre $\left\{ x \in \mathbb{R}^{d+1} \mid \|x\| = 1 \right\}$ des $\mathbb{R}^{d+1}$ ist. Dazu nehmen wir an, daß $\Theta_{d+1} \notin M$ ist. Andernfalls läge nämlich eine Ungleichung der Form $\langle \Theta_d, x \rangle \geq 0$ vor. Eine solche Ungleichung ist redundant und kann weggelassen werden.

Es gilt der

Satz 9.2 *Für jedes t sei $\gamma_t > 0$. Dann ist*

$$P = \left\{ x \in \mathbb{R}^d \mid \langle \gamma_t \cdot a_t, x \rangle \geq \gamma_t \cdot b_t, \quad t \in T \right\}.$$

Beweis Trivial. □

Setzen wir speziell $\gamma_t = \dfrac{1}{\sqrt{\|a_t\|^2 + b_t^2}}$, falls $\Theta_{d+1} \notin M$, dann liegen alle Elemente $\gamma_t \cdot \left(a_t^\top, b_t\right)^\top$ auf der $d+1$-dimensionalen Einheitssphäre. Satz 9.2 besagt, daß es keine Beschränkung der Allgemeinheit ist, von vornherein zu fordern, daß M eine Teilmenge dieser Einheitssphäre ist.

Weiterhin gilt der

Satz 9.3 *Es sei $\left(a^\top, b\right)^\top \in \operatorname{cl} M$. Dann ist $\langle a, x \rangle \geq b$ für alle $x \in P$.*

Beweis Trivial. □

Projizieren wir also M auf die Einheitskugel des $\mathbb{R}^{d+1}$, dann ändert sich P nicht. Bilden wir die abgeschlossene Hülle der Projektion, dann kommen höchstens redundante Ungleichungen hinzu. Als Folgerung aus den Sätzen 9.2 und 9.3 ergibt es sich, daß es keine Einschränkung der Allgemeinheit ist, zu fordern, daß M eine kompakte Teilmenge der $d+1$-dimensionalen Einheitssphäre ist.

Um einen Vektor $x \in P$ zu finden, wenden wir (VF) auf die folgende Aufgabe an:

$$\begin{array}{rcl} \sum p_j \cdot a_{t_j} & = & \Theta_d \\ -\sum p_j \cdot b_{t_j} + p & = & 0. \end{array} \qquad p, p_j \geq 0 \quad \sum p_j + p = 1. \tag{9.2}$$

Ist diese Aufgabe nicht lösbar, dann bricht (VF) nach endlich vielen Schritten mit einer Lösung $x \in P$ ab. Das ist insbesondere dann der Fall, wenn $\operatorname{int} P \neq \emptyset$ ist. Ist (9.2) lösbar mit $p > 0$, dann ist $P \neq \emptyset$ wegen Satz 7.3. Man kann die Lösbarkeit von (9.2) mit dem Verfahren des Abschnitts 8 systematisch untersuchen. (9.2) mit $p > 0$ ist äquivalent zu der Aufgabe

$$\begin{array}{rcl} \sum p_j \cdot a_{t_j} & = & \Theta_d \\ \sum p_j \cdot b_{t_j} & = & 1. \end{array} \qquad p_j \geq 0,$$

und diese Aufgabe hat die Form (AC).

Ist schließlich (9.2) nur mit $p = 0$ lösbar, dann kann man mit Satz 7.5 ein $x \in P$ ermitteln.

Der letztgenannte Fall tritt nicht auf, wenn man fordert, daß $\operatorname{int} P \neq \emptyset$ ist. Diese Forderung spielt in der konvexen Optimierung auch aus anderen Gründen eine Rolle [23, §7, Voraussetzung (V)].

Wir betrachten das folgende Beispiel:

Beispiel 9.1 *Vgl. [23, §10.2].*

$$\begin{array}{rcl} -\exp(x_1) + x_2 & \geq & 0, \\ -\exp(x_2) + x_3 & \geq & 0, \\ x_1 & \geq & 0, \\ x_2 & \geq & 0, \\ 10 \geq x_3 & \geq & 0. \end{array}$$

Das dazugehörige lineare Ungleichungssystem ist

$$\begin{array}{rrrcll} -\exp(t_1) \cdot x_1 + & x_2 & & \geq & \exp(t_1) \cdot (1 - t_1) & \text{(a)} \\ & -\exp(t_2) \cdot x_2 + x_3 & & \geq & \exp(t_2) \cdot (1 - t_2) & \text{(b)} \\ & & -x_3 & \geq & -10 & \text{(c)} \\ x_1 & & & \geq & 0 & \text{(d)} \\ & x_2 & & \geq & 0 & \text{(e)} \\ & & x_3 & \geq & 0. & \text{(f)} \end{array}$$

Es ist $T = \{(t_1, t_2)^\top\} \in \mathbb{R}^2$. Zur Durchführung der Rechnung wurde M wie beschrieben auf die Einheitssphäre projiziert und das Verfahren mit der Auswahlregel $A_{\min}(x)$ auf die durch $t_1 = 0$ in Ungleichung (a) definierte Startlösung angesetzt. Nach 18 Iterationsschritten auf der Rechenanlage CDC 6400 der RWTH Aachen war eine Lösung gefunden worden:

$$\begin{array}{rcl} x_1 & = & 0.195\,633, \\ x_2 & = & 1.290\,415, \\ x_3 & = & 6.058\,477. \end{array}$$

Die Bestimmung eines Punktes $x \in P$ wurde also auf die Lösung einer endlichen Anzahl von Minimierungsproblemen (ohne Restriktionen) zurückgeführt. Es ist natürlich zu überlegen, ob man sich auf Kosten der Anzahl der Iterationen mit der Auswahlmenge $A_0(x)$ zufriedengibt. Im vorliegenden Falle, wo P dargestellt ist in der Form

$$P = \left\{ x \in \mathbb{R}^d \mid g_i(x) \leq 0, \quad i \in I \right\},$$

wobei die $g_i(x)$ konvexe Funktionen sind, kann man sehr einfach ein Element aus $A_0(x)$ angeben, indem man im Punkte x eine Schnittebene konstruiert [23, §10.1].

Kapitel 10

Endlich viele lineare Ungleichungen

Es sei zusätzlich zu den Voraussetzungen der Abschnitte 7 und 8 M eine endliche Menge. Dann gelten natürlich alle in diesen Abschnitten bewiesenen Sätze. Außerdem vereinfacht sich das Verfahren wesentlich. Der Fall endlich vieler linearer Ungleichungen ist auch der Gegenstand der Arbeit von Agmon [2].

Es ist nicht überraschend, daß man hier in jedem Falle Endlichkeit des Verfahrens garantieren kann, wenn man die Iterierten geeignet darstellt. Wegen der in Abschnitt 8 dargestellten Vorgehensweise ist es keine Beschränkung der Allgemeinheit, das homogene Problem

$$\sum_{j=1}^{b} p_j \cdot a_j = \Theta_d, \qquad p_j \geq 0, \qquad \sum_{j=1}^{n} p_j = 1 \tag{10.1}$$

zu betrachten. Dabei ist

$$M = \{a_1, \ldots, a_n\} \subseteq \mathbb{R}^d$$

natürlich beschränkt, so daß auch hier keine Schwierigkeiten entstehen können. Wir setzen ohne Einschränkung der Allgemeinheit voraus, daß $\|a_j\| = 1$ für alle j.

Es sei nun $\sigma = \{i_0, \ldots, i_d\} \subseteq \{1, \ldots, n\}$. Dann definieren die a_i, $i \in \sigma$ ein – eventuell entartetes – Simplex S_σ. Es gibt nur endlich viele solche Simplices. Es sei nun $\Theta_d \notin S_\sigma$ für irgendein σ. Dann können nur endlich viele Iterierte in S_σ liegen. Wenn das Verfahren also nicht abbricht, wird x_r nach einer endlichen Anzahl von Iterationsschritten nur noch in solchen Simplices liegen, die auch Θ_d enthalten. Stellen wir also x_r nach dem Satz von Carathéodory (Satz 2.2) durch höchstens $d+1$ der a_j dar, dann ist auch θ_d als Konvexkombination dieser a_j darstellbar.

Das Problem (10.1) kann man auch mit der Simplexmethode [23, Kap. I] lösen. Diese Methode ist ein Eliminationsverfahren, also anfällig gegen Rundungsfehler. Tatsächlich spielt dieses Problem bei der praktischen Durchführung der Simplexmethode eine große Rolle. Das Verfahren (VF) wird dagegen durch Rundungsfehler kaum beeinflußt, sofern man nach einer gewissen Anzahl von Iterationen dafür sorgt, daß $\sum p_j^{(r)} = 1$ bleibt. Weiterhin ist die große Einfachheit der einzelnen Iterationsschritte bei dem Verfahren von Vorteil. Ein Nachteil ist allerdings die unter Umständen recht langsame Konvergenz.

Wir stellen die Anzahl der Rechenoperationen pro Iterationsschritt in Tabelle 10.1 zusammen. Die Vergleichszahlen für die Simplexmethode (in der sogenannten „modifizierten Form“) stammen von Arden [4].

Tabelle 10.1: Anzahl der Rechenoperationen pro Iterationsschritt für das Verfahren (VF) und die Simplexmethode.

Berechnung von	Additionen	Multiplikationen	Divisionen
y_r	$n \cdot d$	$n \cdot d$	–
μ_r	4	–	1
$\|x_{r+1}\|^2$	1	1	–
x_{r+1}	d	$2 \cdot d$	–
Insgesamt (VF)	$d \cdot (n+1) + 5$	$d \cdot (n+2) + 1$	1
Simplexmethode	$(d+1) \cdot (n + 6 \cdot d + 6)$	$(d+1) \cdot (n + 3 \cdot d + 3)$	$3 \cdot d + 2$

Der Vergleich ist allerdings nicht ganz statthaft. Bei der logisch komplizierteren Simplexmethode hängt die Rechenzeit außerordentlich davon ab, wie gut das Programm den speziellen Eigenschaften der Rechenanlage angepaßt ist.

Um das Verfahren (VF) einfacher zu gestalten, formulieren wir das Problem (10.1) in der Form

$$\begin{aligned} \sum_{j=1}^{n} p_j \cdot \langle a_j, a_i \rangle = 0 \quad i = 1, \ldots, n \\ p_j \geq 0, \quad \sum_{j=1}^{n} p_j = 1, \end{aligned} \tag{10.2}$$

wir bilden also die Gauß-Transformierte $A \cdot A^\top$ der Matrix $A = (a_1, \ldots, a_n)^\top$. Die beiden Aufgaben (10.1) und (10.2) sind offenbar einander äquivalent.

Mit

$$z_i^{(r)} = \sum_{j=1}^{n} p_j^{(r)} \cdot \langle a_j, a_i \rangle = \langle x_r, a_i \rangle$$

und

$$z_r = \left(z_1^{(r)}, \ldots, z_n^{(r)} \right)^\top$$

wird dann

$$z_{r+1} = (\langle x_{r+1}, a_i \rangle)_{i=1}^n = \mu_r \cdot z_r + (1 - \mu_r) \cdot (\langle a_{j_r}, a_i \rangle)_{i=1}^n ,$$

wobei j_r der Index des aus einer passenden Auswahlmenge bestimmten Elements von M ist. Weiterhin ist

$$\|x_{r+1}\|^2 = \mu_r \cdot \langle x_r, a_{j_r} \rangle + 1 - \mu_r = \mu_r \cdot z_j^{(r)} + 1 - \mu_r .$$

Damit erhalten wir die in Tabelle 10.2 angegebenen Rechenoperationen.Die Reduktion der Anzahl der Rechenoperationen geht zu Lasten des benötigten Speicherraumes. Man braucht bei dieser Transformation etwa n^2 Speicherplätze.

Pro Iteration wird jeweils nur eine Spalte der Gauß-Transformierten von A benötigt. Das bedeutet, daß man die Matrix auf einem langsamen Speichermedium abspeichern kann und nur die jeweils benötigte Spalte im Hauptspeicher mitführt.

Tabelle 10.2: Anzahl der Rechenoperationen pro Iterationsschritt für die vereinfachte Version von (VF).

Berechnung von	Additionen	Multiplikationen	Divisionen
j_r	–	–	–
μ_r	4	–	1
$\|x_{r+1}\|^2$	1	1	–
x_{r+1}	n	$2 \cdot n$	–
Insgesamt	$n+5$	$2 \cdot n + 1$	1

Wir betrachten ein Beispiel:

Beispiel 10.1 *Dieses Beispiel stammt von Hadley [54, Chap. 12, Problem 12–15]. Es ist* $d = 8$ *und* $n = 19$ *(in der homogenen Form (10.1)). Die Koeffizientenmatrix ist in Tabelle 10.3 zu finden.*

Nachdem die a_j *und* b *auf die Länge 1 normiert worden waren, wurde die Gauß-Transformation durchgeführt und iteriert. Nach 34 Iterationen war eine Lösung gefunden. Die dazu benötigte Rechenzeit an der Rechenanlage CDC 6400 der RWTH Aachen betrug 0.050 Sekunden (reine CP-Zeit). Dieselbe Aufgabe benötigte mit der Simplexmethode 4 Iterationen in 0.053 Sekunden.*

Zur genaueren Untersuchung wurde weiteriteriert. In Abbildung 10.1 ist die Anzahl der Iterationen aufgetragen, die benötigt wurden, um $\|x_r\|^2$ *um jeweils eine Zehnerpotenz zu verkleinern, Bei* $\|x_r\|^2 \leq 10^{-12}$ *kommt man in die Nähe der Maschinengenauigkeit (etwa 14 Dezimalstellen) und die Ergebnisse sind nicht mehr sinnvoll. Die Genauigkeit von* 10^{-12} *wurde nach 0.33 Sekunden erreicht.*

Wie man sieht, nähert sich die Anzahl der zur Steigerung der Genauigkeit um eine Zehnerpotenz erforderlichen Iterationsschritte dem Wert 278. Daraus kann man mit (6.1) eine obere Schranke für $|\alpha|$ *errechnen. Es ist* $|\alpha| \leq 0.09083$*. Das bedeutet, daß es einen Vektor* y *gibt, so daß der Winkel zwischen* y *und den* a_j *beziehungsweise* $-b$ *größer ist als* 84.39^o*. Das Problem gehört also zu den für das Verfahren nicht günstigen Beispielen.*

In einem weiteren Beispiel soll gezeigt werden, wie sich das Verfahren des Abschnitts 4 auf den endlichen Fall überträgt.

Beispiel 10.2 *Gesucht ist der maximale Eigenwert* λ^* *der Matrix*

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

Hall und Porsching [55] errechneten $\lambda^* = 2.681\ 787$*.*

Tabelle 10.3: Koeffizientenmatrix für Beispiel 10.1.

	1	2	3	4	5	6	7	8
a_1	25	0	0	1	0	0	0	0
a_2	15	0	0	0	1	0	0	0
a_3	10	0	0	0	0	1	0	0
a_4	5	0	0	0	0	0	1	0
a_5	−4	0	0	0	0	0	0	1
a_6	0	15	1	1	0	0	0	0
a_7	0	5	1	0	1	0	0	0
a_8	0	0	1	0	0	1	0	0
a_9	0	−5	1	0	0	0	1	0
a_{10}	0	−14	1	0	0	0	0	1
a_{11}	1	0	0	0	0	0	0	0
a_{12}	0	1	0	0	0	0	0	0
a_{13}	0	0	−1	0	0	0	0	0
a_{14}	0	0	0	1	0	0	0	0
a_{15}	0	0	0	0	1	0	0	0
a_{16}	0	0	0	0	0	1	0	0
a_{17}	0	0	0	0	0	0	1	0
a_{18}	0	0	0	0	0	0	0	1
b	0	0	8	2	4	4	5	3

Mit dem Startvektor $x_0 = (1,1,1,1)^\top$ *erhält man nach Satz 4.2 die Einschließung*

$$2 \le \lambda^* \le 3.$$

$\phi_\lambda(x_0) = \|A\cdot x_0 - \lambda\cdot x_0\|^2$ *ist minimal für* $\lambda_0 = 2.75$. *Nach fünf Iterationen wird* $A\cdot x_5 - \lambda_0\cdot x_5 \le 0$, *also ist wieder Satz 4.2 anwendbar. Es ergibt sich die Einschließung*

$$2.572\ 549 \le \lambda^* \le 2.742\ 612.$$

$\phi_\lambda(x_5)$ *ist minimal für* $\lambda_1 = 2.692\ 243$ *mit* $\phi_{\lambda_1}(x_5) = 0.000\ 847$.

Demgegenüber lieferten zehn Iterationen mit dem Verfahren von v. Mises – der Rechenaufwand zweier Iterationen dieses Verfahrens entspricht etwa dem einer Iteration des Verfahrens (VF) – die Einschließung

$$2.648\ 940 \le \lambda^* \le 2.715\ 440.$$

Das Verfahren von v. Mises ist also hier überlegen. Es sei jedoch darauf hingewiesen, daß die Methode des Abschnitts 4 nicht von der Vielfachheit des gesuchten Eigenwerts beziehungsweise von der Lage der anderen Eigenwerte bezüglich des gesuchten abhängt.

Es ist möglich, die Konvergenz des Verfahrens (VF) unter Umständen zu verbessern. Ist $\alpha = 0$ (beziehungsweise $|\alpha|$ klein), dann müssen gewisse $p_j^{(r)}$ mit $r \to \infty$ gegen Null streben. Mit anderen

Worten: Es wird ein linearer Teilraum identifiziert, so daß die für die Lösung in Frage kommenden a_j in diesem Teilraum liegen (vgl. auch Satz 7.5). Iteriert man nun nach (VF), dann werden nach einer gewissen Anzahl von Iterationen einige der $p_j^{(r)}$ klein gegenüber den anderen sein. Startet man erneut mit einem Startvektor, der sich nur aus denjenigen a_j zusammensetzt, für die die $p_j^{(r)}$ groß sind, dann ist eine Beschleunigung der Konvergenz zu erwarten[1].

Deggin [30] untersuchte diese und andere Möglichkeiten zur Konvergenzverbesserung an einigen Beispielen. Bei der Aufgabe aus Beispiel 10.1 liegt das Quadrat der Norm des Defekts nach 20 Iterationen zwischen 10^{-2} und 10^{-3}. Startet man erneut wie beschrieben mit einem neuen Startvektor, dann verschlechtert sich das Ergebnis zunächst auf 10^{-1}, jedoch bereits nach zehn weiteren Iterationen wird eine Genauigkeit von 10^{-6} erreicht.

[1] Siehe hierzu auch [36].

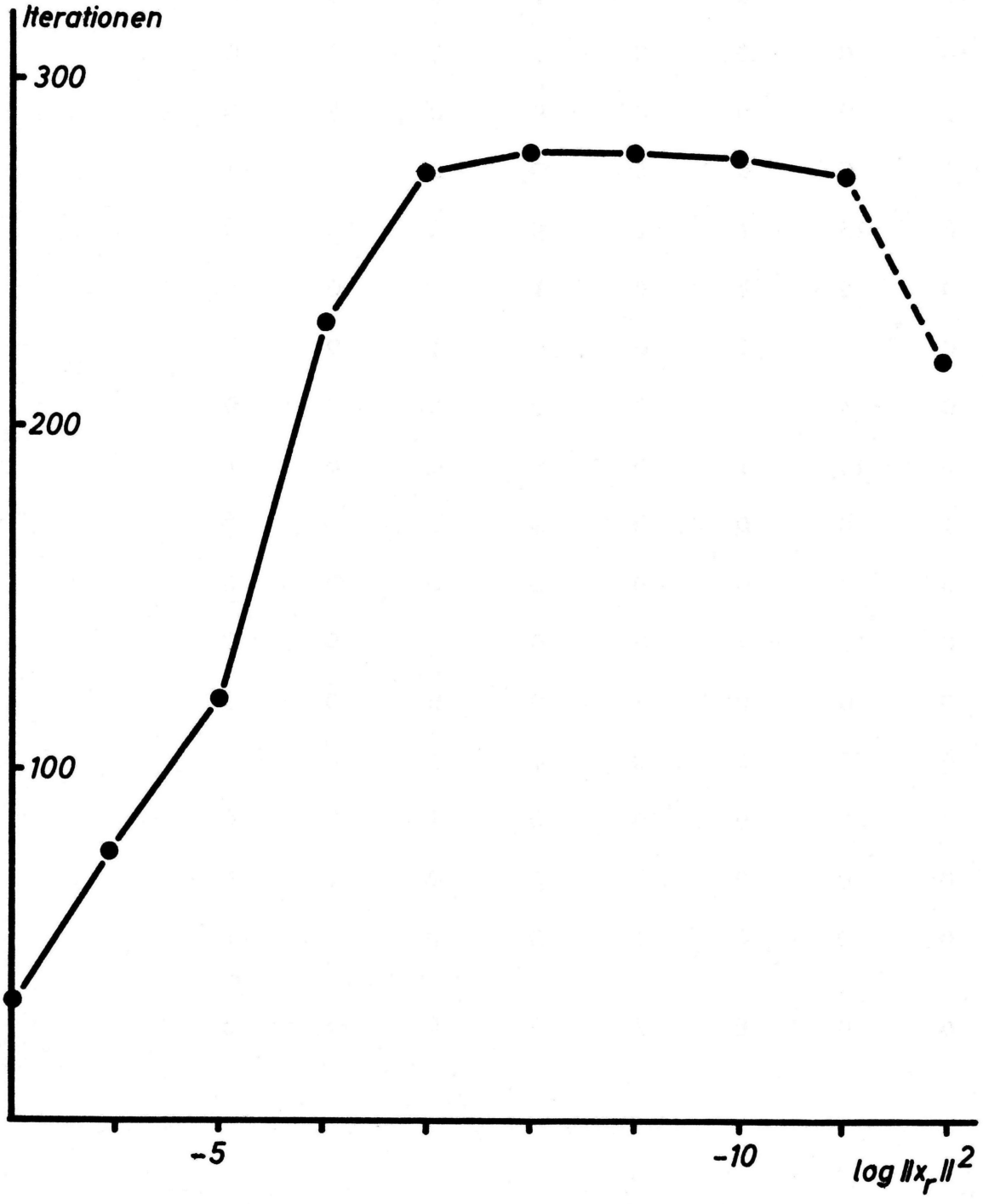

Abbildung 10.1: Anzahl der Iterationsschritte, die erforderlich sind, um $\|x_r\|^2$ auf $\|x_r\|^2/10$ zu verkleinern (Beispiel 10.1).

Kapitel 11

Das Verfahren von Agmon

Zum Abschluß soll das Verfahren (VF) mit der Methode von Agmon [2] verglichen werden. Bei dieser Methode versucht man, das Problem

Gesucht ist ein Vektor $x \in \mathbb{R}^d$, so daß

$$\langle a_j, x \rangle \geq b_j \qquad j = 1, \ldots, n \tag{11.1}$$

zu lösen. Dabei sei wieder $\|a_j\| = 1$ für alle j.

Ausgehend von einem beliebigen Startvektor $x_0 \in \mathbb{R}^d$ setzt man für $r = 0, 1, \ldots$ – falls x_r nicht bereits Lösung von (11.1) ist –

$$x_{r+1} = x_r - \sigma \cdot (\langle x_r, a_{j_r} \rangle - b_{j_r}) \cdot a_{j_r},$$

wobei σ ein fest gewählter Relaxationsfaktor mit $0 < \sigma \leq 2$ ist und j_r so bestimmt, daß $\langle x_r, a_{j_r} \rangle - b_{j_r}$ minimal ist.

Interessant an diesem Verfahren ist, daß es – falls (11.1) lösbar ist – in jedem Falle nach endlich vielen Schritten abbricht und daß die Konvergenz dann stets von der Ordnung $O(\theta^r)$ ist, wo $0 < \theta < 1$. θ ist dabei nicht so einfach geometrisch charakterisierbar wie α. Vielmehr hängt θ davon ab, wo der Grenzwert der Folge $\{x_r\}$ bezüglich der durch die Ungleichungen (11.1) definierten Hyperebenen $\langle a_i, x \rangle = b_i$ liegt. Es wird bei der Definition von θ wesentlich benutzt, daß die Anzahl der Ungleichungen (11.1) endlich ist, weshalb sich das Verfahren auch nicht verallgemeinern käßt. Eine eingehende Untersuchung von θ für verschiedene Vektornormen findet man bei Hoffman [62].

Wählt man $x_0 \in \text{cone } (\{a_j\})$, dann ist auch $x_r \in \text{cone } (\{a_j\})$ wegen $\langle x_r, a_{j_r} \rangle - b_{j_r} < 0$, falls x_r keine Lösung von (11.1) ist. Ist (11.1) nicht lösbar, dann lassen sich keine sinnvollen Schlüsse aus dem Verhalten der Iterierten ziehen. Sicher kann dann die Folge der x_r nicht konvergieren. Wenn (11.1) nicht lösbar ist, gibt es nämlich eine Zahl $\delta > 0$, so daß für jedes r gilt

$$\min_k (\langle x_r, a_j \rangle - b_j) < -\delta.$$

Dann ist

$$\|x_{r+1} - x_r\|^2 = \sigma^2 \cdot (\langle x_r, a_{j_r} \rangle - b_{j_r})^2 \geq \sigma^2 \cdot \delta^2.$$

Wir betrachten ein einfaches Beispiel:

Beispiel 11.1 *Vorgelegt sei das Ungleichungssystem*

$$\begin{array}{rcrcl} - & \dfrac{\xi_1}{\sqrt{2}} & - & \dfrac{\xi_2}{\sqrt{2}} & \geq & \dfrac{1}{\sqrt{2}} \\ & \xi_1 & & & \geq & 0 \\ & & & \xi_2 & \geq & 0. \end{array}$$

Wir wählen als Startvektor $x_0 = \Theta_2$.

1. *Für* $\sigma = 1$ *wird*

$$x_1 = -\frac{1}{2} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad x_2 = -\frac{1}{2} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad x_3 = x_0 = \Theta_2.$$

   Die Folge der Iterierten durchläuft also einen Zyklus.

2. *Ist* $\sigma = \dfrac{3}{2}$, *dann hat die Folge der* x_r *die drei Häufungspunkte*

$$\frac{1}{2} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \qquad \begin{pmatrix} 1/2 \\ -1 \end{pmatrix}.$$

3. *Setzt man schließlich* $\sigma = 2$, *dann ist für* $k = 0, 1, \ldots$

$$x_{3k} = \begin{pmatrix} k \\ k \end{pmatrix}, \qquad x_{3k+1} = \begin{pmatrix} -k-1 \\ -k-1 \end{pmatrix}, \qquad x_{3k+2} = \begin{pmatrix} k+1 \\ -k-1 \end{pmatrix}.$$

   also $\{x_r\}$ *divergent.*

   Es sind damit alle Möglichkeiten der Nichtkonvergenz gegeben.

Von Vorteil erscheint bei dem Verfahren von Agmon, daß die lineare Konvergenz in jedem Falle garantiert ist, wenn überhaupt Konvergenz vorliegt. Nun ist aber die Konvergenzgeschwindigkeit, beziehungsweise der Wert von θ, nur sehr schwer geometrisch zu interpretieren. Man kann also – außer für $\alpha = 0$ – nicht sicher sein, daß das Agmonsche Verfahren besser konvergiert als (VF). $\alpha = 0$ bedeutet aber, daß schon bei einer kleinen Änderung der Problemdaten das Problem unlösbar wird. In diesem Falle würde das Agmonsche Verfahren nicht konvergieren. Man kann also erwarten, daß für kleines $|\alpha|$ auch bei dem Verfahren von Agmon Schwierigkeiten auftreten.

Ein numerischer Test mit dem Beispiel 10.1 in der homogenen Version ergab als untere Schranke für θ den Wert 0.9918, falls $\sigma = 1$ gewählt wurde. Zur Erreichung einer Genauigkeit von 10^{-11} benötigte das Verfahren von Agmon 1 921 Iterationen, das Verfahren (VF) 1836.

Liegt bei dem Verfahren (VF) langsame Konvergenz vor, dann kann man daraus schließen, daß $|\alpha|$ klein ist, eine Situation, die auch bei der Anwendung der Simplexmethode zu Schwierigkeiten führen kann. Durch Rundungsfehler kann es für kleines $|\alpha|$ durchaus geschehen, daß die Simplexmethode mit einer falschen Aussage über die Lösbarkeit abbricht.

Kapitel 12

Lineare Ungleichungssysteme in der numerischen Mathematik

Es gibt eine Reihe von Anwendungen linearer Ungleichungssysteme in der numerischen Mathematik, von denen wir hier einige kurz skizzieren wollen. Naheliegend ist natürlich, daß lineare Ungleichungssysteme im Zusammenhang mit linearen beziehungsweise konvexen Optimierungsaufgaben Bedeutung haben. Es sei nur eine Aufgabe genannt, die eine große praktische Bedeutung hat, nämlich die Identifikation redundanter Ungleichungen eines linearen Ungleichungssystems

$$\sum_{j=1}^{d} a_{ij} \cdot x_j \geq b_i, \qquad i = 1, \ldots, n.$$

Die erste dieser Ungleichungen heißt (streng) redundant, wenn das Ungleichungssystem

$$\begin{aligned} \sum_{j=1}^{d} a_{1j} \cdot x_j &\leq b_1, \\ \sum_{j=1}^{d} a_{ij} \cdot x_j &\geq b_i, \qquad i = 2, \ldots, n \end{aligned}$$

nicht lösbar ist. Eine redundate Ungleichung kann weggelassen werden, ohne die Menge der Lösungen des Ungleichungssystems zu ändern. Praktische Bedeutung hat diese Frage bei der Reduktion großer linearer Ungleichungssysteme, wie sie bei linearen Optimierungsaufgaben auftreten. Zur Identifikation einer redundanten Ungleichung muß man also zeigen, daß ein gewisses Ungleichungssystem nicht lösbar ist. Eine Unlösbarkeitsaussage kann man aber mit (VF) bereits nach endlich vielen Schritten erhalten. Um eine Ungleichung eines Ungleichungssystems grob auf Rdundanz zu testen, kann man etwa eine vorgegebene Anzahl von Iterationsschritten mit (VF) ausführen. Bricht das Verfahren dabei ab, ist die fragliche Ungleichung redundant. Andernfalls muß man sich überlegen, ob man ein schärferes Kriterium anwenden oder aber die Frage auf sich beruhen lassen möchte.

Lineare und rationale Tschebyscheff-Approximationsprobleme kann man auch auf lineare Ungleichungssysteme zurückführen [19, S. 73 und S. 170]. Das Kolmogoroff-Kriterium [78, Satz 16 bis 18] charakterisiert die beste Approximierende durch ein lineares Ungleichungssystem.

Collatz [22, §25.6] gibt an, wie man durch Lösung eines Ungleichungssystems Schranken für die Minimalabweichung finden kann. Auf lineare Ungleichungssysteme kann man insbesondere Approximationsprobleme mit Vorteil zurückführen, wenn noch zusätzliche Nebenbedingungen vorgegeben sind (positive Approximation [71, 73]).

Weitere Anwendungen finden sich im Gebiet der monotonen Operatoren und der Operatoren monotoner Art. Sei etwa $T = T_1 + T_2$, T_1 isoton, T_2 antiton, ein monoton zerlegbarer Operator [22, §21.2]. Zur Einschließung einer Lösung von $u = Tu + r$ sucht man Funktionen v_0, w_0 mit

$$v_0 \leq T_1 v_0 + T_2 w_0 + r \leq T_1 w_0 + T_2 v_0 + r \leq w_0.$$

Sind dabei T_1 und T_2 lineare Operatoren und setzt man $v_0 = \sum c_j \cdot \phi_j$, $w_0 = \sum d_j \cdot \phi_j$, so erhält man ein lineares Ungleichungssystem für die Koeffizienten c_j und d_j (das natürlich nicht unbedingt lösbar sein muß).

Ebenso kann man versuchen, die Lösung einer linearen Operatorgleichung $Tu = r$ durch einen Ansatz $u = \sum c_j \cdot \phi_j$ einzuschließen, wenn T von monotoner Art ist [22, §21.1, §23]. Zu lösen ist dann das lineare Ungleichungssystem

$$\begin{aligned} \sum c_j \cdot T\phi_j &\leq r, \\ \sum d_j \cdot T\phi_j &\geq r. \end{aligned}$$

Bei einer Matrix kann man häufig über das sogenannte Zeilensummenkriterium (beziehungsweise das schwache Zeilensummenkriterium) testen, ob monotone Art vorliegt. Daneben gibt es aber einen allgemeineren Test, der auf ein lineares Ungleichungssystem führt [22, §23.1, Satz 1]. Die Matrix

$$\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}$$

genügt beispielsweise nicht dem Zeilensummenkriterium, jedoch dem allgemeineren Satz, ist also von monotoner Art.

Es sei $C \subseteq \mathbb{R}^d$ eine abgeschlossene beschränkte konvexe Menge, $f : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ eine stetige Funktion. $F(x) = (f_1(x), \ldots, f_d(x))^\top$. Gibt es dann ein $y \in \operatorname{int} C$ mit

$$\langle y - x, F(x) \rangle \geq 0 \qquad \text{für alle } x \in \operatorname{bd} C,$$

dann hat das Gleichungssystem $F(x) = \Theta_d$ eine Lösung in C [88, Theorem 6.3.4]. Beschränkt man sich auf eine Menge C, die durch ein lineares Ungleichungssystem dargestellt wird, dann kann man durch Lösung eines linearen Ungleichungssystems Nullstellen eines nichtlinearen Gleichungssystems einschließen.

Herrn Prof. Dr. W. Krabs möchte ich besonders danken für zahlreiche nützliche Hinweise und Anregungen zu dieser Arbeit.

Herrn Prof. Dr. F. Reutter danke ich für die Übernahme des Korreferats.

Literaturverzeichnis

[1] N. I. Achieser and I. M. Glasmann. *Theorie der linearen Operatoren im Hilbert-Raum.* Mathematische Lehrbücher und Monographien, I. Abteilung, Band IV. Akademie-Verlag, Berlin, dritte, unveränderte Auflage, 1960.

[2] S. Agmon. The relaxation method for linear inequalities. *Canad. J. Math.*, 6:382–392, 1954.

[3] A. Albert. A mathematical theory of pattern recognition. *Ann. Math. Statist.*, 34:284–299, 1963.

[4] D. N. Arden. The solution of linear programming problems. In A. Ralston and H. S. Wilf, editors, *Mathematical Methods for Digital Computers*, pages 263–279. John Wiley & Sons, New York, London, Sydney, 1960.

[5] A. G. Arkadjew and E. M. Brawerman. *Zeichenerkennung und maschinelles Lernen.* Beiheft Elektronische Rechenanlagen, Bd. 13. R. Oldenbourg Verlag, München und Wien, 1966.

[6] В. З. Беленкий, В. А. Волконский, С. А. Иванков, А. Б. Поманский и А. Д. Шапиро. Итеративные методы в теории игр и программировании. «Наука», Москва, 1974.

[7] H. D. Block. The Perceptron: A model for brain functioning. I. *Reviews of Modern Phys.*, 34:123–135, 1962.

[8] R. B. Braithwaite. A terminating iterative algorithm for solving certain games and related sets of linear equations. *Naval Res. Logistics Quarterly*, 6:63–74, 1959.

[9] Э. М. Браверман. Опыты по обучению машины разпознаванию зрительных образов. Автоматика и Телемеханика, XXIII(3):349–364, 1962.

[10] Л. М. Брэгман. Нахождение общей точки выпуклых множеств методом последовательного проектирования. Д. А. Н. СССР, 162:487–490, 1965.

[11] L. M. Bregman. Proof of the convergence of Sheleikhovskii's method for a problem with transportation constraints. *USSR Comput. Math. math. Phys.*, 7(1):191–204, 1967.

[12] L. M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems of convex programming. *USSR Comput. Math. math. Phys.*, 7(3):200–217, 1967.

[13] G. W. Brown and J. v. Neumann. Solutions of games by differential equations. In *Contributions to the theory of games, Vol. I.*, Annals of Mathematics Studies, No. 24, pages 73–79. Princeton University Press, Princeton, New Jersey, 1950.

[14] George W. Brown. Iterative solution of games by fictitious play. In Tjalling C. Koopmans, editor, *Activity Analysis of Production and Allocation. Proceedings of a Conference*, Cowles Commission for Research in Economics, Monograph No. 13, chapter XXIV, pages 374–375. John Wiley & Sons, Inc., New York · London · Sydney, 1951.

[15] В. А. Булавский. Итеративный метод решения задачи линейного программирования. Д. А. Н. СССР, 137:258–260, 1961.

[16] Ju. V. Čaĭkovskiĭ. On a continuous process of matrix game solution. *Soviet Math. Dokl.*, 12:1245–1247, 1971.

[17] C. Carathéodory. Über den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen. *Rendiconti del Circolo Matematico di Palermo*, XXXII(1):193–217, Dezember 1911.

[18] A. Charnes, W. W. Cooper, and K. O. Kortanek. On the theory of semi–infinite programming and a generalization of the Kuhn–Tucker saddle point theorem for arbitrary convex functions. *Naval Res. Logist. Quart.*, 16:41–51, 1969.

[19] E. W. Cheney. *Introduction to Approximation Theory.* McGraw–Hill Book Company, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[20] Elliott Ward Cheney and Will Light. *A Course in Approximation Theory.* Brooks/Cole Publishing Company, Pacific Grove · Albany · Belmont · Boston · Cincinnati · Johannesburg · London · Madrid · Melbourne · Mexico City · New York · Scottsdale · Singapore · Tokyo · Toronto, 2000.

[21] Lothar Collatz. Einschließungssatz für die charakteristischen Zahlen von Matrizen. *Math. Z.*, 48:221–226, 1941.

[22] Lothar Collatz. *Funktionalanalysis und numerische Mathematik.* Die Grundlehren der mathematischen Wissenschaften, Band 120. Springer-Verlag, Berlin, Göttingen, Heidelberg, 1964.

[23] Lothar Collatz and Wolfgang Wetterling. *Optimierungsaufgaben.* Heidelberger Taschenbücher, Band 15. Springer-Verlag, Berlin, Heidelberg, New York, 1966.

[24] Corinna Cortes and Vladimir Vapnik. Support vector networks. *Machine Learning*, 20(3):273–297, September 1995.

[25] Colin W. Cryer. On the calculation of the largest eigenvalue of an integral equation. *Numer. Math.*, 10:435–443, 1967.

[26] Colin W. Cryer. The method of Christopherson for solving free boundary problems for infinite journal bearings by means of finite differences. Technical Report 72, Computer Sciences Department, University of Wisconsin, December 1969.

[27] Colin W. Cryer. The method of Christopherson for solving free boundary problems for infinite journal bearings by means of finite differences. *Mathematics of Computation*, 25:435–443, 1971.

[28] Colin W. Cryer. The solution of a quadratic programming problem using systematic overrelaxation. *SIAM J. Control and Optimization*, 9(3):385–392, 1971.

[29] J. M. Danskin. Fictitious play for continuous games. *Naval Res. Log. Quart.*, 1:313–320, 1954.

[30] D. Deggim. Auflösung linearer Ungleichungssysteme durch Verfahren, die nicht auf Elimination beruhen. Diplomarbeit, RWTH Aachen, 1972.

[31] D. E. D'Esopo and B. Lefkowitz. An algorithm for computing interzonal transfers using the gravity model. *Operations Research*, 11:901–907, 1963.

[32] R. E. Duda and H. Fossum. Pattern classification by iteratively determined linear and piecewise discriminant functions. *IEEE Transactions on Electronic Computers*, EC–15:220–232, 1966.

[33] B. Curtis Eaves. The linear complementarity problem. *Management Science*, 17:612–634, 1971.

[34] Ulrich Eckhardt. Einige Eigenschaften Wilfscher Quadraturformeln. *Numerische Mathematik*, 12:1–7, 1968.

[35] Ulrich Eckhardt. Pseudo–complementary algorithms for mathematical programming. In F. Lootsma, editor, *Numerical Methods for Non–Linear Optimization*, pages 301–312. Academic Press, London, New York, 1972.

[36] Ulrich Eckhardt. Ein Iterationsverfahren zur Berechnung nichtnegativer Lösungen eines linearen Gleichungssystems. *Zeitschrift für Angewandte Mathematik und Mechanik*, 54:T213–T215, 1974.

[37] E. Eisenberg and E. Wong. Iterative synthesis of threshold functions. *J. Math. Anal. Appl.*, 11:226–235, 1965.

[38] H. Engels and R. Mangold. Über eine Klasse Wilfscher Quadraturformeln. Berichte der KFA Jülich, Jül–789–MA, KFA Jülich, 1971.

[39] И. И. Еремин. Обобщение релаксационного метода Моцкина – Агмона. Успехи математических наук, 20(2):183–187, 1965.

[40] И. И. Еремин. Релаксационный метод решения систем неравенств с выпуклими функционалями в левых частях. ДАН СССР, 160:994–996, 1965.

[41] I. I. Eremin. Iterative methods in convex programming. Эконом. и Мат. Методы, 6(2), 1966.

[42] I. I. Eremin. The application of the method of Fejer approximations to the solution of convex programming with nonsmooth constraints. *USSR Comput. Math. math. Phys.*, 9(5):225–235, 1969.

[43] И. И. Еремин и Вл. Д. Мазуров. Итерационный метод решения задаци выпуклого программирования. ДАН СССР, 170:57–60, 1966.

[44] D. K. Faddeew and W. M. Faddejewa. *Numerische Methoden der linearen Algebra.* Mathematik für Naturwissenschaft und Technik Band 10. Deutscher Verlag der Wissenschaften, Berlin, zweite, berichtigte Auflage, 1970.

[45] R. P. Fedorenko. An experiment in the iterational solution of linear programming problems. *USSR Comput. Math. math. Phys.*, 5(4):169–180, 1965.

[46] L. Fejér. Über die Lage der Nullstellen von Polynomen, die aus Minimumforderungen gewisser Arten entspringen. *Math. Ann.*, 85:41–48, 1922.

[47] V. M. Fridman and V. S. Chernina. An iteration process for the solution of the finite–dimensional contact problem. *USSR Comput. Math. Math. Phys.*, 7(1):210–214, 1967.

[48] R. Gastinel. Sur certains procédés itératifs non linéaires de reésolution de systèmes d'equations du premier degré. In C. M. Popplewell, editor, *Information Processing 1962. Proceedings of the IFIP Congress 1962*, pages 97–101. North Holland Publ. Comp., Amsterdam, 1963.

[49] H. J. Greenberg and A. G. Konheim. Linear and nonlinear methods in pattern classification. *IBM J. Res. and Develop.*, 8:299–307, 1964.

[50] J. S. Griffin, Jr., J. H. King, Jr., and C. J. Tunis. A pattern identification system using linear decision functions. *IBM Systems J.*, 2:248–267, 1963.

[51] L. G. Gurin, B. T. Polyak, and E. V. Raik. The method of projections for finding the common point of convex sets. *USSR Comput. Math. math. Phys.*, 7(6):1–24, 1967.

[52] S. Haber. The error in numerical integration of analytic functions. *Quart. Appl. Math.*, 29:411–420, 1971.

[53] G. J. Habetler and A. L. Price. Existence theory for generalized nonlinear complementarity problems. *Journal of Optimization Theory and Applications*, 7:223–239, 1971.

[54] G. Hadley. *Linear Programming.* Addison–Wesley Publishing Company, Reading, Massachusetts · Menlo Park, California · London · Sydney · Manila, 1962.

[55] C. A. Hall and R. A. Porsching. Computing the maximal eigenvalue and eigenvector of a positive matrix. *SIAM J. Numer. Anal.*, 5:269–274, 1968.

[56] M. Hershkowitz. A computational note on von Neumann's algorithm for determining optimal strategy. *Naval Res. Logist. Quart.*, 11:75–78, 1964.

[57] Heinrich R. Hertz. Über die Berührung fester elastischer Körper. *Journal für die reine und angewandte Mathematik*, 92:156–171, 1882.

[58] W. H. Highleyman. Linear decision functions, with applications to pattern recognition. *Proceedings of the IRE*, 50:1501–1514, 1962.

[59] Y. C. Ho and R. L. Kashyap. An algorithm for linear inequalities and its applications. *IEEE Transactions on Electronic Computers*, EC–14:683–688, 1965.

[60] Y. C. Ho and R. L. Kashyap. A class of iterative procedures for linear inequalities. *SIAM J. Control*, 4:112–115, 1966.

[61] A. Hoffman, M. Mannos, D. Sokolowsky, and N. Wiegman. Computational experience in solving linear programs. *SIAM J.*, 1:17–33, 1953.

[62] A. J. Hoffman. On approximate solutions of linear inequalities. *J. Res. National Bureau Standards*, 49:263–265, 1952.

[63] Leonid Hurwicz. Programming in linear spaces. In Kenneth J. Arrow, Leonid Hurwicz, and Hirofumi Uzawa, editors, *Studies in Linear and Non–Linear Programming*, Stanford Mathematical Studies in the Social Sciences, II, pages 38–102. Stanford University Press / Oxford University Press, Stanford, California / London, 1958.

[64] H. Ishida and R. M. Stewart, Jr. A learning network using adaptive threshold elements. *IEEE Transactions on Electronic Computers*, EC–14:481–485, 1965.

[65] L. W. Johnson and R. D. Riess. Minimal quadratures for functions of low-order continuity. *Math. Comp.*, 25:831–835, 1971.

[66] Victor Klee and George J. Minty. How good is the simplex algorithm? In O. Shisha, editor, *Inequalities III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin*, pages 159–175. Academic Press, New York, London, 1972.

[67] K. Knopp. *Theorie und Anwendung der unendlichen Reihen.* Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Band 11. Springer-Verlag, Berlin, Göttingen, Heidelberg, 4. Auflage, 1947.

[68] W. Krabs. Ein Pseudo-Gradientenverfahren zur Lösung des diskreten linearen Tschebyscheff-Problems. *Computing*, 4:216–224, 1968.

[69] В. И. Крылов. Приближенное вычисление интегралов. Физматгиз, Москва, 1959.

[70] C. E. Lemke. On complementary pivot theory. In G. B. Dantzig and A. F. Veinott, editors, *Mathematics of the Decision Sciences, Part I*, Lectures in Applied Mathematics, Volume 11, pages 95–114. American Mathematical Society, Providence, RI, 1968.

[71] R. A. Lorentz. Positive Approximation. Berichte der Gesellschaft für Mathematik und Datenverarbeitung No. 49, BMBW – GMD – 49, Bonn, 1971.

[72] Yu. I. Lyubich and G. D. Maistrovskii. The general theory of relaxation processes for convex functionals. *Russian Mathematical Surveys*, 25:57–117, 1979.

[73] G. Marsaglia. One–sided approximation by linear combinations of functions. In A. Talbot, editor, *Approximation Theory. Proceedings of a Symposium held at Lancaster, July 1969*, pages 233–242. Academic Press, London, New York, 1970.

[74] E. A. Matrose and R. Mukundan. Modelling of recursive games of survival as finite state sequential machines and solutions by fictitious play. *Int. J. Systems Sci.*, 2:369–377, 1972.

[75] C. H. Mays. Effect of adaptation parameters on convergence time and tolerance for adaptive threshold elements. *IEEE Transactions on Electronic Computers*, EC–13:465–468, 1964.

[76] C. H. Mays. The relationship of algorithms used with adjustable threshold elements to differential equations. *IEEE Transactions on Electronic Computers*, EC–14:62–63, 1965.

[77] Günter Meinardus. *Approximation von Funktionen und ihre numerische Behandlung*, Springer Tracts in Natural Philosophy, Ergebnisse der angewandten Mathematik, Volume 4. Springer-Verlag, Berlin, Göttingen, Heidelberg, New York, 1964.

[78] Günter Meinardus. *Approximation of Functions: Theory and Numerical Methods. Translated by Larry L. Schumaker.* Springer Tracts in Natural Philosophy, Volume 15. Springer-Verlag, Berlin, Heidelberg, New York, 1967.

[79] Ju. I. Merzlyakov. On a relaxation method of solving systems of linear inequalities. *USSR Comput. Math. math. Phys.*, 2:504–510, 1962.

[80] Herbert Meschkowski. *Hilbertsche Räume mit Kernfunktion.* Grundlehren der mathematischen Wissenschaften, Band 113. Springer-Verlag, Berlin, Göttingen, Heidelberg, 1962.

[81] Б. Ф. Митчелл, В. Ф. Демьянов и В. Н. Малоземов. Нахождение ближайшей к началу координат точки многогранника. Вестник Ленинградского Университета, Серия Математики, Механики, Астрономии, № 19, выпуск 4:38–45, 1971.

[82] B. F. Mitchell, V. F. Dem'yanov, and V. N. Malozemov. Finding the point of a polyhedron closest to the origin. *SIAM J. Control*, 12:19–26, 1974.

[83] T. S. Motzkin and I. J. Schoenberg. The relaxation method for linear inequalities. *Canad. J. Math.*, 6:393–404, 1954.

[84] John von Neumann. A numerical method to determine optimum strategy. *Naval Res. Log. Quart.*, 1:109–115, 1954.

[85] N. J. Nilsson. *Learning Machines.* McGraw–Hill Book Company, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1965.

[86] A. Novikoff. On convergence proofs for perceptrons. In J. Fox, editor, *Mathematical Theory of Automata. Symposium, Proceedings, New York, 24–26 April 1962*, Microwave Research Institute Symposia Series, 12, pages 615–622. Polytechnic Institute of Brooklyn, Polytechnic Pr., Brooklyn NY, 1963.

[87] W. Oettli. An iterative method, having linear rate of convergence, for solving a pair of dual linear programs. *Math. Programming*, 3:302–311, 1972.

[88] J. M. Ortega and W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations of Several Variables.* Computer Science and Applied Mathematics. Academic Press, New York, London, 1970.

[89] B. T. Polyak. Gradient methods for solving equations and inequalities. *USSR Comput. Math. math. Phys.*, 4(6):17–32, 1964.

[90] B. T. Polyak. Minimization of unsmooth functionals. *USSR Comput. Math. math. Phys.*, 9(3):14–29, 1969.

[91] L. S. Presman. A complex transportation problem. *MATEKON, Translation of Russian and East European Mathematical Economics*, VIII:183–199, 1971. (Экономика и мат. методы, 1970, №. 6.).

[92] Iwan Iwanowitsch Priwalow. *Randeigenschaften analytischer Funktionen.* Hochschulbücher für Mathematik, Band 25. Deutscher Verlag der Wissenschaften, Berlin, 1956.

[93] Э. Раик. Методы фейеровского типа в гилбертовом пространстве. *Eesti NSV Teaduste Akadeemia Toimetised. XVI Köide. Füüsika Matemaatika*, № 3:286–293, 1967.

[94] F. Riesz and B. Sz.-Nagy. *Vorlesungen über Funktionalanalysis.* Hochschulbücher für Mathematik, Band 27. Deutscher Verlag der Wissenschaften, Berlin, 1956.

[95] Julia Robinson. An iterative method of solving a game. *Annals of Mathematics*, 54:296–301, 1951.

[96] R. Tyrrell Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970.

[97] Frank Rosenblatt. *Principles of Neurodynamics.* Spartan Books, Washington, D. C., 1962.

[98] S. Schechter. Minimization of a convex function by relaxation. In J. Abadie, editor, *Integer and Nonlinear Programming*, chapter 7, pages 177–189. North-Holland Publ. Comp., Amsterdam, London, 1970.

[99] H. N. Shapiro. Note on a computation method in the theory of games. *Communications on Pure and Applied Mathematics*, XI:587–593, 1958.

[100] L. S. Shapley. Some topics in two–person games. In M. Dresher, L. S. Shapley, and A. W. Tucker, editors, *Advances in Game Theory*, Annals of Mathematics Studies No. 52, pages 1–28. Princeton University Press, Princeton, New Jersey, 1964.

[101] F. W. Smith. Pattern classifier design by linear programming. *IEEE Transactions on Computers*, C–17:367–372, 1968.

[102] R. R. Sokal. Numerical taxonomy. *Scientific American*, 215(6):106–116, 1966.

[103] Josef Stoer and Christoph Witzgall. *Convexity and Optimization in Finite Dimensions I.* Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Band 163. Springer-Verlag, Berlin, Heidelberg, New York, 1970.

[104] K. Tanabe. Projection method for solving a singular system of linear equations and its applications. *Numer. Math.*, 17:203–214, 1971.

[105] J. J. Torsti and A. M. Aurela. A fast quadratic programming method for solving illconditioned systems of equations. *J. Math Anal. and Appl.*, 38:193–204, 1972.

[106] S. N. Tschernikow. *Lineare Ungleichungen. Übers.: H. Hollatz und H. Weinert.* Hochschulbücher für Mathematik, Band 69. Deutscher Verlag der Wissenschaften, Berlin, 1971.

[107] E. M. Vaisbord. The method of successive projection for the approximate solution of a problem of optimum control. *USSR Comput. Math. math. Phys.*, 6(6):35–48, 1966.

[108] Frederick A. Valentine. *Konvexe Mengen.* B.I. Hochschultaschenbücher 402/402a. Bibliographisches Institut, Mannheim, 1968.

[109] W. Vogel. *Lineares Optimieren.* Akademische Verlagsgesellschaft Geest & Portig K.–G., Leipzig, 1967.

[110] W. G. Wee and K. S. Fu. An adaptive procedure for multiclass pattern classification. *IEEE Transactions on Computers*, C–17:178–182, 1968.

[111] Herbert S. Wilf. Exactness conditions in numerical quadrature. *Numerische Mathematik*, 6:315–319, 1964.

Index