FernUniversität
Gesamthochschule in Hagen

Fachbereich Wirtschaftswissenschaft

Diplomarbeit

im wirtschaftswissenschaftlichen Zusatzstudiengang für Ingenieure und Naturwissenschaftler

Bearbeitungszeit 9 Wochen als Teilzeitstudent

im Fach:	Operations Research
über das Thema:	Der Variable-Size-Simplex-Algorithmus mit Straffunktion zur Lösung restringierter Optimierungsprobleme
eingereicht bei:	Univ.-Prof. Dr. W. Rödder
von:	Dr.-Ing. Ingo D. Rullhusen
Matrikel-Nr.:	4xxxxxx
Anschrift:	xxxxxxxxxxxxxxxxxxx, 28xxx Bremen
Abgabetermin:	9. Juli 2002

Inhaltsverzeichnis

Abbildungsverzeichnis

1 Einführung

Ziel dieser Arbeit ist die Umsetzung und Validierung eines Verfahrens zur Lösung restringierter Optimierungsprobleme. Hierbei sollen die Anforderungen an die Zielfunktion und die Nebenbedingungen möglichst gering sein. In diesem Sinne soll beispielsweise die Differenzierbarkeit nicht vorausgesetzt werden, was jedoch eine notwendige Bedingung für den effizienten Einsatz von Gradientenverfahren wäre. Die Funktionen in Zielfunktion und Nebenbedingungen brauchen noch nicht einmal analytisch zu sein, sondern können auch durch ein Meßverfahren oder einen Algorithmus definiert werden, so daß mithin nur eine direkte Suche möglich ist.

Zur Lösung des unrestringierten Optimierungsproblems in $\mathbb{R}^n$ sind direkte Simplex-Suchalgorithmen entwickelt worden. Diese Verfahren beruhen auf einem direkten Vergleich der Funktionswerte an den Scheitelpunkten eines Simplexes in $\mathbb{R}^n$. Der Simplex wird dann so modifiziert, daß der schlechteste Scheitelpunkt entfernt und durch einen neuen und besseren ersetzt wird, wodurch sich der Simplex auf das Optimum zu bewegt. Ein derartiger Algorithmus ist zunächst von Spendley et al. (1962) beschrieben worden. In dieser ursprünglichen Form war nur die sogenannte Reflexionsoperation enthalten, so daß der verschobene Simplex stets seine Größe beibehielt; das Optimum konnte also nicht beliebig genau erreicht werden. Dieses Verfahren ist dann von Nelder/Mead (1965) zum Variable-Size-Simplex-Algorithmus[1] erweitert worden. Der Variable-Size-Simplex-Algorithmus nach Nelder/Mead (1965) kann mittels zusätzlicher Kontraktion und Expansion die Simplexgröße verändern, so daß das Optimum einerseits mit möglichst wenigen größeren Schritten erreicht und andererseits mit kleiner werdenden Schritten genauer lokalisiert werden kann.

Praktische Optimierungsprobleme sind jedoch häufig unter gewissen Nebenbedingungen zu lösen. Um diese Probleme mit einem unrestringierten Algorithmus wie dem Variable-Size-Simplex-Algorithmus zu lösen, muß man die Zielfunktion derart modifizieren, daß das modifizierte globale Optimum gleichzeitig die Nebenbedingungen erfüllt. Hierzu gibt es die Möglichkeit, einen Strafterm zur ursprünglichen Zielfunktion hinzuzufügen, der eine Verletzung der Nebenbedingungen mittels einer Zielwertverschlechterung vermeiden soll. Beim inneren Straffunktionsverfahren fügt man eine Barrierefunktion hinzu, die eine Annäherung an den Rand des zulässigen Gebietes durch ihren Verlauf verhindern soll. Hierbei muß die Lösungssuche jedoch mit zulässigen Ausgangswerten

[1] Nicht zu verwechseln mit dem Simplexalgorithmus der linearen Optimierung.

begonnen werden, deren Auffindung im höherdimensionalen Fall keineswegs trivial ist. Andererseits sind alle erhaltenen Zwischenlösungen zulässig. Es muß jedoch sichergestellt sein, daß der unrestringierte Löser nicht auf das unzulässige Gebiet zugreift. Beim eigentlichen oder auch äußeren Straffunktionsverfahren fügt man eine Straffunktion hinzu, die eine Zielwertverschlechterung bewirkt, sofern das zulässige Lösungsgebiet verlassen wird. Die Anfangswerte müssen nicht zulässig sein, worin der wesentliche Vorteil dieser Vorgehensweise besteht. Jedoch sind die Zwischenlösungen im allgemeinen nicht zulässig, da das Optimum von außen angenähert wird. Auch das Endergebnis ist nur im Rahmen der gesetzten Fehlerschranke zulässig. Ferner müssen hierbei alle Funktionen im gesamten $\mathbb{R}^n$ definiert sein. Die beliebigen Anfangswerte stellen jedoch einen erheblichen praktischen Vorteil dar, so daß diese Methode der Straffunktion im folgenden zum Einsatz gelangen soll.

Im anschließenden Abschnitt wird zunächst der Variable-Size-Simplex-Algorithmus gemäß Nelder/Mead (1965) prinzipiell und algorithmisch in einer Form beschrieben, wie er im Rahmen dieser Arbeit auch implementiert worden ist. Es folgt die Beschreibung der genutzten Methode der Straffunktion nach Prinzip und Algorithmus. Hierbei handelt es sich um die auf mehrere Skalierungsfaktoren erweiterte Methode von Zangwill (1967). Aus diesem Vorgehen läßt sich anschließend eine Schätzung für die Lagrange-Multiplikatoren gewinnen. Es folgen anschließend die Anwendung des in der Programmiersprache C implementierten Algorithmus auf diverse Beispielprobleme aus der Literatur und eine abschließende Zusammenfassung.

2 Grundlegendes zum Variable-Size-Simplex-Algorithmus

Im Rahmen dieser Arbeit kommt der Variable-Size-Simplex-Algorithmus gemäß Nelder/Mead (1965) zum Einsatz. Dieser ermöglicht die Lösung des unrestringierten Optimierungsproblems

$$\min_{\underline{x}\in\mathbb{R}^n} y = y(\underline{x}), \tag{1}$$

wobei an die Funktion $y(\underline{x})$ keine weiteren Anforderungen gestellt werden. Insbesondere muß y nicht differenzierbar sein, wie beispielsweise für die Gradientenverfahren. Im Gegenzug ist die Konvergenz jedoch nicht garantiert, da es sich um ein heuristisches Lösungsschema handelt.

2.1 Die Funktionsweise

Der Variable-Size-Simplex-Algorithmus ist ein direktes Suchverfahren, daß mit den ermittelten Funktionswerten nach einer Lösungsverbesserung sucht. Um die Anzahl der betrachteten Funktionswerte zu minimieren, wird als Suchgebiet im n-dimensionalen Raum ein n-dimensionaler Simplex verwendet. Bei einem n-dimensionalen Simplex handelt es sich um einen Hyperkörper mit $n+1$ Scheitelpunkten, dessen berandende Hyperflächen jeweils wiederum $(n-1)$-dimensionale Simplexe darstellen. Bei einem nulldimensionalen Simplex handelt es sich um einen Punkt, bei einem eindimensionalen um eine Strecke, bei einem zweidimensionalen um ein Dreieck, bei einem dreidimensionalen um ein Tetraeder und so weiter. Zur Darstellung des Suchgebietes im n-dimensionalen Raum sind mithin nur $n+1$ Scheitelpunkte nötig, wohingegen ein n-dimensionaler Hyperquader beispielsweise aus 2^n Punkten besteht.

Das Grundprinzip des Variable-Size-Simplex-Algorithmus besteht nun darin, aus den bekannten Funktionswerten an den $n+1$ Scheitelpunkten einen neuen Scheitelpunkt mit besserem Funktionswert zu bestimmen, um mit selbigem den schlechtesten der bisherigen zu ersetzen. Hierzu kommen nach Nelder/Mead (1965) die Simplexoperationen Reflexion, Expansion, Kontraktion und totale Kontraktion zu Einsatz. Diese Operationen sind in der Abbildung 1 anhand eines Dreiecks schematisch dargestellt. Das alte gestrichelt gezeichnete Dreieck wird durch ein neues durchgezogen gezeichnetes Dreieck ersetzt, indem der schlechteste Scheitelpunkt $\underline{x}_h$ nach Möglichkeit durch einen besseren ersetzt wird. Die Reflexion stellt die Grundoperation dar, die für die Suchbewegung des Simplexes in Richtung Optimum sorgt. Die Expansion erlaubt es, die Schrittweite in Richtung Optimum zu vergrößern, wenn damit das Optimum schneller erreicht werden kann. Als Gegenspieler sorgt die Kontraktion für eine Verkleinerung des Simplexes, wenn anderenfalls in der Nähe des Optimums keine Verbesserung mehr möglich ist. Die totale Kontraktion bewirkt ähnliches in Richtung der bisher besten Lösung $\underline{x}_l$, wenn ansonsten keine Verbesserung mehr möglich ist.

Der Variable-Size-Simplex-Algorithmus beginnt also mit einem Startsimplex. Dieser wird iterativ mittels der Simplexoperationen in Richtung Optimum verschoben und um das Optimum herum verkleinert. Nach dem Erreichen des Abbruchkriteriums wird aus dem letzten Simplex die optimale Lösung

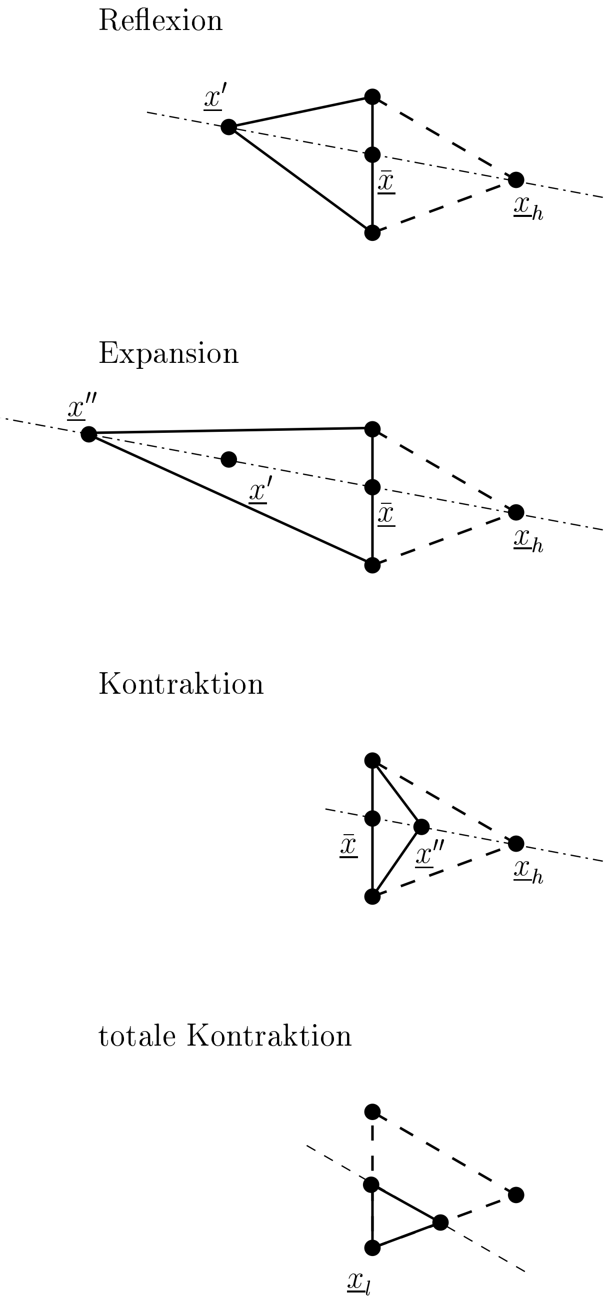

Abbildung 1: Verdeutlichung der im Variable-Size-Simplex-Algorithmus nach Nelder/Mead (1965) durchgeführten Simplexoperationen am Beispiel eines Dreiecks: Das ursprüngliche Dreieck ist gestrichelt, das resultierende Dreieck durchgezogen gezeichnet.

extrahiert. Da es sich hierbei um eine Heuristik handelt, kann die Konvergenz jedoch nicht garantiert werden. Die Konvergenz und die Konvergenzgeschwindigkeit sind vielmehr von der Wahl des Startsimplexes und der Parameter der Simplexoperationen abhängig.

2.2 Der Algorithmus

Nach den allgemeinen Ausführungen soll nun der konkrete Algorithmus erläutert werden, wie er von NELDER/MEAD (1965) vorgestellt worden ist. Der Algorithmus beginnt mit der Aufstellung des Startsimplexes. Als solcher wird im Rahmen dieser Arbeit der ein vorgegebenes Hyperquader umschreibende Simplex gewählt, wodurch dieses Hyperquader mit seinen Abmessungen die anfängliche Schrittweite festlegt.

Nachdem alle Funktionswerte y_i der Scheitelpunkte $\underline{x}_i$ berechnet worden sind, erfolgen die iterativen Simplexschritte zur Verbesserung der Lösung. Das Flußdiagramm eines Schrittes des Variable-Size-Simplex-Algorithmus ist in der Abbildung 2 aufgeführt. Ein solcher Schritt beginnt damit, daß die Indizes des minimalen Funktionswertes l und des maximalen Funktionswertes h bestimmt werden. Desweiteren wird der Mittelpunkt $\underline{\bar{x}}$ aller Scheitelpunkte $\underline{x}_i$ mit $i \neq h$ berechnet. Der Heuristik folgend liegt dieser vom schlechtesten Scheitelpunkt $\underline{x}_h$ aus gesehen in Richtung einer Lösungsverbesserung. Demgemäß bestimmt man einen neuen Scheitelpunktkandidaten durch Reflexion zu

$$\underline{x}' = (1+\alpha)\underline{\bar{x}} - \alpha\underline{x}_h, \tag{2}$$

wobei der Reflexionskoeffizient $\alpha > 0$ sein soll. Ist der zugehörige Funktionswert y' tatsächlich besser als der bisherige beste Wert y_l, so versucht man in der erfolgreichen Richtung noch weiter zu expandieren und ermittelt mit

$$\underline{x}'' = \gamma\underline{x}' + (1-\gamma)\underline{\bar{x}} \tag{3}$$

einen weiteren Kandidaten, wobei für den Expansionskoeffizienten $\gamma > 1$ gelten soll. Wegen der linearen Beziehungen liegen die Reflexions- und Expansionspunkte auf einer Geraden durch den Hilfspunkt $\underline{\bar{x}}$ und den schlechtesten Scheitelpunkt $\underline{x}_h$. Ist der Funktionswert des neuen Punktes $\underline{x}''$ besser als der bisher beste zu $\underline{x}_l$ so wird der bisher schlechteste Scheitelpunkt $\underline{x}_h$ durch $\underline{x}''$ ersetzt; ansonsten wird $\underline{x}_h$ durch $\underline{x}'$ ersetzt. Danach ist der Simplexschritt beendet.

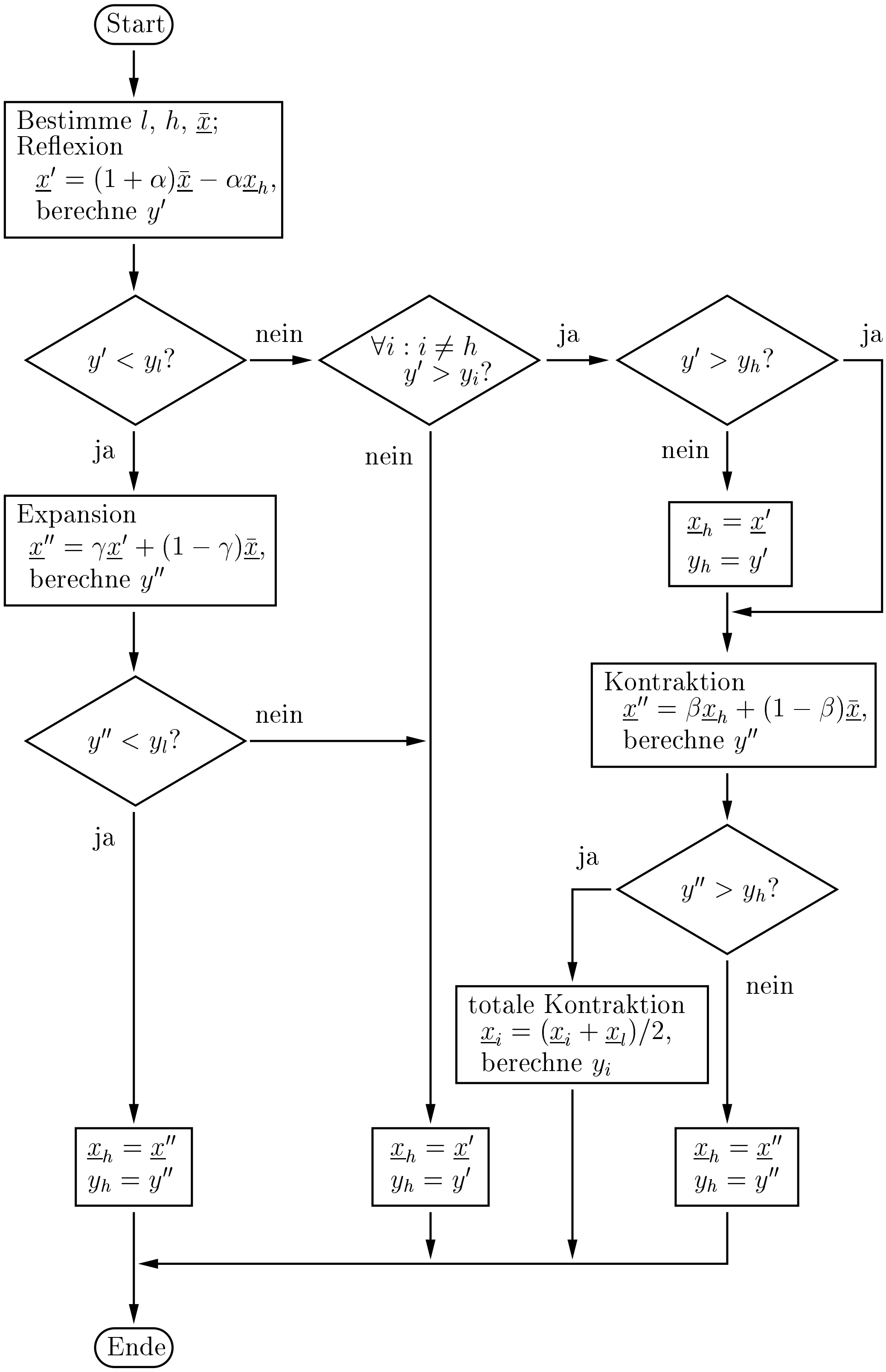

Abbildung 2: Flußdiagramm eines Schrittes des Variable-Size-Simplex-Algorithmus gemäß NELDER/MEAD (1965).

War der Reflexionspunkt $\underline{x}'$ jedoch nicht besser als der bisherige beste Scheitelpunkt $\underline{x}_l$, so sind weitere Untersuchungen notwendig. Ist der Reflexionspunkt wenigstens nicht schlechter als alle bisherigen Scheitelpunkte mit Ausnahme des schlechtesten, so wird er dennoch als Ersatz für den bisher schlechtesten Scheitelpunkt $\underline{x}_h$ akzeptiert und der Simplexschritt ist beendet.

Anderenfalls wird geprüft, ob der Reflexionspunkt $\underline{x}'$ schlechter ist als der bisher schlechteste Scheitelpunkt $\underline{x}_h$. Ist dies nicht der Fall, so wird der bisher schlechteste Scheitelpunkt $\underline{x}_h$ durch $\underline{x}'$ zwecks leichter Verbesserung ersetzt, bevor der Kontraktionspunkt mit

$$\underline{x}'' = \beta \underline{x}_h + (1 - \beta)\bar{\underline{x}} \tag{4}$$

bestimmt wird. Für den Kontraktionskoeffizienten möge hierbei $0 < \beta < 1$ gelten. Ist der Funktionswert des Kontraktionspunktes nicht schlechter als der bisher schlechteste, so wird der bisherige schlechteste Scheitelpunkt $\underline{x}_h$ durch diesen ersetzt und der Simplexschritt ist beendet. Anderenfalls konnte insgesamt kein besserer Scheitelpunkt gefunden werden und als Notlösung wird eine totale Kontraktion zum bisher besten Scheitelpunkt $\underline{x}_l$ hin durchgeführt, indem man alle Scheitelpunkte durch

$$\underline{x}_i = \frac{1}{2}(\underline{x}_i + \underline{x}_l) \tag{5}$$

ersetzt. Der bisher beste Scheitelpunkt $\underline{x}_l$ bleibt somit unverändert und alle übrigen rücken in seine Nähe. Hiernach ist der Simplexschritt in jedem Fall beendet.

Es folgen weitere Schritte, bis das Abbruchkriterium erfüllt ist. Dieses besteht aus zwei Teilen: Zum einen wird gemäß NELDER/MEAD (1965) geprüft, ob

$$\sqrt{\frac{1}{n}\sum_i (y_i - \langle y \rangle)^2} < \epsilon \tag{6}$$

erfüllt ist, wobei $\langle y \rangle$ das arithmetische Mittel aller Funktionswerte der Scheitelpunkte bezeichnet. Bei flachem Verlauf der Zielfunktion kann dieses Kriterium jedoch bereits erfüllt sein, wohingegen im Argumentenvektor $\underline{x}$ noch deutliche Änderungen auftreten. Somit wird neben diesem y-Kriterium zum anderen noch ein analoges x-Kriterium mit

$$\sqrt{\frac{1}{n}\sum_i ||\underline{x}_i - \langle \underline{x} \rangle||^2} < \epsilon \tag{7}$$

definiert. Ist ein Kriterium erfüllt, so erfolgt standardmäßig die Beendigung des Variable-Size-Simplex-Algorithmus mit der Angabe, welches Kriterium erfüllt worden ist. In der Regel wird zunächst das y-Kriterium anschlagen, was in der Praxis auch einer Zielerreichung entspricht. Möchte man jedoch nicht nur den Optimalwert sondern auch den Optimalpunkt hinreichend genau wissen, so muß man die Iteration bis zum Erreichen des x-Kriteriums fortsetzen.

3 Grundlegendes zur Methode der Straffunktion

Die Methode der Straffunktion zur Lösung von restringierten Problemen wurde vermutlich erstmals von Courant (1943) angewandt. Die mathematische Fundierung für den Anwendungsbereich der nichtlinearen Optimierung ist im wesentlichen von Zangwill (1967) begründet worden. Im folgenden wird zunächst ein Abriß dieser Grundlagen gegeben soweit sie die Funktionsweise dieser Methode verdeutlichen. Darauf aufbauend wird dann der angewandte Algorithmus beschrieben.

3.1 Die Funktionsweise

Die Berücksichtigung von Nebenbedingungen bei der Minimierung einer Zielfunktion erfolgt bei der Methode der Straffunktion mittels eines Strafzuschlags zum Funktionswert, falls sich die Variablenwerte außerhalb des zulässigen Bereiches befinden. Somit wird das ursprüngliche restringierte Optimierungsproblem

$$\min f(\underline{x}) \tag{8a}$$

u.d.N.

$$g_i(\underline{x}) \leq 0 \qquad \text{mit} \quad i \in \{1; \ldots; m_1\}, \tag{8b}$$
$$h_j(\underline{x}) = 0 \qquad \text{mit} \quad j \in \{1; \ldots; m_2\}, \tag{8c}$$

wobei $\underline{x} \in \mathbb{R}^n$ sei, in das nichtrestringierte Problem

$$\min_{\underline{x} \in \mathbb{R}^n} q(\underline{x}, \underline{r}) = f(\underline{x}) + \Phi(\underline{x}, \underline{r}) \tag{9}$$

transformiert, wobei für den Skalierungsvektor $\underline{r} \in \mathbb{R}^{m_1+m_2}$ gelte. Bezeichnet man mit X die Menge der gemäß (8b) und (8c) zulässigen Variablenvektoren

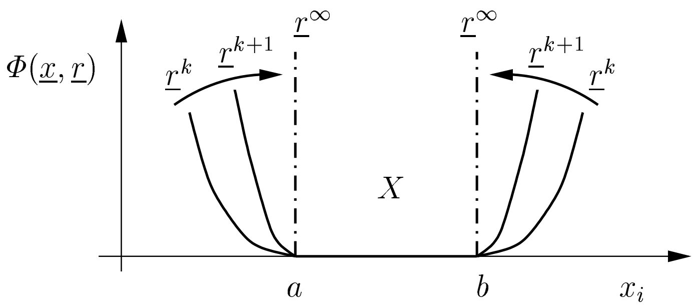

Abbildung 3: Schematische Verdeutlichung der Straffunktion für den Fall eines abgeschlossenen Intervalls $[a; b]$ als Zulässigkeitsbereich für x_i.

$\underline{x}$, so soll die stetige Straffunktion Φ die Eigenschaften

$$\Phi(\underline{x}, \underline{r}) = 0 \qquad \forall \underline{x} \in X, \tag{10}$$

$$\Phi(\underline{x}, \underline{r}) > 0 \qquad \forall \underline{x} \notin X, \tag{11}$$

$$\Phi(\underline{x}, \underline{r}) \to \infty \quad \text{bei} \quad \underline{r} \to \infty \wedge \underline{x} \notin X \tag{12}$$

besitzen. Diese Eigenschaften werden in der Abbildung 3 verdeutlicht. Bei unendlichem $\underline{r}$ würde sich also die Lösung des ursprünglichen Optimierungsproblems (8) unmittelbar als Lösung des unrestringierten Problems (9) ergeben. Wie ZANGWILL (1967) gezeigt hat, ist dies unter bestimmten Umständen theoretisch bereits für ein hinreichend groß gewähltes $\underline{r}$ möglich. Zu dessen Ermittlung muß jedoch die optimale Lösung bereits bekannt sein. In der Rechenpraxis betrachtet man daher eine Folge $\{\underline{r}^k\}$ von Skalierungsvektoren mit der Monotonieeigenschaft

$$\underline{0} < \underline{r}^k \leq \underline{r}^{k+1}, \tag{13}$$

was bedeutet, daß mindestens ein Element des Skalierungsvektors wachsen muß falls die Ungleichung gilt und daß alle Elemente positiv sind. Zu jedem Skalierungsvektor wird eine Optimallösung des unrestringierten Optimierungsproblems (9) bestimmt, so daß man eine Folge $\{\underline{x}^k\}$ von solchen Lösungen erhält. Daß diese Folge tatsächlich gegen die exakte Lösung $\underline{x}^*$ des ursprünglichen restringierten Optimierungsproblems (8) konvergiert, soll im folgenden verdeutlicht werden.

Hierzu kommt es zunächst auf die genauere Struktur der Straffunktion Φ an. Insbesondere muß selbige bezüglich des Skalierungsvektors $\underline{r}$ und der Variablen $\underline{x}$ separierbar sein. Hierfür wird der Ansatz

$$\Phi(\underline{r}, \underline{x}) = \underline{r}^{\mathrm{T}} \cdot \underline{s}(\underline{x}) \tag{14}$$

gewählt, wobei für den Vektor der nichtnegativen Verlustfunktionen $\underline{s}(\underline{x})$

$$\underline{s}(\underline{x}) = \underline{0} \qquad \forall \underline{x} \in X \tag{15}$$

und genau dann bei Verletzung der i-ten Restriktion

$$s_i(\underline{x}) > 0 \tag{16}$$

gelten sollen, um den Bedingungen (10) und (11) zu genügen. Aus dieser Definition folgt unmittelbar die Beziehung

$$\Phi(\underline{x}, \underline{r}^{k+1}) \geq \Phi(\underline{x}, \underline{r}^{k}) \tag{17}$$

als zusätzliche Eigenschaft der Straffunktion, sofern die Beziehung (13) erfüllt ist.

Zunächst einmal sei hiermit das Verhalten der transformierten Zielfunktion q betrachtet. Für selbige gilt:

$$\begin{aligned}
q(\underline{x}^{k+1}, \underline{r}^{k+1}) &= f(\underline{x}^{k+1}) + \Phi(\underline{x}^{k+1}, \underline{r}^{k+1}) && \text{(18a)}\\
&\geq f(\underline{x}^{k+1}) + \Phi(\underline{x}^{k+1}, \underline{r}^{k}) && \text{(18b)}\\
&= q(\underline{x}^{k+1}, \underline{r}^{k}) && \text{(18c)}\\
&\geq q(\underline{x}^{k}, \underline{r}^{k}) && \text{(18d)}\\
&= f(\underline{x}^{k}) + \Phi(\underline{x}^{k}, \underline{r}^{k}). && \text{(18e)}
\end{aligned}$$

Neben der Definition (9) und der Relation (17) ist hierbei nur zu beachten, daß $\underline{x}^k$ per definitionem gerade $q(\underline{x}, \underline{r}^k)$ minimiert. Der Zielfunktionswert des transformierten Problems verhält sich also wegen

$$q(\underline{x}^{k+1}, \underline{r}^{k+1}) \geq q(\underline{x}^{k}, \underline{r}^{k}) \tag{19}$$

während des Iterationsprozesses monoton steigend.

Aus der Ungleichung (18b) (18e) folgt unmittelbar

$$f(\underline{x}^k) + (\underline{r}^k)^{\mathrm{T}} \cdot \underline{s}(\underline{x}^k) \leq f(\underline{x}^{k+1}) + (\underline{r}^k)^{\mathrm{T}} \cdot \underline{s}(\underline{x}^{k+1}) \tag{20}$$

und die Minimierungseigenschaft ergibt erneut

$$f(\underline{x}^{k+1}) + (\underline{r}^{k+1})^{\mathrm{T}} \cdot \underline{s}(\underline{x}^{k+1}) \leq f(\underline{x}^k) + (\underline{r}^{k+1})^{\mathrm{T}} \cdot \underline{s}(\underline{x}^k). \tag{21}$$

Addiert man die letzten beiden Ungleichungen, so erhält man nach einigen Umformungen

$$0 \leq (\underline{r}^{k+1} - \underline{r}^k)^{\mathrm{T}} \cdot \underline{s}(\underline{x}^{k+1}) \leq (\underline{r}^{k+1} - \underline{r}^k)^{\mathrm{T}} \cdot \underline{s}(\underline{x}^k). \tag{22}$$

Wählt man die Folge $\{\underline{r}^k\}$ nun so, daß

$$r_i^{k+1} = \begin{cases} r_i^k & \forall i:\ s_i(\underline{x}^k) = s_i(\underline{x}^{k+1}) = 0 \\ \varrho r_i^k & \text{sonst} \end{cases} \tag{23}$$

gilt, wobei $\varrho > 1$ sei, so wird die Relation (22) zu

$$(\varrho - 1)(\underline{r}^k)^{\mathrm{T}} \cdot \underline{s}(\underline{x}^{k+1}) \leq (\varrho - 1)(\underline{r}^k)^{\mathrm{T}} \cdot \underline{s}(\underline{x}^k)$$

und mithin ergibt sich

$$0 \leq (\underline{r}^k)^{\mathrm{T}} \cdot (\underline{s}(\underline{x}^k) - \underline{s}(\underline{x}^{k+1})). \tag{24}$$

Der Erfüllungsgrad der Restriktionen bezüglich eines Skalierungsvektors wird somit in keinem Iterationsschritt verschlechtert.

Greift man nochmals auf die Ungleichung (20) zurück, so erhält man aus selbiger die Relation

$$f(\underline{x}^{k+1}) - f(\underline{x}^k) \geq (\underline{r}^k)^{\mathrm{T}} \cdot (\underline{s}(\underline{x}^k) - \underline{s}(\underline{x}^{k+1})) \geq 0, \tag{25}$$

sofern man ferner die Beziehung (24) berücksichtigt. Hieraus ergibt sich unmittelbar

$$f(\underline{x}^{k+1}) \geq f(\underline{x}^k). \tag{26}$$

Somit ist die Folge der optimalen Zielfunktionswerte $\{f(\underline{x}^k)\}$ monoton ansteigend, wie es auch bei einer Annäherung aus dem unzulässigen, äußeren Bereich zu erwarten ist.

Das Minimum des restringierten Problems $\underline{x}^* \in X$ existiere, dann gilt offenbar wegen der Minimierungseigenschaft

$$f(\underline{x}^*) = f(\underline{x}^*) + (\underline{r}^k)^{\mathrm{T}} \cdot \underbrace{\underline{s}(\underline{x}^*)}_{=\underline{0}} = q(\underline{x}^*, \underline{r}^k) \geq q(\underline{x}^k, \underline{r}^k) \tag{27}$$

und damit ist die transformierte Zielfunktion q nach oben beschränkt. Ferner ist die Folge $\{q(\underline{x}^k, \underline{r}^k)\}$ gemäß (19) monoton steigend und somit existiert ein endlicher Grenzwert

$$\lim_{k\to\infty} q(\underline{x}^k, \underline{r}^k) = q^\infty. \tag{28}$$

Wegen der Nichtnegativität von $\underline{r}^k$ und $\underline{s}$ ergibt sich desweiteren

$$f(\underline{x}^k) \le f(\underline{x}^k) + (\underline{r}^k)^{\mathrm{T}} \cdot \underline{s}(\underline{x}^k) = q(\underline{x}^k, \underline{r}^k) \le f(\underline{x}^*) \tag{29}$$

und somit besitzt auch die Folge $\{f(\underline{x}^k)\}$ eine obere Schranke. Zusammen mit der Eigenschaft (26) folgt die Existenz eines endlichen Grenzwertes

$$\lim_{k\to\infty} f(\underline{x}^k) = f^\infty. \tag{30}$$

Zusammengenommen folgt hieraus dann die Existenz des Grenzwertes

$$\lim_{k\to\infty} \; (\underline{r}^k)^{\mathrm{T}} \cdot \underline{s}(\underline{x}^k) = q^\infty - f^\infty. \tag{31}$$

Da jedoch $\underline{r}^k \to \infty$ für $k \to \infty$ gelten soll, zumindest soweit die Restriktionen wirksam sind, muß wegen der Nichtnegativität von $\underline{r}^k$ und $\underline{s}$

$$\lim_{k\to\infty} \underline{s}(\underline{x}^k) = \underline{0} \tag{32}$$

sein, um den endlichen Grenzwert (31) zu ermöglichen. Für die unwirksamen Restriktionen gilt $s_i(\underline{x}^k) = 0$ per definitionem.

Der letzte Grenzwert besagt nun, daß für jedes $\epsilon > 0$ ein $K(\epsilon) \in \mathbb{N}$ existiert, so daß $\underline{x}^k \in X^\epsilon$ für alle $k > K(\epsilon)$ gilt, wobei

$$X^\epsilon = \{x | x \in \mathbb{R}^n \wedge s_i(\underline{x}) \le \epsilon \; \forall i = 1; \ldots; m_1 + m_2\} \tag{33}$$

zusätzlich zu X noch eine kleine Umgebung enthält, insofern alle $s_i(\underline{x})$ stetig sind. Ist X^ϵ kompakt, beziehungsweise als Teilmenge von $\mathbb{R}^n$ abgeschlossen und beschränkt, so gibt es eine konvergente Teilfolge $\{\underline{x}^{k_i}\}$ zu $\{\underline{x}^k\}$, die gegen ein $\underline{x}^\infty \in X^\epsilon$ konvergiert.[2] Für diesen Häufungspunkt muß wegen der Beziehung (32) ferner

$$\underline{s}(\underline{x}^\infty) = \underline{0} \tag{34}$$

gelten. Es handelt sich also auch um eine zulässige Lösung. Wegen der Optimalität der ebenfalls zulässigen Lösung $\underline{x}^*$ gilt

$$f(\underline{x}^*) \le f(\underline{x}^\infty). \tag{35}$$

[2] Siehe beispielsweise Bronstein/Semendjajew (1991).

Für die konvergente Teilfolge $\{\underline{x}^{k_i}\}$ gilt

$$\begin{aligned} f(\underline{x}^{k_i}) &\leq f(\underline{x}^{k_i}) + (\underline{r}^{k_i})^{\mathrm{T}} \cdot \underline{s}(\underline{x}^{k_i}) \\ &= q(\underline{x}^{k_i}, \underline{r}^{k_i}) \\ &\leq q(\underline{x}^{*}, \underline{r}^{k_i}) \\ &= f(\underline{x}^{*}) + (\underline{r}^{k_i})^{\mathrm{T}} \cdot \underbrace{\underline{s}(\underline{x}^{*})}_{=\underline{0}} \\ &= f(\underline{x}^{*}), \end{aligned} \tag{36}$$

woraus sich über

$$\lim_{k_i \to \infty} f(\underline{x}^{k_i}) = f(\underline{x}^{\infty}) \leq f(\underline{x}^{*}) \tag{37}$$

und zusammen mit der Relation (35) abschließend

$$f(\underline{x}^{\infty}) = f(\underline{x}^{*}) \tag{38}$$

ergibt. Somit ist der Konvergenzpunkt $\underline{x}^{\infty}$ eine optimale Lösung des restringierten Optimierungsproblems (8).

Die entscheidende Voraussetzung dieser Herleitung und somit auch für die Konvergenz der Straffunktionsmethode ist die Möglichkeit, das Minimum $\underline{x}^k$ von $q(\underline{x}, \underline{r}^k)$ für alle $\underline{r}^k$ aufzufinden. Gemäß Avriel (1976) ist hierzu eine Zielfunktion f notwendig für die $f(\underline{x}) \to \infty$ bei $||\underline{x}|| \to \infty$ gilt, die Menge X^{ϵ} muß desweiteren für einige $\epsilon > 0$ kompakt sein und die transformierte Zielfunktion $q(\underline{x}, \underline{r})$ muß konvex bezüglich $\underline{x}$ sein. Die Tatsache, daß das globale Optimum zu bestimmen ist, erschwert diese Teilaufgabe noch darüber hinaus. Im Rahmen dieser Arbeit kann die Konvergenz jedoch grundsätzlich nicht garantiert werden, da der zur Lösung der unrestringierten Ersatzprobleme eingesetzte Variable-Size-Simplex-Algorithmus seines heuristischen Ursprungs halber selbst keine Konvergenz garantiert, auch wenn das Optimierungsproblem lösbar ist. Im Rahmen der numerischen Anwendung wird davon ausgegangen, daß Konvergenz vorliegt, sofern eine Fehlerschranke innerhalb einer maximalen Iterationsanzahl unterschritten worden ist. Gilt dies für ein einzelnes Ersatzproblem nicht, so wird die Lösung als unrichtig markiert. Bei Erhalt einer zulässigen Lösung kann man somit in der Praxis davon ausgehen, daß die Lösbarkeit gegeben war und das Straffunktionsverfahren gemäß dem zuvor geschilderten Verhalten abgelaufen ist.

3.2 Der Algorithmus

Um die Methode der Straffunktion als Algorithmus zu spezifizieren, sind zusätzlich zu den obigen Angaben noch die Verlustfunktionen $\underline{s}(\underline{x})$ zu definieren. Im Rahmen dieser Arbeit werden jene als

$$s_i(\underline{x}) = \begin{cases} (\max(0; g_i(\underline{x})))^2 & \text{für} \quad i \in \{1; \ldots; m_1\} \\ (h_{i-m_1}(\underline{x}))^2 & \text{für} \quad i \in \{m_1 + 1; \ldots; m_1 + m_2\} \end{cases} \tag{39}$$

angesetzt, was den Forderungen (15) und (16) genügt. Auf die Differenzierbarkeit muß eigentlich bei Verwendung des Variable-Size-Simplex-Algorithmus keine Rücksicht genommen werden. Zur späteren Schätzung der Lagrange-Multiplikatoren ist jedoch die stetige Differenzierbarkeit erforderlich. Die Quadrierung ergibt einen hinreichend starken Anstieg der Straffunktion im unzulässigen Bereich. Der monotone Verlauf der Verlustfunktionen im unzulässigen Bereich unterstützt die Auffindung des globalen Minimums.

Der Ablauf des Algorithmus ist im Flußdiagramm der Abbildung 4 schematisch dargestellt: Nach dem Aufruf des Algorithmus wird zunächst einmal die Zielfunktion q des unrestringierten Ersatzproblems (9) gemäß den Spezifikationen gebildet. Es sind alle Abbruch- und Inkrementparameter zu initialisieren sowie die Parameter des Variable-Size-Simplex-Algorithmus festzulegen. Im nächsten Schritt wird der Variable-Size-Simplex-Algorithmus initialisiert und ein Startsimplex gemäß den externen Vorgaben gebildet. Diese Vorgaben bestehen in einem n-dimensionalen Hyperquader, das vom Startsimplex umschrieben wird.

Der Lösungsprozeß beginnt nun mit dem Aufruf des Variable-Size-Simplex-Algorithmus zur Bestimmung des globalen Optimums $\underline{x}^k$ der transformierten Zielfunktion $q(\underline{x}, \underline{r}^k)$ bei gegebenem Skalierungsvektor $\underline{r}^k$. Wird keine Lösung gefunden, so terminiert der Prozeß mangels Konvergenz, da die Voraussetzung für die Straffunktionsmethode nicht erfüllt ist. Wird eine Optimallösung des Ersatzproblems gefunden, so wird deren Zulässigkeit bezüglich der Fehlerschranke ϵ geprüft. Jede verletzte Restriktion i führt zu einer Erhöhung des beigeordneten Skalierungsparameters r_i^k um einen Inkrementfaktor ϱ, wie es im Ansatz (23) gefordert wird. Findet keine einzige derartige Erhöhung statt, so werden alle Restriktionen auf ϵ genau eingehalten und der Algorithmus endet erfolgreich. Anderenfalls wird der Iterationszähler um eins erhöht und mit dem Lösungsprozeß fortgefahren, solange die maximal zulässige Anzahl an Iterationen noch nicht erreicht worden ist.

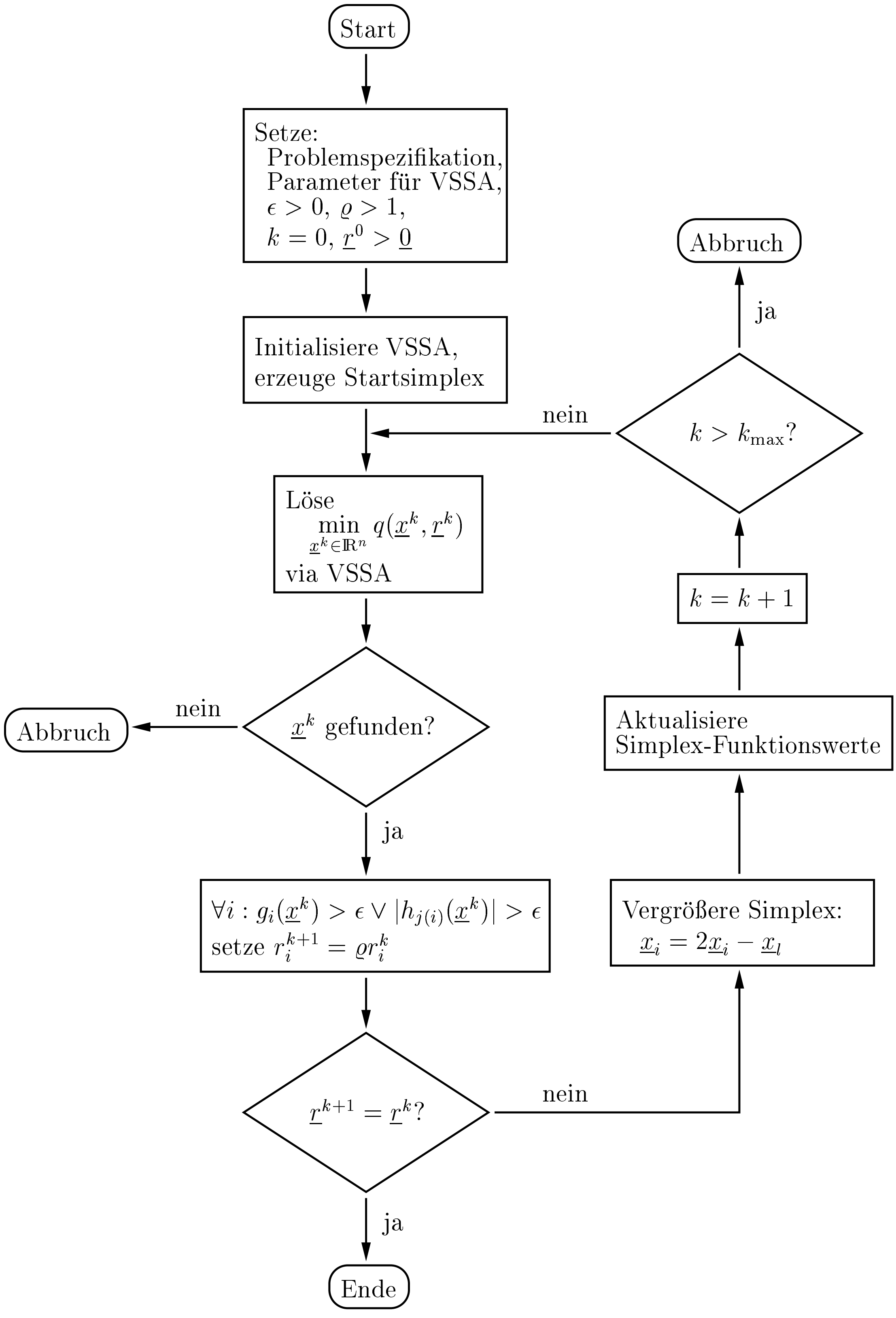

Abbildung 4: Flußdiagramm der Straffunktionsmethode zur Lösung restringierter Optimierungsprobleme mittels des Variable-Size-Simplex-Algorithmus (VSSA); siehe auch Abbildung 2.

Zu beachten ist, daß im Iterationsprozeß jeweils der vergrößerte Lösungssimplex des vorherigen Durchlaufs als Startsimplex für den Variable-Size-Simplex-Algorithmus im nächsten Durchlauf dient, da sich das Optimum zum vorherigen Durchlauf nur wenig verschieben sollte. Diese Vergrößerung stellt eine Umkehrung der totalen Kontraktion des Variable-Size-Simplex-Algorithmus dar und ist notwendig, damit die Simplexgröße über dessen Abbruchkriterium liegt und selbiger überhaupt anläuft. Die anfänglichen Zielfunktionswerte im Simplex sind natürlich im Rahmen des Variable-Size-Simplex-Algorithmus zuvor zu aktualisieren. Diese Aktualisierung kann mit der Aktualisierung des Skalierungsvektors, der Feststellung der Zielerreichung und der Simplexvergrößerung zusammengefaßt werden.

Nach der erfolgreichen Beendigung ist auf der Grundlage des letzten Simplexes die Lösung zu erzeugen. Der Optimalpunkt $\underline{x}^*$ ist der Scheitelpunkt mit dem minimalen Funktionswert. Der optimale Funktionswert ist danach nochmals mit der ursprünglichen Zielfunktion $f(\underline{x})$ zu berechnen, um Störungen durch Überreste der Straffunktion auszuschließen. Aus Vergleichen mit den übrigen Scheitelpunkten können die Genauigkeiten des Funktionswertes und des Optimalpunktes abgeschätzt werden.

3.3 Schätzung der Lagrange-Multiplikatoren

Wie von Fiacco/McCormick (1968) dargelegt worden ist, kann man auf der Grundlage der Methode der Straffunktion eine Schätzung der Schattenkosten bzw. Lagrange-Multiplikatoren gewinnen, sofern alle betrachteten Funktionen im Optimum differenzierbar sind. Diese Schätzung kann als Ansatzpunkt für eine optimierende Modifikation der Restriktionen interessant sein.

Hierzu müssen die Verlustfunktionen lediglich gewisse Bedingungen erfüllen. Bildet man den Gradienten der transformierten Zielfunktion q bezüglich des Variablenvektors $\underline{x}$, so erhält man nämlich

$$\operatorname{grad}_{\underline{x}} q(\underline{x}, \underline{r}) = \operatorname{grad}_{\underline{x}} f(\underline{x}) + \sum_i r_i \operatorname{grad}_{\underline{x}} s_i(\underline{x}), \tag{40}$$

sofern man die Gleichungen (9) und (14) berücksichtigt. Die Kettenregel der

Differentialrechnung führt dann zu

$$\begin{aligned}\operatorname{grad}_{\underline{x}} q(\underline{x},\underline{r}) = \operatorname{grad}_{\underline{x}} f(\underline{x}) + \sum_{i=1}^{m_1} r_i \frac{\mathrm{d}s_i}{\mathrm{d}g_i} \operatorname{grad}_{\underline{x}} g_i(\underline{x}) \\ + \sum_{j=1}^{m_2} r_{m_1+j} \frac{\mathrm{d}s_{m_1+j}}{\mathrm{d}h_j} \operatorname{grad}_{\underline{x}} h_j(\underline{x}), \quad (41)\end{aligned}$$

sofern man die funktionale Verkettung der Restriktionsfunktionen mit den Verlustfunktionen an dieser Stelle aufhebt. Wird das Optimum erreicht, so gilt ferner

$$\operatorname{grad}_{\underline{x}} q(\underline{x},\underline{r}) = \underline{0}. \quad (42)$$

Werden desweiteren beim Erreichen des globalen Optimums auch noch die Randbedingungen eingehalten, so entspricht dieses der Kuhn-Tucker-Bedingung

$$\begin{aligned}\underline{0} &= \operatorname{grad}_{\underline{x}} L(\underline{x},\underline{u},\underline{v}) \\ &= \operatorname{grad}_{\underline{x}} f(\underline{x}) + \sum_{i=1}^{m_1} u_i \operatorname{grad}_{\underline{x}} g_i(\underline{x}) + \sum_{j=1}^{m_2} v_j \operatorname{grad}_{\underline{x}} h_j(\underline{x}) \quad (43)\end{aligned}$$

mit der dem Problem (8) zugeordneten Lagrange-Funktion

$$L(\underline{x},\underline{u},\underline{v}) = f(\underline{x}) + \sum_{i=1}^{m_1} u_i g_i(\underline{x}) + \sum_{j=1}^{m_2} v_j h_j(\underline{x}) \quad (44)$$

und den nichtnegativen Lagrange-Multiplikatoren $\underline{u}$ sowie den nicht vorzeichenbeschränkten Lagrange-Multiplikatoren $\underline{v}$. Ein Koeffizientenvergleich der Beziehung (43) mit der Beziehung (41) liefert somit

$$u_i = \lim_{k\to\infty} r_i^k \frac{\mathrm{d}s_i(g_i(\underline{x}^k))}{\mathrm{d}g_i} \quad \text{mit} \quad i \in \{1;\ldots;m_1\}, \quad (45)$$

$$v_j = \lim_{k\to\infty} r_{m_1+j}^k \frac{\mathrm{d}s_{m_1+j}(h_j(\underline{x}^k))}{\mathrm{d}h_j} \quad \text{mit} \quad j \in \{1;\ldots;m_2\} \quad (46)$$

als Schätzung für die Lagrange-Multiplikatoren, sofern diese Grenzwerte existieren. Hierzu müssen die Verlustfunktionen s_i in der Umgebung des Nullpunktes stetig differenzierbar sein und ihre Ableitungen müssen dann im zulässigen Gebiet verschwinden. Diesen Anforderungen werden die zuvor gewählten Verlustfunktionen (39) vollkommen gerecht.

Die numerische Auswertung dieser Grenzwerte bereitet jedoch Schwierigkeiten, da Ausdrücke der Art $\infty \cdot 0$ zu bilden sind. Dieser Problematik wird in der programmtechnischen Implementation dadurch begegnet, daß die Lagrange-Multiplikatoren laufend mit den Skalierungsfaktoren r_i aktualisiert werden, sofern die relative Differenz zum vorherigen Wert betragsmäßig kleiner wird als bei der vorherigen Aktualisierung. Wenn diese Differenz größer wird, ist wegen der endlichen Rechengenauigkeit keine Verbesserung des Schätzwertes mehr möglich. Dennoch kann es hierbei zu deutlichen Fehlern kommen, insbesondere wenn eine nicht bindende Restriktion zwischenzeitlich als bindend behandelt worden ist. Diesem Problem kann man sofern möglich mit einem ausreichend groß gewähltem $\underline{r}^0$ entgehen, mit welchem die nicht bindenden Restriktionen von Anfang an erfüllt werden.

4 Anwendung der dargestellten Methode

Die zuvor dargestellte Methode ist in Form der Programmbibliothek `libvssa` realisiert worden. Im folgenden Abschnitt werden mit selbiger einige Beispielaufgaben aus der Literatur gelöst, um die Funktionsweise zu demonstrieren. Die grundsätzlichen Ergebnisse werden abschließend zusammengefaßt. Die Implementationen der Beispielprobleme sowie die Programmbibliothek selbst sind auf der Begleit-CD als Quellkode beigefügt.

4.1 Problem 1

Dieses Problem ist Horst (1999) entnommen worden und dort als Beispiel 2.13 aufgeführt. Es zeichnet sich insbesondere durch eine analytische Lösung inklusive Lagrange-Multiplikatoren aus. Der Quellkode der Implementation ist in der Datei `fuh1.c` zu finden.

Zielfunktion:

$$\min f(x_1; x_2) = 2{x_1}^2 + 2x_1x_2 + {x_2}^2 - 10x_1 - 10x_2. \tag{47}$$

Nebenbedingungen:

$$g_1(x_1; x_2) = {x_1}^2 + {x_2}^2 - 5 \leq 0, \tag{48}$$

$$g_2(x_1; x_2) = 3x_1 + x_2 - 6 \leq 0. \tag{49}$$

Lösung:

	diese Arbeit	analytisch
f	$-20{,}000000$	-20
x_1	1,000000	1
x_2	2,000000	2
u_1	0,999967	1
u_2	0,000000	0

Dieses Ergebnis ergibt sich mit dem Parametersatz

ϵ	α	β	γ	r_i^0	ϱ
10^{-14}	1,0	0,5	2,0	10^3	10,0

und den Startwerten

$$\underline{x}_{\text{min}} = (\ 0{,}0\ \ 0{,}0\ )^{\mathrm{T}},$$
$$\underline{x}_{\text{max}} = (\ 0{,}1\ \ 0{,}1\ )^{\mathrm{T}}$$

nach 11 Iterationen und 1099 Zielfunktionsaufrufen. Es sei nochmals vermerkt, daß die Lagrange-Multiplikatoren nur Beiprodukte sind und in ihrer Genauigkeit nicht kontrolliert werden. Dennoch ergeben sich für selbige akzeptable Schätzwerte.

4.2 Problem 2

Dieses nichtkonvexe Problem ist Gould (1972) Abschnitt 4.5 entnommen. Der Quellkode für das Problem ist in der Datei `GouldNC2.c` enthalten.

Zielfunktion:

$$\min f(x_1; x_2) = (x_1 - 10)^3 + (x_2 - 20)^3. \tag{50}$$

Nebenbedingungen:

$$g_1(x_1; x_2) = -x_1 + 13 \leq 0, \tag{51}$$
$$g_2(x_1; x_2) = -(x_1 - 5)^2 - (x_2 - 5)^2 + 100 \leq 0, \tag{52}$$
$$g_3(x_1; x_2) = (x_1 - 6)^2 + (x_2 - 5)^2 - 82{,}81 \leq 0, \tag{53}$$
$$g_4(x_1; x_2) = -x_2 \leq 0. \tag{54}$$

Lösung:

	diese Arbeit	GOULD (1972)
f	$-6961{,}813876$	$-6961{,}81$
x_1	14,095000	14,095
x_2	0,842961	0,842961
u_1	0,000000	—
u_2	1096,835121	—
u_3	1229,458331	—
u_4	0,000000	—

Dieses Ergebnis ergibt sich mit dem Parametersatz

ϵ	α	β	γ	r_i^0	ϱ
10^{-10}	1,0	0,5	2,0	10^4	10,0

und den Startwerten

$$\underline{x}_{\text{min}} = (\ 20{,}1\ \ 5{,}84\ )^{\mathrm{T}},$$
$$\underline{x}_{\text{max}} = (\ 20{,}2\ \ 5{,}94\ )^{\mathrm{T}}$$

nach 9 Iterationen und 1281 Zielfunktionsaufrufen. Die Lagrange-Multiplikatoren sind trotz fehlender Referenz dennoch mit angegeben worden.

4.3 Problem 3

Dieses Problem ist dem Anhang a in HIMMELBLAU (1972) entnommen und wird dort als Problem 1 bezeichnet. Es zeichnet sich durch eine Gleichheits- und Ungleichheitsrestriktion aus. Der Quellkode für dieses Problem ist in der Datei `Himmelblau1.c` enthalten.

Zielfunktion:

$$\min f(x_1; x_2) = (x_1 - 2)^2 + (x_2 - 1)^2. \tag{55}$$

Nebenbedingungen:

$$g_1(x_1; x_2) = \frac{{x_1}^2}{4} + {x_2}^2 - 1 \leq 0, \tag{56}$$
$$h_1(x_1; x_2) = x_1 - 2x_2 + 1 = 0. \tag{57}$$

Lösung:

	diese Arbeit	HIMMELBLAU (1972)
f	1,393465	1,393
x_1	0,822876	0,823
x_2	0,911438	0,911
u_1	1,844848	—
v_1	1,592782	—

Dieses Ergebnis ergibt sich mit dem Parametersatz

ϵ	α	β	γ	r_i^0	ϱ
10^{-10}	1,0	0,5	2,0	10,0	10,0

und den Startwerten

$$\underline{x}_{\min} = (\ 2{,}0\ \ 2{,}0\ )^{\mathrm{T}},$$
$$\underline{x}_{\max} = (\ 2{,}1\ \ 2{,}1\ )^{\mathrm{T}}$$

nach 9 Iterationen und 826 Zielfunktionsaufrufen. Die Lagrange-Multiplikatoren sind trotz fehlender Referenz erneut mit angegeben worden.

4.4 Problem 4

Dieses Problem ist ebenfalls dem Anhang a in HIMMELBLAU (1972) entnommen und wird dort als Problem 5 bezeichnet. Es zeichnet sich durch zwei Gleichheitsrestriktionen und vorzeichenbeschränkte Variablen aus. Der Quellkode für dieses Problem ist in der Datei `Himmelblau5.c` enthalten.

Zielfunktion:

$$\min f(x_1; x_2; x_3) = 1000 - {x_1}^2 - 2{x_2}^2 - {x_3}^2 - x_1 x_2 - x_1 x_3. \tag{58}$$

Nebenbedingungen:

$$g_i(x_1; x_2; x_3) = -x_i \leq 0 \qquad \text{mit} \quad i \in \{1; 2; 3\}, \tag{59}$$
$$h_1(x_1; x_2; x_3) = {x_1}^2 + {x_2}^2 + {x_3}^2 - 25 = 0, \tag{60}$$
$$h_2(x_1; x_2; x_3) = 8x_1 + 14x_2 + 7x_3 - 56 = 0. \tag{61}$$

Lösung:

	diese Arbeit	HIMMELBLAU (1972)
f	961,715172	961,715
x_1	3,512109	3,512
x_2	0,216989	0,217
x_3	3,552184	3,552
u_1	0,000000	—
u_2	0,000000	—
u_3	0,000000	—
v_1	1,223738	—
v_2	0,274547	—

Dieses Ergebnis ergibt sich mit dem Parametersatz

ϵ	α	β	γ	r_i^0	ϱ
10^{-10}	1,0	0,5	2,0	10^2	10,0

und den Startwerten

$$\underline{x}_{\text{min}} = (\ 2{,}0\ \ 2{,}0\ \ 2{,}0\ )^{\mathrm{T}},$$
$$\underline{x}_{\text{max}} = (\ 2{,}1\ \ 2{,}1\ \ 2{,}1\ )^{\mathrm{T}}$$

nach 8 Iterationen und 1200 Zielfunktionsaufrufen. Die Lagrange-Multiplikatoren sind trotz fehlender Referenz erneut mit angegeben worden und zeigen an, daß die Vorzeichenrestriktionen nicht bindend sind.

4.5 Problem 5

Dieses Problem ist auch dem Anhang a in HIMMELBLAU (1972) entnommen und wird dort als Problem 21 bezeichnet. Es zeichnet sich durch eine kompliziertere Zielfunktion aus. Der Quellkode für dieses Problem ist in der Datei `Himmelblau21.c` enthalten.

Zielfunktion:

$$\min f(x_1; x_2; x_3) = \sum_{i=1}^{99} \left(\mathrm{e}^{-\frac{(a_i - x_2)^{x_3}}{x_1}} - 0{,}01i \right)^2$$
$$\text{mit} \quad a_i = 25 + (-50 \ln 0{,}01i)^{\frac{1}{1,5}}\,. \quad (62)$$

Nebenbedingungen:

$$g_1(x_1; x_2; x_3) = 0{,}1 - x_1 \leq 0, \tag{63}$$

$$g_2(x_1; x_2; x_3) = -x_2 \leq 0, \tag{64}$$

$$g_3(x_1; x_2; x_3) = -x_3 \leq 0, \tag{65}$$

$$g_4(x_1; x_2; x_3) = x_1 - 100{,}0 \leq 0, \tag{66}$$

$$g_5(x_1; x_2; x_3) = x_2 - 25{,}6 \leq 0, \tag{67}$$

$$g_6(x_1; x_2; x_3) = x_3 - 5{,}0 \leq 0. \tag{68}$$

Lösung:

	diese Arbeit	Himmelblau (1972)
f	0,000000	0,0
x_1	50,000000	50,0
x_2	25,000000	25,0
x_3	1,500000	1,5
u_1	0,000000	—
u_2	0,000000	—
u_3	0,000000	—
u_4	0,000000	—
u_5	0,000000	—
u_6	0,000000	—

Dieses Ergebnis ergibt sich mit dem Parametersatz

ϵ	α	β	γ	r_i^0	ϱ
10^{-20}	1,0	0,5	2,0	1,0	10,0

und den Startwerten

$$\underline{x}_{\text{min}} = (\ 100{,}0\ \ 12{,}5\ \ 3{,}0\ )^{\mathrm{T}},$$
$$\underline{x}_{\text{max}} = (\ 100{,}1\ \ 12{,}6\ \ 3{,}1\ )^{\mathrm{T}}$$

sofort ohne eine einzige Iteration, da alle Restriktionen auf Anhieb erfüllt werden. Es erfolgen 546 Zielfunktionsaufrufe. Die Lagrange-Multiplikatoren sind wie erwartet wegen der erfüllten Restriktionen sämtlich gleich null.

4.6 Allgemeine Ergebnisse

Zusätzlich zu den zuvor präsentierten prägnanten Beispielen sind weitere Probleme untersucht worden und auf der Begleit-CD enthalten. Aus diesen Untersuchungen ergeben sich einige Bemerkungen zum dargestellten Verfahren.

Der Parametersatz

ϵ	α	β	γ	r_i^0	ϱ
10^{-10}	1,0	0,5	2,0	1,0	10,0

erweist sich als gute Wahl, sofern keine weiteren Erkenntnisse vorliegen. Hiermit kann aber nicht immer die maximal mögliche Genauigkeit erreicht werden. Zur Erhöhung der Genauigkeit bei gleichzeitiger Reduzierung der Iterationszahl ist r_i^0 gerade so weit zu erhöhen, daß alle nicht bindenden Restriktionen von Anfang an erfüllt sind. Dies ist daran zu erkennen, daß die zugehörigen Lagrange-Multiplikatoren exakt null sind. Eine weitere Erhöhung von r_i^0 darüber hinaus führt aber zu einer starken Verschlechterung der Lösung bis hin zu fehlender Konvergenz. Danach kann man versuchen die Genauigkeit ϵ mit der die Nebenbedingungen erfüllt und das Optimierungsproblem gelöst werden so weit wie möglich zu reduzieren, bis das Verfahren gerade noch konvergiert. Hierbei erhöht sich natürlich die Anzahl der Zielfunktionsaufrufe. In seltenen Fällen kann eine genauere Lösung auch noch durch eine Erhöhung des Kontraktionskoeffizienten β auf Werte von etwa 0,7 erreicht werden. Im allgemeinen stellen die Standardparameter des Variable-Size-Simplex-Algorithmus jedoch eine sehr gute Wahl dar. Sonstige Veränderungen der Parameter scheinen nicht sinnvoll zu sein.

Die Suche nach einer möglichst genauen Lösung wirft in der Praxis jedoch das Problem fehlender Vergleichswerte auf. Die im Rahmen des präsentierten Algorithmus generierten Fehlergrößen beziehen sich in der Regel auf Differenzen im Iterationsprozeß und gestatten somit keine direkten Aussagen bezüglich des tatsächlichen Fehlers. Es hat sich jedoch gezeigt, daß die Minimierung möglichst aller dieser Fehlergrößen tendenziell auf die exakte Lösung führt. Dies gilt auch für die Fehler der geschätzten Lagrange-Multiplikatoren, die jedoch mitunter sehr groß bleiben, da sie keiner eigenen Fehlerkontrolle unterliegen.

5 Zusammenfassung

Im Rahmen dieser Arbeit ist der Variable-Size-Simplex-Algorithmus nach Nelder/Mead (1965) in Verbindung mit einer Straffunktion zur Anwendung gebracht worden, um restringierte Optimierungsprobleme zu lösen. Der heuristische Variable-Size-Simplex-Algorithmus gestattet die Behandlung von Zielfunktionen, ohne besondere Anforderungen beispielsweise an ihre Differenzierbarkeit zu stellen. Die Methode der Straffunktion erlaubt die Berücksichtigung nahezu beliebiger Nebenbedingungen, sofern diese mit stetigen Funktionen beschreibbar sind. Für den iterativen Lösungsprozeß ist ferner die Kenntnis von zulässigen Startwerten nicht erforderlich. Dieses stellt einen großen Vorteil der präsentierten Methode dar, da das Auffinden von zulässigen Startwerten bei komplexer strukturierten Randbedingungen aufwendig sein kann. Die wesentliche Beschränkung der dargestellten Methode liegt lediglich in der Notwendigkeit begründet, daß alle Funktionen im gesamten $\mathbb{R}^n$ definiert sein müssen. Dies kann bei experimentellen Prozeßfunktionen durchaus problematisch sein.

Nach einer allgemeinen Einführung sind dann zunächst die Grundlagen des Variable-Size-Simplex-Algorithmus dargestellt worden. Anschließend ist eine Erörterung der Straffunktionsmethode erfolgt, wobei die prinzipielle Funktionsweise näher dargelegt worden ist. Auf dieser theoretischen Grundlage ist die Programmbibliothek `libvssa` entwickelt und in der Programmiersprache C implementiert worden. Diese Programmbibliothek enthält alle dargestellten Algorithmen und kann zur restringierten Optimierung eingesetzt werden. Die Anwendung ist dann anhand einiger Beispiele aus der Literatur demonstriert worden, die hierzu ebenfalls in der Programmiersprache C implementiert worden sind. Den Quellkode entnehme man der Begleit-CD, auf welcher auch noch weitere nicht explizit angesprochene Beispiele enthalten sind.

Die Anwendungen haben aufgezeigt, daß sich der Variable-Size-Simplex-Algorithmus mit Straffunktion unmittelbar zur Lösung nichtkonvexer, nichtlinearer Optimierungsprobleme einsetzen läßt, ohne daß die beteiligten Funktionen differenzierbar sein müßten. Der Parametersatz

ϵ	α	β	γ	r_i^0	ϱ
10^{-10}	1,0	0,5	2,0	1,0	10,0

erweist sich grundsätzlich als gute Wahl. Die optimalen Parameter können jedoch in den Bereichen

ϵ	α	β	γ	r_i^0	ϱ
$10^{-20} \ldots 10^{-6}$	1,0	0,5…0,8	2,0	$10^0 \ldots 10^4$	2,0…100,0

liegen. So kann bei einem zu kleinen Kontraktionskoeffizienten β das Optimum verfehlt werden, wohingegen ein zu großer die Konvergenz verhindert. Die initialen Skalierungsparameter r_i^0 sollten groß genug gewählt werden, um die nicht bindenden Restriktionen von Anfang an zu vermeiden, was die Genauigkeit der Lösung erhöht. Bei einer zu großen Wahl kann das Optimum mitunter jedoch nicht mehr erreicht werden. Der Genauigkeitsparameter ϵ definiert die einzuhaltende Fehlerschranke der Nebenbedingungen sowie das Abbruchkriterium des Variable-Size-Simplex-Algorithmus. Er sollte mithin so klein wie möglich gewählt werden, sofern dies zu Lasten der Iterationszahl möglich ist. Die untere Schranke für ϵ liegt dabei in etwa bei der Rechengenauigkeit der verwendeten `double`-Variablen und zeigt sich im Ausbleiben der Konvergenz. Falls die Straffunktionsiteration ordnungsgemäß endet, so werden die Nebenbedingungen jeweils garantiert auf ϵ genau erfüllt. Ob tatsächlich das Optimum erreicht wurde, ist durch Variation der Parameter insbesondere von β validierbar.

Der Variable-Size-Simplex-Algorithmus mit Straffunktion stellt also ein probates Mittel zur Lösung nahezu beliebiger Optimierungsprobleme dar. Eine geeignete Parameterwahl ist jedoch im voraus nicht möglich. Um das Optimum tatsächlich zu erreichen, sind mitunter mehrere Lösungsläufe mit unterschiedlichen Parametersätzen notwendig. Die Nebenbedingungen werden jedoch immer bei ordnungsgemäßer Beendigung garantiert auf ϵ genau eingehalten. Die Startwerte sind nur hinsichtlich der Iterationszahl von Bedeutung und müssen keinen sonstigen Bedingungen genügen, sofern der Startsimplex ein Hypervolumen größer null besitzt. Weiterführend könnte es von Interesse sein, die Straffunktion durch eine Barrierefunktion zu ersetzen, um Optimierungsprobleme mit Funktionen auf eingeschränkten Definitionsbereichen zu lösen. Hierbei muß zuvor jedoch auch eine zulässige Lösung ermittelt werden und sichergestellt sein, daß der Variable-Size-Simplex-Algorithmus keinesfalls den Definitionsbereich verläßt.

Literatur

Avriel, Mordecai (1976): *Nonlinear Programming: Analysis and Methods*, Automatic Computation, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Bracken, Jerome und McCormick, Garth P. (1970): *Ausgewählte Anwendungen nichtlinearer Programmierung*, Verlag Berliner Union GmbH, Verlag W. Kohlhammer GmbH.

Bronstein, I. N. und Semendjajew, K. A. (1989): *Taschenbuch der Mathematik*, Verlag Harri Deutsch, Thun, 24. Aufl.

Bronstein, I. N. und Semendjajew, K. A. (1991): *Ergänzende Kapitel zu Taschenbuch der Mathematik*, Verlag Harri Deutsch, Thun, 6. Aufl.

Courant, R. (1943): *Variational Methods for the Solution of Problems of Equilibrium and Vibrations*, Bull. Amer. Math. Soc., 49:1–23.

Elster, Karl-Heinz (1978): *Nichtlineare Optimierung*, Bd. 15 d. Reihe *Mathematik für Ingenieure, Naturwissenschaftler, Ökonomen und Landwirte*, BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1. Aufl.

Fiacco, Anthony V. und McCormick, Garth P. (1968): *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley and Sons, Inc.

Gill, Philip E.; Murray, Walter und Wright, Margaret H. (1989): *Practical Optimization*, Academic Press Limited, 24/28 Oval Road, London NW1 7DX.

Gould, F. J. (1972): *Non-linear Tolerance Programming*, in: Lootsma, F. A.; Hrsg.: *Numerical Methods for Non-linear Optimization*, Kap. 24, S. 349–366, Academic Press Inc. (London) Ltd., 24/28 Oval Road, London NW1.

Henoch, Mark (2001): *Der Variable-Size Simplex Algorithmus*, Fernuniversität Hagen (Seminararbeit).

Himmelblau, David M. (1972): *Applied Nonlinear Programming*, McGraw-Hill, Inc.

Horst, Reiner (1999): *Nichtlineare Optimierung*, Fernuniversität Hagen, Kurs 0854.

NELDER, J. A. und MEAD, R. (1965): *A simplex method for function minimization*, The Computer Journal, 7:308–313.

NEUMANN, KLAUS und MORLOCK, MARTIN (1993): *Operations Research*, Carl Hanser Verlag München Wien.

PRESS, WILLIAM H.; TEUKOLSKY, SAUL A.; VETTERLING, WILLIAM T. und FLANNERY, BRIAN P. (1992): *Numerical Recipes in FORTRAN*, Cambridge University Press, 2. Aufl.

RYAN, D. M. (1974): *Penalty and Barrier Functions*, in: GILL, P. E. und MURRAY, W.; Hrsg.: *Numerical Methods for Constrained Optimization*, Kap. VI, S. 175–190, Academic Press Inc. (London) Ltd., 24/28 Oval Road, London NW1.

SPENDLEY, W.; HEXT, G. R. und HIMSWORTH, F. R. (1962): *Sequential Application of Simplex Designs in Optimisation and Evolutionary Operation*, Technometrics, 4(4):441–461.

SWANN, W. H. (1974): *Constrained Optimization by Direct Search*, in: GILL, P. E. und MURRAY, W.; Hrsg.: *Numerical Methods for Constrained Optimization*, Kap. VII, S. 191–217, Academic Press Inc. (London) Ltd., 24/28 Oval Road, London NW1.

WOLFE, MICHAEL ANTHONY (1978): *Numerical Methods for Unconstrained Optimization*, Van Nostrand Reinhold.

ZANGWILL, WILLARD I. (1967): *Non-Linear Programming via Penalty Functions*, Management Science, 13A(5):344–358.

ZANGWILL, WILLARD I. (1969): *Nonlinear Programming: A Unified Approach*, Prentice-Hall, Inc., Englewood Cliffs, N. J.

Erklärung

Ich versichere, daß ich diese Diplomarbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Die Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen.

Bremen, den 25. Mai 2002

(Dr.-Ing. Ingo D. Rullhusen)