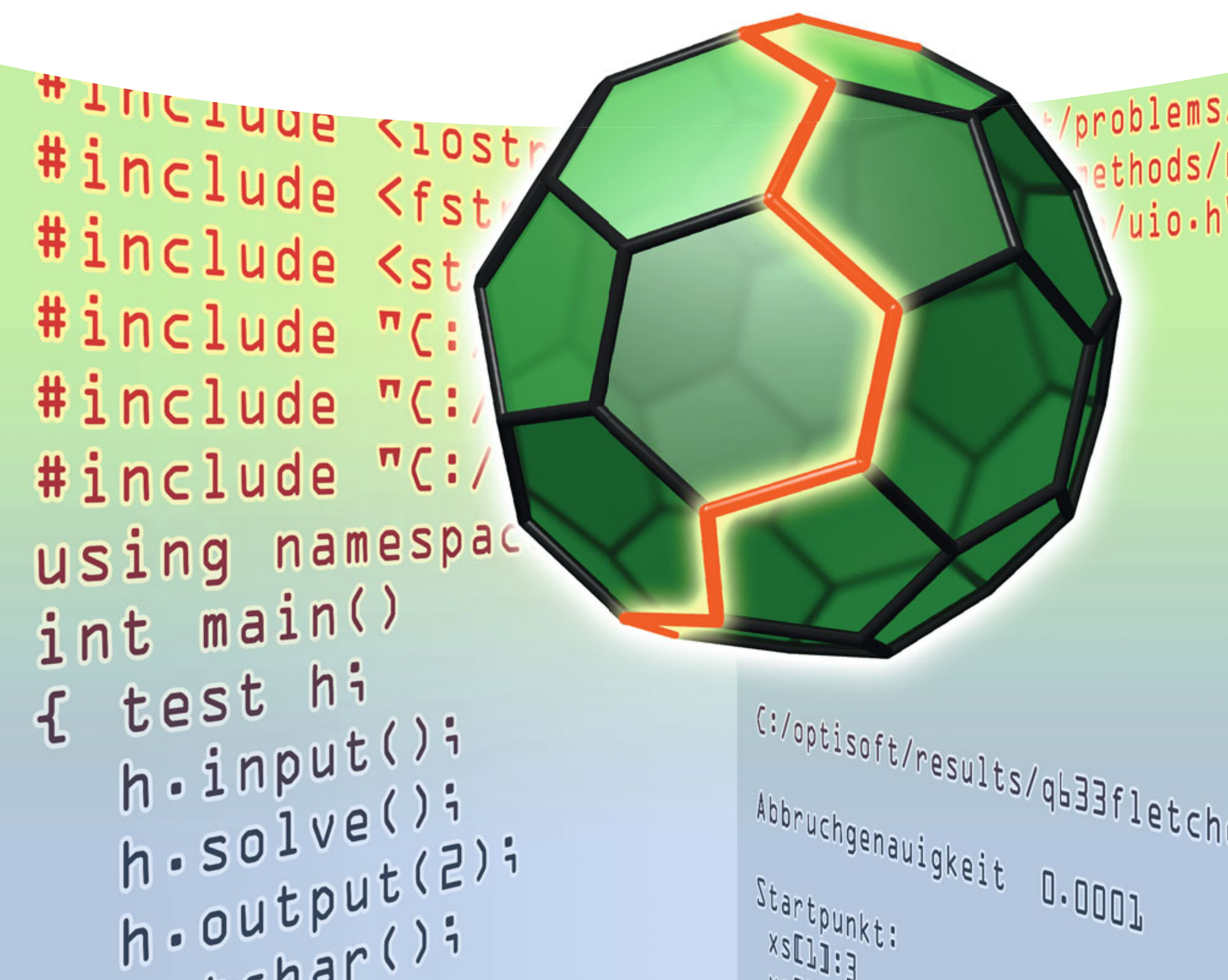
WILEY-VCH

Claus Richter

Optimierung in C++

Grundlagen und Algorithmen



Claus Richter

Optimierung in C++

Claus Richter

Optimierung in C++

Grundlagen und Algorithmen

WILEY-VCH
Verlag GmbH & Co. KGaA

Autor

Claus Richter
Gerhart-Hauptmann-Str. 1
01219 Dresden
Deutschland

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Umschlaggestaltung Formgeber, Mannheim, Deutschland
Satz le-tex publishing services GmbH, Leipzig, Deutschland

Print ISBN 978-3-527-34107-8
ePDF ISBN 978-3-527-80079-7
ePub ISBN 978-3-527-80080-3
Mobi ISBN 978-3-527-80081-0

Gedruckt auf säurefreiem Papier.

Für meine liebe Frau Hannelore

Inhaltsverzeichnis

Vorwort

Bücher zur Implementierung numerischer Verfahren der Optimierung sind seit vielen Jahren gefragt. Die Behandlung mathematisch-naturwissenschaftlicher, technischer und ökonomischer Fragestellungen erfordert in wachsendem Umfang die Lösung linearer oder nichtlinearer Optimierungsaufgaben. Gegenüber den ersten Bemühungen in den 40er- und 50er-Jahren haben sich hierfür die Voraussetzungen auf dem Gebiet der Informatik wesentlich verbessert. Nicht nur Rechenzeit und Speicherplatz haben eine andere Bewertung erfahren, auch Programmierparadigmen und die Nutzung von Dialogmöglichkeiten haben sich geändert. Dieser Entwicklung folgend, werden im vorliegenden Buch Probleme und Lösungsverfahren als Klassen der objektorientierten Programmierung aufgefasst. Die Formulierung der zu lösenden Optimierungsaufgabe und die Auswahl der Lösungsmethode erfolgt im Dialog, die Ergebnisse der Berechnung werden automatisch gespeichert. Im Unterschied zu komplexen Systemen, wie Matlab sind die einzelnen Routinen modifizierbar und separat nutzbar. Seit den Arbeiten von Kantorovich [1] und Dantzig [2] zum Simplexverfahren hat auch die Entwicklung effektiver numerischer Verfahren der Optimierung eine stürmische Entwicklung genommen. Ihre theoretische Begründung und sachgerechte Implementierung stellt inzwischen einen eigenständigen Problemkreis dar, welcher als Numerik der Optimierung (in englischer Sprache als „Computational Mathematical Programming“, in russischer Sprache als „Vycislitelnye metody programmirovanija“) bezeichnet wird. Die Aneignung der auf diesem Gebiet vorhandenen Erkenntnisse, noch mehr aber das Erleben des Zusammenhangs von beschriebenem Algorithmus, umgesetztem Programm und bereitgestellter Nutzeroberfläche werden zum Bedürfnis des an der Optimierung interessierten Praktikers. Gegenstand des Buches sind deshalb nicht in erster Linie theoretische Grundlagen, sondern Fragen der praktischen Realisierung der Verfahren mit modernen Mitteln der Informatik. Es soll einen Einstieg in die Behandlung von Optimierungsaufgaben auf Computern ermöglichen.

Für praktische Hilfeleistungen beim Zustandekommen des Buches bin ich Klaus Schönefeld zu Dank verpflichtet. In gleicher Weise danke ich Thomas Cassebaum für die Möglichkeit, die von ihm bereitgestellte Entwicklungsumgebung „SmallCpp“ nutzen zu können und in C++-Fragen in ihm jederzeit einen guten Gesprächspartner gefunden zu haben. Ermutigende Worte und gute Ratschläge

vieler Kollegen, insbesondere von Diethard Pallaschke, Oleg Burdakov, Manfred Grauer und Gerd Langensiepen, haben den Entstehungsprozess befördert. Dem Wiley-Verlag danke ich für die Möglichkeit, die Ergebnisse meiner Überlegungen zu publizieren. Schließlich möchte ich meiner Frau Hannelore für das Verständnis danken, mit dem sie die Belastung mitgetragen hat, welche dem Autor aus dem Schreiben eines Buches erwächst. Die Publikation von Algorithmen und Programmen schließt zu erwartende Kritiken und Hinweise von vornherein ein. Sie werden von mir sorgfältig berücksichtigt und in die Aufbereitung weiterer Programmversionen eingearbeitet.

Claus Richter

1
Einleitung

1.1
Das lineare und das nichtlineare Optimierungsproblem

Im vorliegenden Buch werden Optimierungsaufgaben betrachtet, die dadurch charakterisiert sind, dass eine lineare oder nichtlineare Zielfunktion f unter linearen oder nichtlinearen Ungleichungsnebenbedingungen minimiert wird, d. h.

$$f(x) = \min! \quad \text{bei} \quad x \in G = \{x : g_i(x) \leq 0 \quad i \in I\} \ , \tag{1.1}$$

wobei I

$$I = \{i : i = 1, \dots , m\}$$

die Indexmenge der Ungeichungsrestriktionen bezeichnet. Gleichungsrestriktionen werden der Übersichtlichkeit halber zunächst weggelassen. An geeigneten Stellen werden sie zusätzlich berücksichtigt.

1.2
Definitionen und Bezeichnungen

Für die weiteren Überlegungen benötigen wir folgende Bezeichnungen:

- n-dimensionaler Euklidischer Raum: R^n,
- Menge der reellen Zahlen: R,
- nichtnegativer Orthant des n-dimensionalen Euklidischen Raumes: R^n_+,
- Euklidische Norm: $\| x \|_2 = (x^\mathrm{T} x)^{1/2}$,
- Betragssummennorm: $\| x \|_1 = \sum_{i=1}^{m} |x_i|$,
- (m, n)-Matrix A: rechteckiges Zahlenschema $A = (a_{i,j})$ von $m * n$ Zahlen, angeordnet in m Zeilen und n Spalten,
- quadratische Matrix: (m, n)-Matrix A mit $m = n$,
- Diagonalmatrix A: quadratische Matrix A mit $a_{ij} = 0$ für $i \neq j$ und $a_{ii} \neq 0$,
- Einheitsmatrix I: Diagonalmatrix A mit $a_{ii} = 1$,
- obere Dreiecksmatrix A: quadratische Matrix A mit $a_{ij} = 0, i > j$,
- untere Dreiecksmatrix A: quadratische Matrix A mit $a_{ij} = 0, i < j$,

Optimierung in C++, 1. Auflage. Claus Richter.

- positiv definite Matrix A: quadratische Matrix A mit $x^{\mathrm{T}}Ax > 0$ für alle $x \neq 0$,
- symmetrische Matrix A: quadratische Matrix A mit $a_{ij} = a_{ji}$,
- transponierte Matrix A^{T} zu A: Matrix A^{T} mit $A^{\mathrm{T}} = (a_{ji})$,
- inverse Matrix A^{-1} zur Matrix A: Matrix mit der Eigenschaft $A * A^{-1} = I$,
- nichtsinguläre Matrix A: die inverse Matrix A^{-1} zu A existiert,
- orthogonale Matrix A: Matrix mit der Eigenschaft $A^{\mathrm{T}} = A^{-1}$,
- transponierter Vektor: $x^{\mathrm{T}} = (x_1, \ldots, x_n)$,
- Gradient einer Funktion $f : R^n \to R$

$$\nabla f(x) := \left(\frac{\partial}{\partial x_1} f(x), \ldots, \frac{\partial}{\partial x_n} f(x) \right)^{\mathrm{T}} ,$$

- Hesse-Matrix einer Funktion $f : R^n \to R$

$$\nabla^2 f(x)_{ij} := \frac{\partial^2}{\partial x_i \partial x_j} f(x) \quad i, j = 1, \ldots, n ,$$

- Lagrange-Funktion für die Aufgabe (1.1)

$$L(x, u) = f(x) + \sum_{i=1}^{m} u_i g_i(x) ,$$

- Ableitung der Lagrange-Funktion nach den Komponenten des 1. Arguments

$$\nabla_1 L(x, u) = \nabla f(x) + \sum_{i=1}^{m} u_i \nabla g_i(x) ,$$

- zweite Ableitung der Lagrange-Funktion nach den Komponenten des 1. Arguments

$$\nabla_{11} L(x, u) = \nabla_{11} f(x) + \sum_{i=1}^{m} u_i \nabla_{11} g_i(x) ,$$

- Indexmenge $I(x)$ der in x aktiven Restriktionen

$$I(x) = \{ i : g_i(x) = 0,\ i \in \{1, \ldots, m\} \} ,$$

- Vektor, dessen Komponenten alle gleich 1 sind: $e = (1, \ldots, 1)^{\mathrm{T}}$,
- i-ter Einheitsvektor: $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^{\mathrm{T}}$,
- die Menge $G^0 := \{x : g_i(x) < 0, i = 1, \ldots, m\}$.

1.3
Spezialfälle linearer und nichtlinearer Optimierungsaufgaben

Besitzen Zielfunktion f und der zulässige Bereich G bzw. Nebenbedingungen g_i und g_j eine spezielle Gestalt, so können zur Lösung von (1.1) spezielle Verfahren herangezogen werden. Für die Zielfunktion f sind folgende Strukturen interessant:

1. Allgemeine nichtlineare Zielfunktion $f(x)$.
2. Lineare Zielfunktion $f(x) = c^T x$.
3. Quadratische Zielfunktion $f(x) = \frac{1}{2} x^T C x + d^T x$.
4. Quadratsumme (Regression)

$$f(x) = \sum_{i=1}^{m} (y_i - f(x, t_i))^2 \ , \quad y_i \text{ – Messwerte zum Messpunkt } t_i \ .$$

5. Maximum von Funktionen $f(x) = \max f_j(x) \quad (j = 1, \ldots, l)$.

In Bezug auf die Nebenbedingungen N sind folgende Situationen typisch:

1. Allgemeine nichtlineare Nebenbedingungen.
2. Lineare Nebenbedingungen $g_i(x) = a_i^T x + b_i \quad (i = 1, \ldots, m)$.
3. Keine Nebenbedingungen $G = R^n$.

In den folgenden Kapiteln werden spezielle Kombinationen von Zielfunktion und Nebenbedingungen eine besondere Rolle spielen:

- lineare Optimierung (L): $f2 + N2$,
- quadratische Optimierung (Q): $f3 + N2$,
- allgemeine nichtlineare Optimierungsaufgabe (C): $f1 + N1$,
- unbeschränkte Minimierung (U): $f1 + N3$,
- Regressionsprobleme (P): $f4 + N3$, $f4 + N1$.

Die Spezifikationen L, Q, C, U und P werden in der Charakterisierung der implementierten Beispiele im Programmsystem „Optisoft" verwendet. Über die dargestellten Kombinationen von Zielfunktion und Nebenbedingungen hinaus spielen Aufgaben der nichtglatten Optimierung eine besondere Rolle. Diese finden im vorliegenden Buch keine Beachtung. Gleiches gilt auch für Optimierungsaufgaben mit sehr vielen Variablen: $n > 100$, sofern sie nicht als Teilprobleme zur Lösung von (1.1) auftreten.

Obwohl die spezifische Gestalt von Zielfunktion und Nebenbedingungen interessant ist, wie etwa in der geometrischen Optimierung

$$f(x) = \sum_{k=1}^{r} c_{k(o)} \prod_{i=1}^{n} x_i a_{ik}^0$$
$$g_j(x) = \sum_{k=1}^{r} c_{k(j)} \prod_{i=1}^{n} x_i a_{ij} \quad (j = 1, \ldots, m) \ ,$$

wird diese nicht explizit berücksichtigt.

In der Betrachtung von Optimierungsverfahren gehen wir von dem Grundmodell (1.1) aus. Für Least-Square-Probleme in Differenzialgleichungsmodellen und bei Strukturoptimierungsproblemen liegen spezielle Aufgaben zugrunde. Diese werden in den folgenden Kapiteln näher erläutert.

1.4 Anwendungen

Nichtlineare Optimierungsprobleme spielen in vielen Anwendungsbereichen eine wichtige Rolle, z. B. in der

- Luft- und Raumfahrt (Steuerung, Konstruktion),
- Mechanik (Optimierung mechanischer Strukturen, z. B. von Tragwerken),
- Elektrotechnik (Transformatorkonstruktion),
- Chemie (Gleichgewichtsprobleme),
- Medizin, Soziologie (Statistische Probleme),
- Betriebswirtschaft (Planungsmodelle),
- Physik (Kernforschung),
- Energiewesen (Energieverteilung).

Typische Anwendungsbeispiele finden sich in den Büchern von Bracken und McCormick [3] oder Beightler und Phillips [4]. Einige mathematische Fragestellungen, welche bei der Lösung praktischer Probleme auf Optimierungsverfahren zurückgreifen, werden im Buch näher betrachtet:

1.4.1 Strukturoptimierung

Die Strukturoptimierung wird schon seit einigen Jahren in der computergestützten Konstruktion eingesetzt. In der zugrundeliegenden Aufgabenstellung wird dabei zwischen Querschnitts-, Form-, und Topologieoptimierung (der eigentlichen Strukturoptimierung) unterschieden. Grundlegende Fragestellung ist dabei, die Struktur und die Abmessungen von Konstruktionen derart zu wählen, dass zum einen die mechanischen Randbedingungen erfüllt und zum anderen der Materialeinsatz und damit die Kosten möglichst gering sind.

Obwohl die Berücksichtigung der Nebenbedingungen oft die Koppelung mit komplizierten Berechnungsvorschriften – z. B. FEM-Solvern – erfordert, soll das Grundprinzip an folgendem Beispiel erläutert werden:

Beispiel 1.1 Ziel ist die Erstellung von Bemessungstafeln für geschweißte I-Träger mit Querschnitten minimalen Gewichts (Abb. 1.1).
Da das Gewicht eines Trägers mit vorgegebener Länge proportional zum Querschnitt ist, lautet die Zielfunktion

$$f(x) = (x_1 - 2x_4)x_2 + 2x_3x_4 .$$

Tragsicherheitsnachweise (g_1, g_2), Beulsicherheitsnachweise (g_3, g_4) und konstruktive Restriktionen $(g_5 - g_{10})$ führen zu den Nebenbedingungen $g_i \leq 0$:

$$
\begin{aligned}
g_1(x) &= \begin{cases} M_{pl}(x) - M_v & \text{für} \quad |N_v|/f(x) < 0.1\sigma_F \\ M_{pl}(x)(1.1 - |N_v|/(f(x)\sigma_F)) - M_v & \text{sonst} \end{cases} \\
g_2(x) &= 0.8f(x)\sigma_F - N_v \\
g_3(x) &= 17 - \frac{x_3}{x_4} \\
g_4(x) &= \begin{cases} 43 - x_1/x_2 & \text{für} \quad |N_v|/f(x) < 0.27\sigma_F \\ 70(1.1 - 1.4|N_v|/(f(x)\sigma_F)) - x_1/x_2 & \text{sonst} \end{cases} \\
g_5(x) &= 1000 - x_1 \\
g_6(x) &= 60 - x_2 \\
g_7(x) &= 60 - x_4 \\
g_8(x) &= x_2 - 5 \\
g_9(x) &= x_4 - 5 \\
g_{10}(x) &= \frac{x_1}{4} - x_3
\end{aligned}
$$

wobei $M_{pl}(x) := \sigma_F((x_1 - x_4)x_3x_4 + (x_1 - 2x_4)^2 x_2/4)$.
Die Größen M_v, N_v und σ_F sind konstante Parameter. Die Anzahl der Variablen ist 4, und es liegen 10 Ungleichheitsrestriktionen vor.

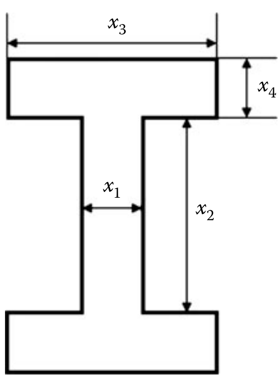


Abb. 1.1 Stahlträger.

1.4.2
Das Least-Squares-Problem

Spezielle nichtlineare Optimierungsaufgaben treten bei der Parameterbestimmung von Modellen auf, die einen in Natur- oder Technikwissenschaften vorliegenden Zusammenhang qualitativ beschreiben. Sind über diesen Zusammenhang Resultate von Experimenten bekannt, kann man die Methode der kleinsten