



ISNM
Internationale Schriftenreihe zur
Numerischen Mathematik
Lehrbuch

Herausgegeben von
K.-H. Hoffmann, München
H. D. Mittelmann, Tempe
J. Todd, Pasadena

Numerische Verfahren der nichtlinearen Optimierung

P. Spellucci

**Birkhäuser Verlag
Basel · Boston · Berlin**

Autor

Prof. Dr. Peter Spellucci
Fachbereich Mathematik
AG Numerische Mathematik
T.H. Darmstadt
Schlossgartenstrasse 7
D-W-6100 Darmstadt

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Spellucci, Peter:

Numerische Verfahren der nichtlinearen Optimierung / P.

Spellucci. – Basel ; Boston ; Berlin : Birkhäuser, 1993

(Internationale Schriftenreihe zur numerischen Mathematik :
Lehrbuch)

ISBN-13: 978-3-0348-7215-7 e-ISBN-13: 978-3-0348-7214-0

DOI: 10.1007/978-3-0348-7214-0

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Uebersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechts.

© 1993 Birkhäuser Verlag, Postfach 133, CH-4010 Basel, Schweiz

Camera-ready Vorlage erstellt vom Autor

Gedruckt auf säurefreiem Papier, hergestellt aus chlorfrei gebleichtem Zellstoff

ISBN-13: 978-3-0348-7215-7

9 8 7 6 5 4 3 2 1

Vorwort

Dieses Buch ist aus Vorlesungen entstanden, die der Verfasser in den letzten 15 Jahren an verschiedenen Hochschulen und vor verschiedenen Hörerkreisen gehalten hat.

Es wendet sich sowohl an den angewandten Mathematiker als auch an den mathematisch interessierten Anwender. Deshalb wurde der Text so gestaltet, daß er mit den Vorkenntnissen aus einem mathematischen Grundstudium zugänglich ist. Aus der Analysis werden nur die Differentialrechnung mehrerer Veränderlicher, insbesondere die Kettenregel, die TAYLOR-Formel mit Integralrestglied für Vektorfunktionen und der Hauptsatz über implizite Funktionen, sowie das Arbeiten mit den LANDAU-Symbolen \mathcal{O} und o vorausgesetzt. Aus der linearen Algebra benötigen wir Vertrautheit mit der Matrizenalgebra und aus der Matrizennumerik die Kenntnis der LR-Zerlegung, der QR-Zerlegung nach GRAM-SCHMIDT und HOUSEHOLDER, der CHOLESKY-Zerlegung positiv definiter Matrizen, der Singulärwert-Zerlegung und den Normbegriff (auf \mathbb{R}^n). Leser, die mit diesen Begriffen nicht vertraut sind, finden eine leicht verständliche Einführung dazu etwa in [222]. Ziel der Darstellung ist es, den Leser mit den heute praxisrelevanten numerischen Optimierungsverfahren vertraut zu machen. Je nach Intention des Lesers kann dabei die detaillierte Beschäftigung mit den Konvergenzeigenschaften der Verfahren im Vordergrund stehen oder auch übergangen werden. Ich bin jedoch der Meinung, daß aus der Kenntnis der Konvergenzbeweise auch ein vertieftes Verständnis für die praktischen Grenzen der Verfahren erwächst. Leider ist unser gegenwärtiger Kenntnisstand nicht ausreichend, um dem Anwender einen Optimierungscodes als niemals versagendes Werkzeug zur Verfügung stellen zu können.

Um dem Leser die Möglichkeit zu geben, sich zunächst einen Gesamtüberblick zu verschaffen und danach je nach seinen Bedürfnissen praktische oder theoretische Details zu erarbeiten, wurde der Text wie folgt strukturiert: mit “(*)” gekennzeichnete Abschnitte können ausgelassen werden, ohne den Gesamtüberblick zu gefährden. Mit $\ll \gg$ eingeschlossene Textabschnitte enthalten Detailanalysen, die zur Motivation eines Beweises oder zur Herleitung eines Verfahrens dienen. Der Inhalt dieser Passagen wird jeweils gesondert zusammengefaßt dargestellt. Der an den Details der Herleitung weniger Interessierte kann sie deshalb ohne weiteres übergangen. Für einen ersten Überblick bieten sich also Kapitel 1, Kapitel 2 und die Hauptteile von Kapitel 3 an, ohne die mit “(*)” oder “ $\ll \gg$ ” gekennzeichneten Teile und ohne die Lektüre der Beweise.

Um einem nicht vorinformierten Leser einen in sich geschlossenen Text anbieten zu können, habe ich die Grundtatsachen aus der Optimierungstheorie in \mathbb{R}^n in Kapitel 2 dargestellt, wobei ich mich auf das Notwendigste beschränkt habe. Insbesondere wurde auf eine Vertiefung der Darstellung konvexer Optimierungsaufgaben verzichtet, zumal dafür bereits mehrere schöne Lehrbücher existieren.

Bei der Darstellung der Verfahren wurde als Schwerpunkt die Herausarbeitung der Eigenschaften derjenigen Verfahren gewählt, die nach dem heutigen Kenntnisstand als zuverlässig und effizient gelten. Bewußt wurde darauf verzichtet, größtmögliche Allgemeinheit in den Aussagen anzustreben. Stattdessen habe ich da, wo es nach meiner Erfahrung für die praktische Umsetzung wichtig ist, auch Implementierungsfragen diskutiert.

Verfahren, die nur für spezielle Problemstellungen geeignet sind, wie die Schnittebenenverfahren, habe ich bewußt übergangen.

Da die lineare Optimierung meistens nicht Gegenstand des mathematischen Grundstudiums ist, habe ich mich entschlossen, die wichtigsten Verfahren dazu in einem Abschnitt kurz anzudiskutieren, da später an einigen Stellen auf diese Verfahren Bezug genommen wird. Zu diesem für die Anwendung äußerst wichtigen Spezialgebiet gibt es mehrere hervorragende Monographien, auf die im Text hingewiesen ist.

Die numerischen Verfahren der nichtlinearen Optimierung haben seit den Anfängen vor dreißig Jahren eine schnelle und sich immer mehr beschleunigende Entwicklung erfahren. Dies drückt sich auch in der Fülle der damit befaßten Veröffentlichungen aus. Auf eine allgemeine Literaturübersicht habe ich verzichtet und nur solche Arbeiten zitiert, deren Ergebnisse benutzt werden oder die weitergehende und meiner Meinung nach wichtige Ergebnisse enthalten.

Mehr als die Hälfte der zitierten Arbeiten stammt aus den Jahren 1980–1992, und so erscheinen hier viele Ergebnisse zum ersten Mal in Buchform. Die Abschnitte 3.4 und 3.6 enthalten zum Teil bisher nicht veröffentlichte Ergebnisse des Verfassers. Die in diesem Buch wiedergegebenen numerischen Resultate wurden vom Verfasser selbst auf verschiedenen Rechnern erarbeitet. Zur Erzeugung der Graphiken diente CA-DISSPLA auf einer HP 9000–345.

Zahlreiche meiner Studenten haben mir mit kritischen Bemerkungen bei der Verbesserung des Textes geholfen. Ihnen gilt mein Dank ebenso wie vier anonymen Referenten, deren Hinweise und Verbesserungsvorschläge die endgültige Form des Buches beeinflußt haben.

Dieses Buch wäre nicht ohne den unermüdlichen Einsatz und die großen Fertigkeiten von Frau Gudrun Schumm entstanden, die den gesamten Text in Latex geschrieben hat. Ihr gilt mein besonderer Dank.

Mein Dank gilt auch den Herausgebern für die Aufnahme des Werkes in die Serie ISNM und dem Birkhäuser Verlag für die angenehme Zusammenarbeit.

Darmstadt, im Juli 1992

P. Spellucci

Inhaltsverzeichnis

1 Einführung

1.1 Auftreten von Optimierungsproblemen in der Praxis	1
1.2 Das Modell des allgemeinen NLO-Problems	10
1.3 Geometrische Veranschaulichung einfacher Optimierungsprobleme	12

2 Theorie

2.1 Extremalkriterien für differenzierbare Probleme	15
(*) Anhang 1 zu 2.1: Alternativsätze für Systeme linearer Ungleichungen	38
(*) Anhang 2 zu 2.1: Alternative Herleitung der Multiplikator-Regel	42
2.2 Lagrange-Dualität I	46
2.3 Konvexe Optimierungsaufgaben	50
2.4 Lagrange-Dualität II	69
2.5 (*) Sensitivitäts- und Stabilitätsbetrachtungen	77

3 Verfahren

3.0 Übersicht	87
3.1 Verfahren der unrestringierten Minimierung	91
3.1.1 Schrittweisenverfahren	97
3.1.2 Verfahren zur Richtungsbestimmung	111
3.1.2.1 (*) Die Methode des koordinatenweisen Abstiegs. Das SOR-Newton-Verfahren	114
3.1.2.2 Verfahren mit gradientenbezogenen Richtungen	119
3.1.2.3 Newton- und Newton-ähnliche Verfahren	130
3.1.2.4 Quasi-Newton-Verfahren, insbesondere das BFGS-Verfahren	133
3.1.2.5 Verfahren konjugierter Richtungen. Das cg-Verfahren	156
3.1.2.6 (*) Weitere Quasi-Newton-Verfahren	172
3.1.2.7 (*) Verfahren, die die notwendigen Bedingungen zweiter Ordnung erfüllen	178
3.1.3 (*) Die Methode der Vertrauensbereiche	185

3.1.4	Spezielle Verfahren für Ausgleichsprobleme	192
3.1.4.1	Lineare Ausgleichsrechnung	193
3.1.4.2	Das Gauß–Newton–Verfahren	195
3.1.4.3	Schrittweitenverfahren für Ausgleichsaufgaben	201
3.1.4.4	(*) Das Verfahren von Levenberg und Marquardt in der Implementierung von J.J. Moré	202
3.1.4.5	(*) Ein spezielles Verfahren für die Ausgleichung mittels rationaler Funktionen	217
3.1.4.6	(*) Verfahren für Ausgleichsprobleme mit separierbaren Variablen	223
3.1.4.7	(*) Orthogonale Regression	225
3.1.5	Ergänzungen	233
3.1.5.1	Konvergenztheorie	233
3.1.5.2	Das Problem der Skalierung	235
3.1.5.3	Numerische Differentiation	238
3.1.5.4	Grenzgenauigkeit und Abbruchkriterien	240
3.2	Verfahren zur linearen Optimierung	242
3.2.1	Normalform einer LO–Aufgabe und Transformation auf Normalform	242
3.2.2	Struktur der zulässigen Menge und der Lösungsmenge der LO–Aufgabe	244
3.2.3	Das Simplexverfahren	249
3.2.4	Ermittlung einer zulässigen Ausgangsecke	255
3.2.5	Simplex–Verfahren mit LR–Zerlegung	257
3.2.6	Einiges über duale lineare Programme und Anwendungen	259
3.2.7	Die algebraische Berechnungskomplexität der LO–Aufgabe. Die Verfahren von Barnes, Khachiyan und Karmarkar	261
3.3	Verfahren zur quadratischen Optimierung	285
3.3.1	Ein primales Verfahren vom Projektionstyp	285
3.3.2	Das duale Verfahren von Goldfarb und Idnani zur Lösung streng konvexer quadratischer Optimierungsaufgaben	293
3.3.3	(*) Ein Verfahren für lineare Ausgleichsaufgaben mit linearen Restriktionen	304
3.3.4	(*) Verfahren zur Lösung quadratischer Optimierungsprobleme mit Mehrfachinaktivierung	312
3.3.5	(*) Weitere Verfahren zur Lösung von konvexen QP–Problemen	318
3.3.6	(*) Ein polynomiales Verfahren für konvexe quadratische Optimierungsprobleme	319
3.3.7	(*) Das indefinite quadratische Optimierungsproblem	324
3.4	Projektions–und Reduktionsverfahren für NLO	331
3.4.1	Allgemeine Konvergenztheorie eines primalen Abstiegsverfahrens	331
3.4.2	Konstruktion eines zulässigen Kurvenbogens	347

Inhaltsverzeichnis

3.4.3	Ein Schrittweisenverfahren für (nichtlinear) restringierte Optimierungsprobleme	351
3.4.4	Q-superlinear konvergente Varianten von GGPRV	357
3.4.5	Verfahren vom Typ der reduzierten Gradienten	361
3.4.6	(*) Ein Reduktionsverfahren mit Mehrfachinaktivierung	375
3.4.7	(*) Das Projektionsverfahren von Bertsekas	383
3.4.8	Ergänzende Bemerkungen	389
3.5	Penalty- und Multiplikator-Verfahren	393
3.5.1	Klassische Penalty-Verfahren	393
3.5.2	Die Multiplikator-Methode von Hestenes und Powell für gleichungsrestringierte Probleme	410
3.5.3	Die Multiplikator-Methode von Rockafellar	428
3.5.4	(*) Exakte differenzierbare Penalty-Funktionen	440
3.5.5	Weitere Hinweise und Bemerkungen	451
3.6	Die Methode der sequentiellen quadratischen Minimierung	455
3.6.1	(*) Exakte nichtdifferenzierbare Penalty-Funktionen	455
3.6.2	Die Methode der sequentiellen quadratischen Optimierung	474
3.6.2.1	Allgemeine Vorüberlegungen	474
3.6.2.2	Die SQP-Methode für konvexe Optimierungsaufgaben	483
3.6.2.3	Die SQP-Methode für nichtkonvexe NLO-Probleme. Regularisierungstechniken für inkompatible QP-Probleme	487
3.6.2.4	(*) Die Konstruktion der Matrizenfolge $\{A_k\}$	507
3.6.2.5	(*) Der Maratos-Effekt	509
3.6.2.6	(*) Zur Schrittweitenbestimmung	513
3.6.2.7	(*) Zur Konvergenzgeschwindigkeit der SQP-Methode	514
3.6.2.8	Weitere Hinweise und Bemerkungen	523
3.7	Hinweise zur Praxis von NLO	528
3.7.1	Problemformulierung	528
3.7.2	Skalierung	529
3.7.3	Numerische Differentiation	530
3.7.4	Grenzgenauigkeit und Abbruchkriterien	530
Anhang 1	: Übersicht über verfügbare Software	532
Anhang 2	: Übersicht über themenspezifische Zeitschriften und Buchreihen	533
Anhang 3	: Notationen	534
Literaturverzeichnis	539
Sachverzeichnis	553