

Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben

2., überarbeitete Auflage
Von Prof. Dr. rer. nat. Peter Kosmol
Universität Kiel



Springer Fachmedien Wiesbaden GmbH 1993

Prof. Dr. rer. nat. Peter Kosmol

Geboren 1942 in Ratiborhammer/Schlesien. Von 1960 bis 1965 Studium der Mathematik und Physik mit anschließender Assistententätigkeit bis 1967 an der Universität Wroclaw (Breslau). 1970 Promotion, 1974 Habilitation und 1979 Ernennung zum apl. Professor in Kiel. Seit 1971 Betreuung des Arbeitsbereiches Optimierungs- und Approximationstheorie einschließlich der dazugehörigen numerischen Verfahren am Mathematischen Seminar der Universität Kiel.

ISBN 978-3-519-12085-8

ISBN 978-3-663-12239-5 (eBook)

DOI 10.1007/978-3-663-12239-5

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Kosmol, Peter:

Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben / von Peter Kosmol. – 2., überarb. Aufl. – Stuttgart : Teubner, 1993
(Teubner Studienbücher : Mathematik)
ISBN 978-3-519-12085-8

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt besonders für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

© Springer Fachmedien Wiesbaden 1989

Ursprünglich erschienen bei B. G. Teubner Stuttgart in 1989

Gesamtherstellung: Druckhaus Beltz, Hemsbach/Bergstraße
Umschlaggestaltung: M. Koch, Reutlingen

Vorwort

Das Ziel der hier vorliegenden Abhandlung ist eine einfache einheitliche Darstellung der Konvergenzbeweise für numerische Verfahren nichtlinearer Optimierungsaufgaben und der damit verbundenen nichtlinearen Gleichungen. Im wesentlichen werden Verfahren betrachtet, die auf der Idee des Gradienten- und Newton-Verfahrens beruhen. Es wurde dabei nach möglichst einfachen Beweisen für die Konvergenz und die Konvergenzgeschwindigkeit von Algorithmen für Aufgaben in dem Euklidischen Raum \mathbb{R}^n gesucht. Es hat sich aber herausgestellt, daß gerade die einfachen Beweise nicht die spezielle Struktur des \mathbb{R}^n benutzen und in allgemeinen normierten Räumen gültig sind. Das zentrale Beweismittel ist hier der Mittelwertsatz der Differentialrechnung in der Integralform, der auch in Banachräumen gilt. Wir setzen den Begriff eines Vektorraumes (linearen Raumes) als bekannt voraus und wollen mit der Definition eines normierten Raumes die Einführung beginnen. Die Auswahl der Eigenschaften eines normierten Raumes wird sich an der Tatsache orientieren, daß die Numerik in \mathbb{R}^n im Vordergrund stehen soll. Unter einem Vektorraum wird im gesamten Text ein Vektorraum über dem Körper der reellen Zahlen verstanden. Es wird empfohlen sofort mit dem eigentlichen Text (ab Kapitel 1) anzufangen und die Einführung nur als Nachschlagewerk zu benutzen. Denn die Einführung ist an einigen Stellen als Ergänzung gedacht. So werden z.B. im Abschnitt 0.8.6 uniform konvexe Funktionen eingeführt, die auch für die Numerik in \mathbb{R}^n wichtig sind. Jedoch zur vollen Entfaltung kommt dieser Begriff erst im Rahmen der normierten Räume und die allgemeine Sicht kann auch zum besseren Verständnis führen. Weiter sind die im Text behandelten Optimierungsaufgaben nichtrestringiert und in der Einführung wird gezeigt, wie restringierte Aufgaben auf nichtrestringierte zurückgeführt werden können (s. 0.9). Die Stabilitätssätze (s. 0.3.2) weisen z.B. einen Weg, wie man nichtdifferenzierbare Aufgaben durch eine Folge differenzierbarer ersetzen kann.

Viele in den Anwendungen vorkommende Aufgaben sind nichtdifferenzierbar, aber sie lassen sich oft mit der dazugehörigen Theorie auf das Lösen von nichtlinearen Gleichungen zurückführen, so z.B. die Aufgaben der Čebyšev- und L_1 -Approximation s. [G-G], [G-S], [H-Z] und [K4].

Dieser Text wendet sich an Leser, die eine zweisemestrige Mathematik-Vorlesung im Rahmen eines Studiums der Mathematik, Informatik oder der Naturwissenschaften gehört haben.

Das Kapitel 1, 2 und 12 beruhen auf einer früheren Ausarbeitung einer meiner Vorlesungen von Bärbel Schröder.

Das mühevoll Schreiben haben Anne-Katrin Främbs, Inken Höhrmann und Sabine Thielk übernommen und mit großer Sorgfalt durchgeführt. Ihnen sei herzlich gedankt.

Von den Hörern meiner Vorlesungen habe ich zahlreiche Hinweise und Korrekturen zu der früheren Version des Skriptes erhalten. Reinhard Lohse hat mir durch sein sorgfältiges Studieren des Textes sehr geholfen

Ihnen allen sei herzlich gedankt.

Kiel, im Januar 1989

Peter Kosmol

Vorwort zur zweiten Auflage

Die wesentlichen Änderungen dieser Neuauflage bestehen aus neuen Algorithmen, die in den Abschnitten 2.5, 10.8 und Kapitel 13 zu finden sind. Ich habe die Hoffnung, die meisten Druckfehler jetzt beseitigt zu haben. Mein herzlicher Dank gilt meinen Lesern, die mir dabei geholfen haben, und vor allem Reinhard Lohse, der mich stets unterstützt hat.

Den Mitarbeitern des Verlages danke ich für die gute Zusammenarbeit und für das Entgegenkommen bei der Herstellung dieser Auflage.

Kiel, im Februar 1993

Peter Kosmol

Inhaltsverzeichnis

| | | |
|----------|--|----|
| 0 | Einführung | |
| 0.1 | Normierte Räume | 1 |
| 0.2 | Prä-Hilbert-Räume | 3 |
| 0.3 | Konvexe Funktionen | 5 |
| 0.3.1 | Jensen'sche Ungleichung | 6 |
| 0.3.2 | Äquivalenz der Normen und Stetigkeit konvexer Funktionen in \mathbb{R}^n . Stabilitätssätze. | 7 |
| 0.3.3 | Das Schnittebenenverfahren | 12 |
| 0.4 | Richtungsableitung und Fréchet-Differenzierbarkeit | 13 |
| 0.5 | Differentialrechnung in \mathbb{R}^n . Matrix- und Operator-schreibweise. | 16 |
| 0.6 | Mittelwertsatz in der Integralform | 17 |
| 0.7 | Matrizen | 20 |
| 0.7.1 | Eigenwerte und positiv definite Matrizen | 20 |
| 0.7.2 | Spur einer Matrix | 24 |
| 0.7.3 | Frobenius-Norm | 25 |
| 0.7.4 | Neumann-Lemma | 26 |
| 0.7.5 | Störungslemma | 26 |
| 0.7.6 | Lösung linearer Gleichungen - Cholesky-Zerlegung | 27 |
| 0.8 | Elemente der Optimierungstheorie | 28 |
| 0.8.1 | Existenz von Minimallösungen. Der Satz von Weierstraß | 29 |
| 0.8.2 | Eindeutige Lösbarkeit von Optimierungsaufgaben | 29 |
| 0.8.3 | Notwendige Optimalitätsbedingungen | 30 |
| 0.8.4 | Hinreichende Optimalitätsbedingungen. Charakterisierungssatz der konvexen Optimierung | 31 |
| 0.8.5 | Approximation in Prä-Hilbert-Räumen | 33 |
| 0.8.6 | Uniform konvexe Funktionen. Starke Lösbarkeit. | 34 |
| 0.9 | Restringierte Optimierungsaufgaben. Lagrange- und Penalty-Methode | 42 |
| 0.9.1 | Lagrange-Methode | 43 |

| | | |
|----------|--|-----------|
| 0.9.2 | Lagrange-Lemma bei Gleichungen und Ungleichungen | 48 |
| 0.9.3 | Zurückführung von Ungleichungsrestriktionen auf Gleichungsrestriktionen | 50 |
| 0.9.4 | Penalty-Methode | 50 |
| 0.9.5 | Numerische Behandlung restringierter Optimierungsaufgaben | 51 |
| 1 | Eindimensionale Bestimmung von Nullstellen | 52 |
| 1.1 | Bisektionsverfahren | 52 |
| 1.2 | Newton-Verfahren | 52 |
| 1.3 | Regula-Falsi | 53 |
| 2 | Konvergenzordnung. Eindimensionale Minimierung | 55 |
| 2.1 | Q-Konvergenz für Folgen (Quotientenkriterium) | 55 |
| 2.2 | R-Konvergenzordnung (Wurzelkriterium) | 55 |
| 2.3 | Algorithmen | 56 |
| 2.4 | Konvergenzordnung für Algorithmen | 57 |
| 2.5 | Schnell und global konvergente Verfahren | 57 |
| 2.6 | Lemma von Dennis-Moré | 64 |
| 2.7 | Eindimensionale Minimierung | 64 |
| 2.8 | Verfahren des goldenen Schnitts | 64 |
| 2.9 | DSCP-Verfahren | 65 |
| 3 | Newton-Verfahren und Newton-ähnliche Verfahren | 67 |
| 3.1 | Newton-Verfahren | 67 |
| 3.2 | Charakterisierung der Q-superlinearen Konvergenz. Newton-ähnliche Verfahren. | 70 |
| 3.3 | Charakterisierung der quadratischen Konvergenz | 73 |
| 3.4 | Q-superlineare Konvergenz bei Matrix-Richtungen | 75 |
| 3.5 | Einfluß der Störungen beim Newton-Verfahren | 76 |
| 3.6 | Das Newton-Verfahren mit Differenzenquotienten | 78 |
| 3.7 | Gauß-Newton-Verfahren | 80 |

| | | |
|----------|---|------------|
| 4 | Verallgemeinerte Gradientenverfahren | 84 |
| 4.1 | Einführung | 84 |
| 4.2 | Einige Schrittweitenregeln (Schrittweiten-Algorithmien) | 86 |
| 4.2.1 | Minimierungsregel (Regel der optimalen Schrittweite) (M) | 86 |
| 4.2.2 | Limitierte Minimierungsregel (limitierte optimale Schrittweite) (LM) | 87 |
| 4.2.3 | Curry Minimierungsregel (C) | 87 |
| 4.2.4 | Armijoregel (AR) | 88 |
| 4.2.5 | Armijoregel mit Aufweitung (ARA) | 89 |
| 4.2.6 | Goldsteinregel (G) | 90 |
| 4.2.7 | Powell-Wolfe-Regel (PW) | 90 |
| 4.3 | Realisierbarkeit der Schrittweitenregeln | 91 |
| 4.3.1 | Minimierungsregel und modifizierte Minimierungsregel | 91 |
| 4.3.2 | Limitierte Minimierungsregel | 91 |
| 4.3.3 | Armijoregel | 91 |
| 4.3.4 | Goldsteinregel und Armijoregel mit Aufweitung | 92 |
| 4.3.5 | Powell-Wolfe-Regel | 92 |
| 5 | Klassifikation der Schrittweitenregeln | 94 |
| 5.0 | Vorbetrachtungen | 94 |
| 5.1 | Effiziente Schrittweitenregeln | 94 |
| 5.2 | Konvergenzverhalten bei effizienten Schrittweitenregeln | 98 |
| 6 | Konvergenzbetrachtungen für verallgemeinerte Gradientenverfahren | 100 |
| 6.0 | Vorbemerkungen | 100 |
| 6.1 | Konvergenz verallgemeinerter Gradientenverfahren | 101 |
| 6.2 | R-lineare Konvergenz bei stark konvexen Funktionen | 102 |
| 6.3 | Lineare Konvergenz verallgemeinerter Gradientenverfahren | 104 |
| 6.4 | Spacer step | 107 |
| 6.5 | Eigenschaft (G) | 108 |

| | | |
|-----------|---|------------|
| 7 | Konvergenzverhalten von verallgemeinerten Gradientenverfahren bei quadratischen Funktionen | 111 |
| 7.1 | Kantorovich-Ungleichung | 111 |
| 7.2 | Konvergenzrate bei quadratischen Funktionen | 113 |
| 7.3 | Beschleunigung durch Maßstabsänderung | 116 |
| 8 | Global und Q-superlinear konvergente Abstiegsverfahren | 118 |
| 8.1 | Stark konvexe Optimierungsaufgaben | 118 |
| 8.2 | Lokale Minimallösungen | 123 |
| 8.3 | Superlineare Konvergenz bei der Curry-Regel | 124 |
| 8.4 | Globale Varianten der Nullstellenbestimmung | 125 |
| 9 | Global konvergente Modifikationen des Newton-Verfahrens | 130 |
| 9.1 | Gedämpftes Newton-Verfahren für konvexe Funktionen | 130 |
| 9.2 | Gedämpftes diskretisiertes Newton-Verfahren | 131 |
| 9.3 | Gedämpftes Newton-Verfahren für Gleichungen | 132 |
| 9.4 | Gedämpftes Gauß-Newton-Verfahren | 134 |
| 9.5 | Gedämpftes Newton-Verfahren für nichtkonvexe Funktionen | 136 |
| 9.6 | Schrittweitenabhängige Suchrichtungen | 138 |
| 9.7 | Positiv definite Störungen der Hesse-Matrix | 141 |
| 9.8 | Verfahren von Levenberg/Marquardt | 142 |
| 10 | Quasi-Newton-Verfahren | 145 |
| 10.1 | Quasi-Newton-Gleichung und Aufdatierungsmatrizen | 145 |
| 10.2 | Sekantenverfahren minimaler Änderung und ihre Geometrie | 149 |
| 10.3 | Q-superlineare Konvergenz linear konvergenter Sekantenverfahren minimaler Änderung | 152 |
| 10.4 | Symmetrische Aufdatierungen. PSB-Formel. | 154 |
| 10.5 | Quasi-Newton-Methoden für schwach-besetzte Matrizen | 156 |
| 10.6 | Lokale und Q-superlineare Konvergenz von Sekanten- verfahren minimaler Änderung | 162 |

| | | |
|-----------|--|------------|
| 10.7 | Variable Sekantenverfahren minimaler Änderung | 165 |
| 10.8 | Global konvergente Methoden für Gleichungen. MAN-Verfahren. | 170 |
| 11 | Sekantenverfahren bei nichtrestringierter Minimierung | 183 |
| 11.0 | Modifizierte Sekantenverfahren | 183 |
| 11.1 | Positive Definitheit der DFP- und BFGS-Aufdatierungen | 187 |
| 11.2 | Globale und lineare Konvergenz des BFGS-Verfahrens | 189 |
| 11.3 | Ordnungsmonotonie der BFGS- und der DFP-Formel | 193 |
| 11.4 | Q-superlineare Konvergenz des BFGS-Verfahrens | 197 |
| 11.5 | Beispiele für global konvergente Modifikationen der Sekantenverfahren | 199 |
| 11.6 | Die Verfahren der Broyden-Klasse für quadratische Funktionen | 203 |
| 12 | Verfahren der konjugierten Gradienten | 205 |
| 12.1 | Konjugierte Richtungen | 205 |
| 12.2 | Verfahren der konjugierten Gradienten zur Lösung linearer Gleichungssysteme | 207 |
| 12.3 | Verfahren der konjugierten Gradienten für nicht- Minimierungsaufgaben | 210 |
| 13 | Sekantenverfahren für lineare Gleichungen | 214 |
| | Literaturangaben | 221 |
| | Zeichenliste | 228 |
| | Algorithmenliste | 228 |
| | Namens- und Sachverzeichnis | 229 |