Nonconvex All-Quadratic Global Optimization Problems:

Solution Methods, Application and Related Topics

Dissertation

Zusammenfassung

Volltext im pdf-Format abrufbar unter:
http://ub-dok.uni-trier.de/diss/diss55/19990901/19990901.pdf

Ulrich Raber

Trier, 1999

ZUSAMMENFASSUNG

Die vorliegende Arbeit beschäftigt sich im Wesentlichen mit der Entwicklung und der sowohl theoretischen als auch numerischen Untersuchung von Lösungsmethoden für sogenannte *nichtkonvexe all-quadratische Optimierungsprobleme*, d.h. für Probleme der Form

$$\begin{aligned} \min\ x^T Q^0 x + (d^0)^T x & \\ x^T Q^l x + (d^l)^T x + c^l \ \le\ 0 \qquad & l = 1, \dots, p \\ x \in P\ , & \end{aligned} \tag{QP}$$

wobei $Q^l \in \mathbb{R}^{n\times n}$ symmetrisch, $d^l \in \mathbb{R}^n$ $(l = 0, \dots, p)$, $c^l \in \mathbb{R}$ $(l = 1, \dots, p)$ und $P = \{x \in \mathbb{R}^n : Ax \le b\}$ mit $A \in \mathbb{R}^{m\times n}$ und $b \in \mathbb{R}^m$ ein nichtleeres, volldimensionales Polytop ist.

In einem einleitenden Kapitel werden zunächst einige Anwendungen dieses Typs globaler Optimierungsprobleme vorgestellt. Des Weiteren werden nach der Beschreibung einiger grundlegender Konzepte in der deterministischen globalen Optimierung bereits bekannte Verfahren zur Lösung von (QP) kurz beleuchtet. Das 1. Kapitel schließt mit der Beschreibung der Konstruktion von zufällig erzeugten Test-Beispielen, die zur numerischen Untersuchung verschiedener Verfahren innerhalb dieser Dissertation verwendet wurden.

Im 2. Kapitel wird die Klasse sogenannter *unärer Optimierungsprobleme* untersucht. Diese ist im Zusammenhang mit all-quadratischen Problemen von Interesse, da jedes Problem vom Typ (QP) zu einem unären Problem transformiert werden kann. Mit einigem technischem Aufwand werden neue Verfahren zur Lösung solcher unärer Probleme entwickelt. Im Unterschied zu der äußeren Approximationsmethode, die von Ramana [Ram93] vorgeschlagen wurde und bisher der einzige bekannte Zugang zur Lösung unärer Probleme ist, kann die Konvergenz der in dieser Arbeit hergeleiteten Verfahren gesichert werden. Für das Verfahren von Ramana ist die Frage der Konvergenz nach wie vor offen. Von besonderem Interesse ist im 2. Kapitel die Konstruktion eines regulären n-Simplex, dessen Ecken auf dem Rand der Einheitskugel liegen. Die Eigenschaften dieser Menge sind in der Literatur bisher zwar untersucht worden (siehe z.B. [Som29, Sle69, GKL95]), jedoch wurde ein solcher Simplex bisher nicht explizit konstruiert – außer in [HR98].

Leider ist der Zugang zur Lösung von (QP) über unäre Probleme nur von theoretischem Interesse. Die numerischen Untersuchungen haben gezeigt, dass diese Verfahren nicht in der Lage sind, die unären Probleme, die aus der Transformation der zufällig erzeugten all-quadratischen Test-Beispiele resultieren, mit akzeptablem numerischem Aufwand zu lösen. Dabei ist zu beachten, dass der numerische Aufwand zur Lösung unärer Probleme entscheidend von der affinen Matrix-Abbildung abhängt, die die einzige nichtlineare Nebenbedingung solcher Probleme beschreibt. Im Fall der unären Probleme, die aus der Transformation von Problemen vom Typ (QP) resultieren, hat diese Abbildung eine sehr unangenehme Gestalt. Für unäre Probleme mit einer einfacheren Matrix-Abbildung ist es wahrscheinlich, dass die in dieser Arbeit entwickelten Verfahren ein wesentlich besseres numerisches Verhalten zeigen. Allerdings sind bisher außer den all-quadratischen Problemen keine Anwendungen bekannt, die zu unären Problemen führen. Somit wäre die Untersuchung unärer Probleme mit einer einfacheren Matrix-Abbildung nicht von praktischem Interesse.

Neben der äußeren Approximationsmethode von Ramana gibt es auch andere Schnittebenenverfahren zur Lösung globaler Optimierungsprobleme, bei denen die Konvergenz nicht nachgewiesen werden kann (siehe hierzu [HT96]). Die im 2. Kapitel vorgestellten Ideen zur Sicherung der Konvergenz der neu entwickelten Verfahren für unäre Probleme, die eine Kombination aus äußerer Approximation und schrittweiser Aufteilung der zulässigen Menge sind, können unter Umständen auch genutzt werden, um bei äußeren Approximationsmethoden mit unbekanntem Konvergenzverhalten die Konvergenz zu erzwingen. Dies ist ein weiterer interessanter Aspekt der Ergebnisse im zweiten Kapitel dieser Arbeit.

Im dritten Kapitel wird eine Lösungsmethode untersucht, die im Unterschied zum angesprochenen indirekten Zugang über unäre Probleme versucht, das Problem (QP) direkt zu lösen. Dieser neue simpliziale Branch-and-Bound Algorithmus zeigt ein wesentlich besseres numerisches Verhalten als die zuvor untersuchten Methoden. All-quadratische Probleme mit $n < 9$ können mit akzeptablem Rechenaufwand gelöst werden. Das vorgestellte Verfahren ist sogar, zumindest in Bezug auf unsere Test-Beispiele, oft entscheidend besser als eine Rechteck-Branch-and-Bound Methode, die von Al-Khayyal et al. [AKLV95] entwickelt worden ist. Diese Methode war bisher einer der wenigen Ansätze, die versuchen (QP), direkt zu lösen. Die meisten in der Literatur vorgeschlagenen Methoden betrachten (QP) nur als einen Spezialfall einer allgemeineren Klasse von Problemen. Diese allgemeineren Probleme sind z.B. bilineare Probleme [AK92], polynomiale Probleme [ST92], Probleme mit bikonvexen Funktionen [FV93] oder unäre Probleme [Ram93].

Die simpliziale Methode im 3. Kapitel basiert zwar auf den gleichen Konzepten wie der Ansatz in [AKLV95]. Durch die Verwendung von Simplices als Aufteilungsmengen anstelle von Rechtecken ist es allerdings möglich, die Komplexität der zur lösenden linearen Teilprobleme soweit zu reduzieren, dass das Verfahren schneller ist, obwohl mehr Iterationen durchlaufen werden müssen. Die Komplexität der Teilprobleme, d.h. deren Dimension und die Anzahl der linearen Nebenbedingungen, hängt in unserem Algorithmus linear von den Inputgrößen n und p ab. In Al-Khayyal et al.'s Zugang müssen Probleme mit der Dimension $(p + 2)n$ und $4(p + 1)n + p + m$ Nebenbedingungen gelöst werden, d.h. die Komplexität der Teilprobleme hängt multiplikativ von diesen Inputgrößen ab.

Eine wesentliche Schwierigkeit bei der Lösung all-quadratischer Probleme besteht darin, dass das Problem, einen zulässigen Punkt für (QP) zu finden, genau so schwer ist wie das Optimierungsproblem selbst. Deshalb muß man in der Praxis damit zufrieden sein, wenn ein Verfahren zur Lösung von (QP) nach endlicher Zeit entweder feststellt, dass der zulässige Bereich von (QP) leer ist, oder einen Punkt $\bar{x}$ findet, der approximativ zulässig ist, d.h. alle Nebenbedingungen bis zu einer gewissen Toleranz erfüllt, und zudem garantiert, dass der Optimalwert von (QP) nur im Rahmen einer bestimmten Genauigkeit unterhalb des Funktionswertes der Zielfunktion im Punkt $\bar{x}$ liegen kann. Alle in dieser Doktorarbeit vorgestellten Verfahren können so modifiziert werden, dass sie diese Eigenschaft erfüllen. Diese Modifizierung ist ohnehin notwendig, um die vorgeschlagenen Methoden numerisch zu testen.

Um die Konvergenz des oben angesprochenen simplizialen Branch-and-Bound Verfahrens zu sichern und somit die Endlichkeit dieser Methode hinsichtlich approximativer Lösungen zu gewährleisten, wird in der Formulierung des Algorithmus eine sogenannte *ausschöpfende* (exhaustive) Aufteilungsregel für die relevanten n-Simplices verwendet. Von der sogenannten *Bisektion*, bei der ein n-Simplex bezüglich des Mittelpunktes der längsten Kante in zwei Teilsimplices unterteilt wird, weiß man, dass sie ausschöpfend ist. In der Literatur wird im Zusammenhang mit simplizialen Branch-and-Bound Methoden zur Minimierung einer konkaven Funktion über einem Polytop, d.h. zur Lösung sogenannter *konkaver Minimierungsprobleme*, von einigen Autoren eine andere Aufteilungsregel, die sogenannte ω-*Subdivision*, favorisiert. Bei dieser Regel wird ein n-Simplex S in bis zu

$n+1$ Teilsimplices unterteilt, wobei als Aufteilungspunkt die Optimallösung der gelösten LP-Relaxierung bezüglich der Menge S verwendet wird. Da diese Aufteilungsregel nicht ausschöpfend ist, war es bisher nicht klar, ob ein simplizialer Branch-and-Bound Algorithmus, der nur diese Regel verwendet, konvergent ist.

Wir sind in erster Linie an Verfahren zur Lösung nichtkonvexer all-quadratischer Probleme interessiert, für die die Konvergenz gesichert werden kann bzw. für die zumindest gezeigt werden kann, dass sie in endlicher Zeit approximative Lösungen liefern, so dass sie numerisch untersucht werden können. Um die ω-Subdivision in unserem simplizialen Branch-and-Bound Verfahren anzuwenden, müssen wir uns deshalb zunächst mit der Frage der Konvergenz befassen. Im 4. Kapitel der vorliegenden Arbeit wird diese Fragestellung untersucht. Die Konzepte, die wir im dritten Kapitel zur Herleitung eines Lösungsverfahrens für (QP) angewendet haben, können in analoger Weise verwendet werden, um ein Verfahren zur Lösung sogenannter *allgemeiner d.c. Probleme* zu entwickeln. Diese Klasse von Problemen wird durch Funktionen beschrieben, die sich als Differenz konvexer Funktionen (**d**ifference of **c**onvex functions) darstellen lassen. Somit enthält diese Klasse insbesondere die all-quadratischen und die zuvor erwähnten konkaven Minimierungsprobleme.

Im 4. Kapitel wird anhand eines Gegenbeispiels gezeigt, dass die Variante des vorgestellten Verfahrens zur Lösung von d.c. Problemen, die nur ω-Subdivisionen verwendet, im allgemeinen nicht konvergent ist und zudem nach endlicher Zeit nicht notwendigerweise eine approximative Lösung liefert. Wenn man die Klasse der zur lösenden Probleme allerdings einschränkt, kann die Endlichkeit dieser Variante hinsichtlich approximativer Lösungen dennoch gesichert werden. Es wird bewiesen, dass diese Methode nach endlicher Zeit entweder feststellt, dass der zulässige Bereich des untersuchten Problems leer ist, oder eine approximative Lösung liefert, solange Probleme betrachtet werden, bei denen die Zielfunktion d.c. ist – mit einem strikt konkaven Anteil – und alle nichtlinearen Nebenbedingungen konkav sind. Für konkave Minimierungsprobleme gilt dasselbe Resultat auch ohne die Annahme der strikten Konkavität der Zielfunktion. Dies sind die wesentlichen theoretischen Ergebnisse des vierten Kapitels dieser Dissertation.

Neben diesem Endlichkeitsresultat hinsichtlich approximativer Lösungen wird in diesem Kapitel zudem gezeigt, dass die Variante des simplizialen Branch-and-Bound Verfahrens, die nur ω-Subdivisionen verwendet, für konkave Minimierungsprobleme nach endlicher Zeit sogar die optimale Lösung liefert, sofern zwei zusätzliche Bedingungen erfüllt sind. Die erste dieser Bedingungen ist keine Einschränkung der Allgemeinheit, da das untersuchte konkave Minimierungsproblem immer so äquivalent umgeformt werden kann, dass diese Bedingung erfüllt ist. Die zweite Bedingung kann im Allgemeinen nicht garantiert werden. Allerdings gilt sie für bestimmte Klassen derartiger Optimierungsprobleme, z.B. für konkave Minimierungsprobleme über höherdimensionalen Rechtecken.

Im 4. Kapitel wird das vorgeschlagene Verfahren zur Lösung von d.c. Problemen auch numerisch untersucht. In dieser Methode kann neben der ω-Subdivision auch die Bisektion als Aufteilungsregel verwendet werden, wobei für diese Variante die Konvergenz in jedem Fall gesichert werden kann, genauso wie im Fall des im 3. Kapitel untersuchten Algorithmus. Da der Algorithmus im vierten Kapitel im Unterschied zu dem Algorithmus im dritten Kapitel konvexe anstelle linearer Relaxierungen verwendet, ist es zunächst einmal interessant zu untersuchen, welche von beiden Methoden bei der Anwendung auf Probleme vom Typ (QP) die effizientere ist. Die numerischen Ergebnisse in Kapitel 4 zeigen, dass wir erwarten können, dass der simpliziale Branch-and-Bound Algorithmus mit konvexen Teilproblemen numerisch effizienter ist, zumindest solange ein Verfahren zur Lösung der konvex quadratischen Teilprobleme verwendet wird, das die quadratische Struktur ausnutzt.

Der wesentliche Grund für die Untersuchung der ω-Subdivision in dieser Arbeit ist die Hoffnung, dass die Anwendung dieser Regel zu einem schnelleren Verfahren zur Lösung von all-quadratischen Problemen führt. Wir haben zwar gezeigt, dass die Variante der im 4. Kapitel vorgestellten Methode, die nur ω-Subdivisionen durchführt, nicht notwendigerweise nach endlicher Zeit eine approximative Lösung für allgemeine quadratische Probleme vom Typ (QP) liefert. Allerdings ist es möglich, unter Verwendung der theoretischen Ergebnisse, die in dieser Arbeit hergeleitet werden, eine gemischte Aufteilungsstrategie zu entwickeln, die sowohl ω-Subdivisionen als auch Bisektionen verwendet und zudem die Endlichkeit des simplizialen Branch-and-Bound Algorithmus garantiert. Dabei ist zu beachten, dass die gemischte Strategie, die wir (MGWSR) nennen, nicht mit der gemischten Strategie übereinstimmt, die in sogenannten *normalen* Algorithmen verwendet wird (siehe hierzu z.B. [HT96]).

Wir haben diese gemischte Strategie und einige Varianten davon numerisch anhand unserer all-quadratischen Beispiele getestet. Leider sind die Resultate recht enttäuschend. Unser simpliziales Branch-and-Bound Verfahren hat zwar unter Verwendung dieser gemischten Strategie in einigen wenigen Test-Beispielen ein sehr gutes numerisches Verhalten gezeigt. Im Durchschnitt war diese Strategie aber zum Teil erheblich schlechter als die Bisektion. Es ist bisher keine Aufteilungsstrategie bekannt, die im Allgemeinen besser ist als die Bisektion. Insbesondere führt keine der in dieser Arbeit untersuchten Varianten von (MGWSR) zu besseren numerischen Ergebnissen. Wir können zwar die Endlichkeit des vorgestellten Verfahrens hinsichtlich approximativer Lösungen sichern, selbst wenn eine nicht notwendigerweise ausschöpfende Aufteilungsregel verwendet wird. Jedoch erfüllt sich unsere Hoffnung nicht, dass diese Regel zu einer numerischen Verbesserung der untersuchten simplizialen Branch-and-Bound Verfahren führt.

Es gibt eine Reihe von praktischen Anwendungen, die zu Problemen vom Typ (QP) führen. Die vorliegende Arbeit schließt mit der Untersuchung einer solchen Anwendung, dem sogenannten *Packing Problem*. Bei diesem Problem geht es um die Frage, wie n Kreise ($n \in \mathbb{N}$) im Einheitsquadrat so positioniert werden können, dass sich diese Kreise nicht überlappen und zudem der Radius der Kreise maximal ist. Das zu diesem Problem äquivalente *Punkt-Positionierungsproblem* läßt sich als all-quadratisches Optimierungsproblem in der Dimension $2n + 1$ mit linearer Zielfunktion und $\binom{n}{2}$ konkav quadratischen Nebenbedingungen formulieren. Die Lösungen des Packing Problems mit bis zu 20 Kreisen sind bekannt. Insofern ist es nur interessant, Probleme mit mehr als 20 Kreisen zu untersuchen. Dies bedeutet, dass globale Optimierungsprobleme mit mehr als 40 Entscheidungsvariablen und mit mehr als 190 nichtlinearen Nebenbedingungen gelöst werden müßten. Die in den vorangegangenen Kapiteln entwickelten allgemeinen Verfahren zur Lösung von Problemen vom Typ (QP) sind nicht in der Lage, Probleme in dieser Größenordnung zu lösen. Man sollte beachten, dass aus der Sicht der deterministischen globalen Optimierung solche Probleme sehr groß sind.

Um das Packing Problem bzw. das äquivalente Punkt-Positionierungsproblem dennoch für $n > 20$ zu lösen, wird im 5. Kapitel der vorliegenden Arbeit ein speziell auf dieses Problem zugeschnittenes Rechteck-Branch-and-Bound Verfahren vorgestellt. Dieses ist in der Lage, auf einer leistungsstarken Maschine für das Packing Problem mit bis zu 27 Kreisen innerhalb von zwei Stunden approximative

Lösungen zu ermitteln. Mit der Entwicklung dieses Verfahrens wird insbesondere die an mehreren Stellen innerhalb dieser Dissertation aufgestellte Behauptung, dass die Ausnutzung der Struktur spezieller Probleme zu einer erheblichen Verbesserung der numerischen Effizienz aller untersuchten Verfahren führen kann, untermauert. Im 5. Kapitel werden zunächst einige theoretische Eigenschaften des Punkt-Positionierungsproblems untersucht. Es wird gezeigt, dass Lösungen existieren, die ein ganz bestimmtes Verhalten am Rand des Einheitsquadrates zeigen. Diese Eigenschaften bestätigen die intuitiv naheliegende Vermutung, dass bei einer optimalen Lösung des Packing Problems möglichst viele Kreise den Rand des Einheitsquadrates berühren oder zumindest nahe am Rand liegen.

Unter Ausnutzung der speziellen Struktur des untersuchten Problems wird im 5. Kapitel eine spezielle LP-Relaxierung vorgestellt, die besser ist als jene, die man durch allgemeine Ansätze (siehe z.B. [ST92, AKLV95]) erhält. Des Weiteren werden spezielle Aufteilungsstrategien für die verwendeten höherdimensionalen Rechtecke entwickelt, die bei diesem speziellen Problem zu einer erheblich besseren numerischen Performance des Algorithmus führen als z.B. die bekannte Rechteckbisektion. Zudem werden unter Ausnutzung der Struktur weitere Strategien zur Manipulierung der Aufteilungsmengen entworfen, die insbesondere die angesprochenen theoretischen Eigenschaften bestimmter Lösungen ausnutzen. Bei der Untersuchung von Lösungsverfahren für allgemeine Problemklassen ist es normalerweise nicht möglich, solche zusätzlichen Manipulierungsstrategien zu erhalten.

Der resultierende Algorithmus ist theoretisch in der Lage, sogenannte *ϵ-optimale* Lösungen für das Packing Problem mit beliebig vielen Kreisen in endlicher Zeit zu bestimmen. Es ist zu beachten, dass die meisten Lösungen für das Packing Problem mit mehr als 20 Kreisen nicht bekannt sind. In der Literatur sind zwar bereits "gute" Lösungen für bis zu 50 Kreise vorgestellt worden [NO97], allerdings ist die Qualität dieser Lösungen hinsichtlich ihrer Optimalität nicht bekannt. Die garantierte ϵ-Optimalität der berechneten Lösungen ist der wesentliche Vorteil des hier vorgestellten neuen Verfahrens zur Lösung des Packing Problems im Unterschied zu den verwendeten Lösungsansätzen zur Bestimmung "guter" Lösungen. Ferner kann, wie schon zuvor gesagt, mit dem vorgestellten Verfahren das Packing Problem mit bis zu 27 Kreisen in 2 Stunden approximativ gelöst werden. Zudem wurden – zwar mit einem erheblich höherem Rechenaufwand – Lösungen mit bis zu 31 Kreisen bestimmt und für 32 Kreise wurde sogar eine Lösung gefunden, die besser ist als die bisher beste bekannte. Einschränkend muß allerdings gesagt werden, dass die gegenwärtige Implementierung des Lösungsverfahrens die ϵ-Optimalität der berechneten Lösungen nicht ohne Einschränkung garantieren kann, da Rechenfehler, die aus der Maschinengenauigkeit resultieren, nicht speziell beachtet worden sind. Es ist allerdings mit einigem Aufwand möglich, die Implementierung entsprechend anzupassen. Da wir in erster Linie daran interessiert sind, ob das vorgestellte Verfahren überhaupt in der Lage ist, Probleme in der gewünschten Größenordnung mit akzeptablem numerischen Aufwand global zu lösen, ist diese Anpassung der Implementierung bisher nicht erfolgt.

Ein interessanter Aspekt bei der Anpassung des verwendeten allgemeinen Rechteck-Branch-and-Bound Ansatzes zu einer speziellen Lösungsmethode für das Packing Problem ist, dass die Entwicklung guter LP-Relaxierungen zur Berechnung von Schranken nicht entscheidend für die Verbesserung der numerischen Effizienz des Verfahrens ist. Die Aufteilungs- und insbesondere die zusätzlichen Manipulierungsstrategien haben einen wesentlich größeren Anteil an der Performanceverbesserung. Es ist möglich, dass auch bei anderen Problembeispielen, die aus Anwendungen resultieren, die Ausnutzung der speziellen Problemstruktur zur Ableitung solcher Strategien zur Manipulierung der Aufteilungsmengen im Hinblick auf das numerische Verhalten eines Branch-and-Bound Verfahrens erfolgreicher ist als die Entwicklung spezifischer, guter Schranken.

LITERATUR

[AK92] Faiz A. Al-Khayyal. Generalized Bilinear Programming: Part I. Models, Applications and Linear Programming Relaxation. *European Journal of Operational Research*, 60:306–314, 1992.

[AKLV95] Faiz A. Al-Khayyal, C. Larsen, and T. van Voorhis. A Relaxation Method for Nonconvex Quadratically Constrained Quadratic Programs. *Journal of Global Optimization*, 6:215–230, 1995.

[FV93] C.A. Floudas and V. Visweswaran. Primal-Relaxed Dual Global Optimization Approach. *Journal of Optimization Theory and Applications*, 78(2):187–225, 1993.

[GKL95] P. Gritzmann, V. Klee, and D. Larman. Largest j-Simplices in n-Polytopes. *Discrete Comput. Geom.*, 13:477–513, 1995.

[HR98] R. Horst and U. Raber. Convergent Outer Approximation Algorithms for Solving Unary Problems. *Journal of Global Optimization*, 13:123–149, 1998.

[HT96] R. Horst and H. Tuy. *Global Optimization: Deterministic Approaches*. Springer, Heidelberg, 3rd enlarged edition, 1996.

[NO97] K.J. Nurmela and P.R.J. Oestergard. Packing up to 50 Equal Circles in a Square. *Discrete Comput. Geom.*, 18:111–120, 1997.

[Ram93] M. Ramana. *An Algorithmic Analysis of Multiquadratic and Semidefinite Programming Problems*. PhD thesis, The John Hopkins University, Baltimore, 1993.

[Sle69] D. Slepan. The Content of some Extreme Simplices. *Pacific J. Math.*, 31:795–808, 1969.

[Som29] D.M.Y. Sommerville. *An Introduction to the Geometry of N Dimensions*. Methuen, London, 1929.

[ST92] H.D. Sherali and C.H. Tuncbilek. A Global Optimization Algorithm for Polynomial Programming Problems Using a Reformulation-Linearization Technique. *Journal of Global Optimization*, 2:101–112, 1992.