

# Optimierungsalgorithmen - Deterministische Globale Optimierung

Tim Dafler

## ZUSAMMENFASSUNG

Der Zweig der Deterministischen Globalen Optimierung befasst sich mit Problemen, bei denen exakte Verfahren für die Lösung notwendig sind. Anstelle einer lokal optimalen Lösung sind globale Optima von Interesse. Anwendungen findet die Theorie unter anderem in Prozessen der Chemie, bei der chemische Gleichgewichte eine exakte Bestimmung des Optimalwerts erfordern. Aber auch in der Robotik, in Netzwerken und bei der Strukturvorhersage in Clustern findet die Theorie praktische Anwendung. Die Arbeit befasst sich mit den grundlegenden Ansätzen und Vorgehensweisen um die Lösung des vorliegenden Problems möglichst effizient zu berechnen. Hierfür werden verschiedene Verfahren des Branching, Branch-and-Bound und Branch-and-Reduce aufgezeigt und deren praktischer Nutzen für die Abschätzung von Schranken dargelegt. Anschließend werden Anwendungen und Software-Lösungen für die Optimierungsprobleme vorgestellt, ehe in einer kurzen Diskussion die Ergebnisse und Erkenntnisse zusammengetragen werden. Geschlossen wird mit einem kurzen Ausblick der weiteren Entwicklung.

## KEYWORDS

**Operations Research, Optimierungsprobleme, Deterministische Globale Optimierung, Branching, DIRECT, MCL, LGO, Branch-and-Bound, Branch-and-Reduce, Constraint Propagation, Konvexe Relaxierung, Intervallarithmetik, Lagrange-Dualität**

## 1 MOTIVATION

Tagtägliches Verhalten und Handeln ist stets geprägt von zielgerichteten, rationalen Entscheidungsprozessen, welche das Ziel der Ermittlung optimaler Handlungsstrategien verfolgt [4, S. 1][5, S. 1][23, S. 1].

Der Wissenszweig, welcher sich mit der Lösung und Untersuchung in der Praxis, unter Verwendung systematisch-methodischen Vorgehens beschäftigt, wird als **Operations Research (OR)** bezeichnet [5, S. 1][21, S. 1]. Als interdisziplinäre Wissenschaft verbindet Operations Research hierfür Methoden und Anwendungen der Mathematik, der Wirtschaftswissenschaften, der Informatik und der Ingenieurwissenschaften [21, S. 1][23, S. 1].

Das Vorgehen des Entscheidungsfindungsprozesses ist in Abbildung 1 skizziert. Die in der Realität zu treffende Entscheidung wird zunächst in ein vereinfachtes Modell übertragen. Dieses beinhaltet nunmehr nur die erforderlichen Aspekte und stellt eine Abstraktion der Realität dar. Durch Anwendung quantitativer Optimierungsverfahren kann eine gute oder etwaig optimale Lösung bestimmt werden. Diese wird folgend für die reale Umsetzung interpretiert, ehe sie schließlich Anwendung findet. [4, S. 2-3][5, S. 1-2]

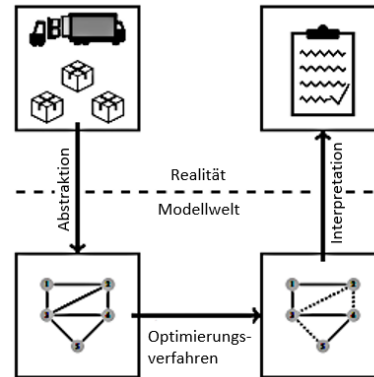


Abbildung 1: Prozess der Entscheidungsfindung [4, S. 2]

Die Optimalität der Lösung wird unter anderem durch Zeitanforderungen und Rechenaufwand beeinflusst [4, S. 3][5, S. 2]. Anstelle der besten Lösung kann aus Gründen der Effizienz daher auch eine hinreichend optimale Lösung gewählt werden [5, S. 2][23, S. 1]. Hierbei kommt das ökonomische Prinzip zu tragen, welches in Form des Maximum- oder Minimum-Prinzips auftritt [23, S. 1].

Das **Maximum-Prinzip** beschreibt das Streben mit gegebenem Einsatz an Gütern ein maximales Ergebnis zu erreichen [23, S. 1].

Das **Minimum-Prinzip** hingegen ersucht das Erreichen eines bestimmten Ergebnisses unter minimalem Einsatz begrenzter Güter [23, S. 1].

Die Verfahren für Optimierungsprobleme, und somit zur Bestimmung der Lösung, teilen sich insgesamt in **drei unterschiedliche Typen** auf [4, S. 10]:

- Optimale / Exakte Verfahren
- Approximationsverfahren
- Heuristiken

**Optimale oder exakte Verfahren** garantieren zu jeder Problem Instanz eine optimale Lösung, sind zumeist aber mit viel Aufwand verbunden. [4, S. 10]

**Approximationsverfahren** dagegen finden potentielle suboptimale Lösungen, die zu einem bestimmten Güte-Verhältnis zum Wert der optimalen Lösung stehen. Sie garantieren für jede Instanz eine beschränkte maximale Fehlerabweichung (Worst Case Abschätzung). [4, S. 10][5, S. 135]

**Heuristiken** ergeben zu einer Instanz potentiell beliebig schlechte Lösungswerte, benötigen aber geringsten Aufwand. Im Gegensatz zu Approximationsverfahren können sie keine bestimmte Güte der Lösung garantieren. [4, S. 10-11]

Operations Research gliedert sich unter anderem in **folgende Teilgebiete** [8, S. 1][21, S. 2]:

- lineare und nicht-lineare Optimierung
- ganzzahlige und kombinatorische Optimierung
- konvexe Optimierung
- differenzierbare Optimierung
- evolutionäre Algorithmen

Die Teilgebiete betrachten jeweils unterschiedliche Modelle, welche beispielsweise in den Zielfunktion variieren können [5, S. 8-9].

In den folgenden Abschnitten werden **deterministische globale Optimierungsmethoden** aus den Bereichen der **linearen und nicht-linearen Optimierung** sowie der **ganzzahligen linearen und nicht-linearen Optimierung** genauer untersucht, sowie praktische Anwendungen aufgezeigt.

**Deterministisch** bedeutet in diesem Zusammenhang, dass die angewandten Methoden in endlicher Zeit eine Approximation an ein globales Optimum, innerhalb vorgegebener Toleranzen, liefern können [15, S. 2]. Das Gebiet der deterministischen globalen Optimierung fokussiert sich somit auf das Finden eines globalen Optimums einer nicht-konvexen Zielfunktion  $f$  unter bestimmten Bedingungen  $R$ , auch Restriktionen oder Nebenbedingungen genannt [7, S. 2][13, S. 11]. Es handelt sich somit um optimale/exakte Verfahren und Approximationsverfahren, wobei die Methoden von Neumaier auch alternativ mit folgenden Klassen beschrieben werden [18, S. 275]:

- **Vollständige Verfahren** erreichen ein globales Minimum bei exakten Berechnungen und unendlicher Laufzeit mit Sicherheit, wobei nach endlicher Zeit bekannt ist, ob ein globales Minimum gefunden wurde.
- **Rigorose Verfahren** erreichen ein globales Minimum innerhalb vorgegebener Toleranzen mit Sicherheit, selbst unter Einbezug von Rundungsfehlern (ausgenommen entartete Fälle, in denen Toleranzen überschritten werden).

Im nächsten Abschnitt wird das typische Problem der Deterministischen Globalen Optimierung formuliert.

## 2 PROBLEME DER DETERMINISTISCHEN GLOBALEN OPTIMIERUNG

Für die unterschiedlichen Methoden wird folgende Form der Optimierungsprobleme zugrundegelegt (vgl. [18, S. 282]):

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{g}(\mathbf{x}) \in \mathbf{F} \quad (1) \\ & \mathbf{x} \in \mathbf{X} \end{array}$$

mit  $f$  und  $g$  Funktionen mit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  und  $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

$\mathbf{X}$  ist beschränkte oder unbeschränkte Box in  $\mathbb{R}^n$ , das heißt  $\mathbf{X} = [\underline{\mathbf{x}}, \bar{\mathbf{x}}]$  mit  $\underline{\mathbf{x}} \in (\mathbb{R} \cup \{-\infty\})^n$  und  $\bar{\mathbf{x}} \in (\mathbb{R} \cup \{-\infty\})^n$ .

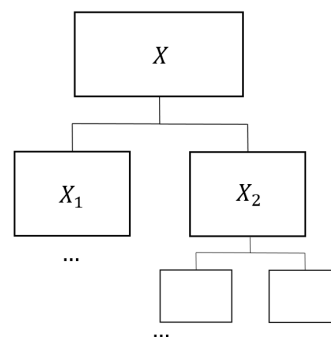
Während  $f$  die Zielfunktion darstellt, bildet  $g$  einen Vektor von  $m$  Nebenbedingungen (Gleichungs- und Ungleichungsform), wobei durch  $\mathbf{F}$  eine Box im  $\mathbb{R}^m$  repräsentiert wird, welche die Bedingungen von  $g$  definiert [18, S. 282]. Die zulässige Lösungsmenge entspricht dann  $S := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \in \mathbf{F}, \mathbf{x} \in \mathbf{X}\}$ , sowie deren beinhaltenden  $\mathbf{x} \in S$  zulässigen Punkten (vgl. [18, S. 282]).

Die Betrachtung eines Minimierungsproblems stellt hierbei keinerlei Einschränkung dar, da jedes Maximierungsproblem in ein duales Minimierungsproblem umgeformt werden kann [13, S. 11-12].

Die drei folgenden Abschnitte zeigen unterschiedliche Methoden zur Lösung der Deterministischen Globalen Optimierung.

## 3 REINES BRANCHING

Reine Branching-Methoden sind die einzigen Verfahren, welche auch bei lediglich lokalen Informationen funktionieren [18, S. 299]. Hierbei wird die Idee verfolgt, dass die Ermittlung einer global optimalen Lösung in der gesamten Box  $\mathbf{X}$  zu kompliziert ist, und diese daher in Teilboxen unterteilt wird (siehe Abbildung 2) [4, S. 179]. Diese Teilboxen werden dahingehend untersucht ob lokal optimale Lösungen vorliegen oder nicht und gegebenenfalls iterativ weiter aufgeteilt [4, S. 179-182]. Ein Durchsuchen aller Teilboxen liefert dann die global optimale Lösung [4, S. 179]. Die Darstellung des Verfahrens erfolgt üblicherweise in einem Baumdiagramm (siehe Abbildung 2) [4, S. 180].



**Abbildung 2: Aufteilung der Box  $\mathbf{X}$  in Teilboxen**

Mithilfe des später betrachteten Boundings können die zu durchsuchenden Teilboxen eingeschränkt werden, was in einem Effizienzgewinn resultiert [4, S. 179].

Als lokale Informationen für die Branching-Methoden werden zum Beispiel folgende Punktauswertungen verwendet [18, S. 299]:

$$(f(\mathbf{x}), g(\mathbf{x}), \nabla f(\mathbf{x}), \nabla g(\mathbf{x}), \nabla^2 f(\mathbf{x}), \nabla^2 g(\mathbf{x}), \dots)$$

Eine notwendige und hinreichende Bedingung für vollständige Methoden basiert auf dem Dichtheitstheorem von Törn und Zinlaskas (vgl. [18, S. 299]):

### Theorem 3.1 Dichttheoreme von Törn und Zinlinskas

Jede Methode, welche auf lokalen Informationen basiert und für jede stetige Funktion  $f$  gegen ein globales Minimum innerhalb eines zulässigen Bereiches  $C$  konvergiert, muss eine Folge von Punkten  $(x_i)_{i \in \mathbb{N}}$  erzeugen, die dicht in  $C$  liegt.

Es gilt:

$$\liminf_{i \rightarrow \infty} f(x_i) = \min\{f(x) \mid x \in C\}$$

Erfüllt eine globale Optimierungs-Methode, welche auf lokalen Informationen basiert, das Dichttheorem, so heißt die Methode konvergent. Hierbei gilt, dass Konvergenz nur minimale Anforderung ist und neben dieser auch die Berechnungszeit relevant ist. [18, S. 299]

Eine mögliche Vorgehensweise ist die **Gittersuche**, bei der über eine rechteckige Domäne ein Gitter mit äquidistanten Abständen errichtet wird. Die Gitterpunkte werden anschließend evaluiert, wobei der beste Punkt eine gute Approximation des globalen Minimums der Box  $X$  darstellt. Ein Problem des Verfahrens ist allerdings, dass der gefundene Punkt gegebenenfalls stark vom globalen Minimum abweichen kann, wenn die Region innerhalb des Gitters eine Nadelgestalt aufweist. Des Weiteren steigt die Anzahl der Gitterpunkte mit ansteigender Dimension exponentiell, womit die Gittersuche ab Dimension  $\dim > 2$  keinen effektiven Lösungsansatz darstellt. [9, S. 74]

Gute allgemeine, asymptotische, globale Optimierungsalgorithmen, welche sich auf lokalen Informationen stützen, sind hingegen [18, S. 299]:

- DIRECT
- MCS
- LGO

Während LGO Probleme der oben vorgestellten Form 1 betrachtet, behandeln DIRECT und MCS Probleme folgender einfacherer Form [10, S. 2][11, S. 159]:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in X \end{array}$$

Neben den Unterschieden in der betrachteten Problemstellung gibt es außerdem Differenzen in der Art und dem Zeitpunkt der Aufteilung der Boxen sowie der Vorgehensweise innerhalb jeder entstehenden Box. Die Konvergenz wird hierbei durch die angewendeten Branching-Schemata erzwungen, welche eine Folge von Bäumen von Boxen generieren, deren Blätter die gesamte Box  $X$  des Problems abdecken. Für jede der entstehenden Boxen bzw. Blätter wird mindestens ein innerer Punkt ausgewertet. Die Bäume selbst entstehen iterativ durch Aufteilen der Ursprungsbox in Blätter, welche wiederum in Blätter zerfallen können. Die Konvergenz des Durchmessers wird durch geeignete Aufteilungsregeln gesichert, welche schließlich die Konvergenz des Verfahrens induzieren. Neben Konvergenz ist für die Effizienz auch ein ausgewogenes Gleichgewicht zwischen globaler und lokaler Suche notwendig. [18, S. 300]

### 3.1 DIRECT

DIRECT, ein Akronym für **D**ividing **R**ectangles, erreicht dies indem in jeder Iteration alle Boxen, welche nicht durch eine andere Box dominiert werden, aufgesplittet werden (siehe Abbildung 3) [18, S. 300]. Die Boxen  $B_i$  werden hierbei durch Paare  $(v_i, f_i)$  aus dem Funktionswert  $f_i$  am Mittelpunkt der Box und dem Volumen  $v_i$  beschrieben [18, S. 300].

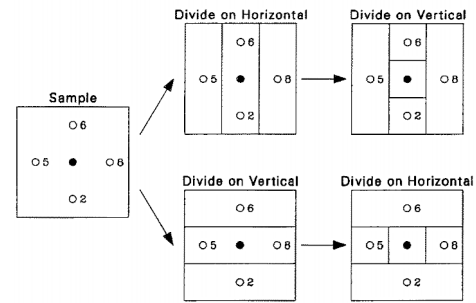


Abbildung 3: Splitting der Box nach DIRECT [11, S. 168]

Es gilt [18, S. 300]:

**$(v, f)$  wird von  $(v', f')$  dominiert, wenn  $v' < v$  und  $f' > f$ .**

Dies bedeutet insbesondere, dass die Boxen mit dem größten Volumen oder dem besten Funktionswert niemals dominiert werden und somit stets aufgeteilt werden [18, S. 300].

### 3.2 MCL

MCL (**M**ultivlevel **C**oordinate **S**earch) zielt ebenfalls auf eine Balance zwischen globaler (Splitting von Boxen mit großen ununtersuchten Gebieten) und lokaler Suche (Splitting von Boxen mit guten Funktionswerten) ab, wobei hierfür ein Multilevel-Ansatz Anwendung findet. Hierbei wird jeder Box eine Zahl, genannt Level, von durchgeführten Teilungsoperationen zugewiesen, wobei eine Maximalzahl festgelegt wird, bei der eine weitere Teilung nicht sinnvoll erscheint, da die Box zu klein wird. Das heißt das Level  $l$  einer Box entspricht einer Zahl aus der Menge  $\{0, 1, \dots, s_{\max}\}$ . Wird eine Box mit Level  $0 < l < s_{\max}$  geteilt, wird ihr Level auf  $l = 0$  gesetzt und die Nachkommen erhalten das Level  $l + 1$  oder  $\min(l + 2, s_{\max})$ , abhängig davon, ob die Teilung der Nachkommen sinnvoll ist oder nicht. Die globale Suche entspricht der Auswahl der Box mit kleinstem Level, während die lokale Suche die Wahl des kleinsten Funktionswertes umsetzt. [10, S. 4-11]

Es gilt somit folgende Ordnungsrelation [10, S. 7-8]:

$$(l, f) < (l', f') : \Leftrightarrow f < f' \text{ und } l < l'$$

Die Umsetzung des Splittings erfolgt durch sogenannte Sweeps, wobei die ungeteilten Boxen eines Levels als mögliche Split-Kandidaten untersucht und gegebenenfalls geteilt werden. Die Teilung erfolgt indem bekannte Punkte quadratisch interpoliert werden und eine Aufteilung der Box am Minimum erfolgt. [10, S. 8-11]

### 3.3 LGO

LGO (Lipschitz Global Optimization) hat wenig bis keine analytischen Informationen über die Funktion  $f$ , womit die Funktionswerte eine Art Black-Box bilden. Einzige Annahme über die Funktion ist, dass die Variabilität in Bezug auf die Eingabevariable  $x$  beschränkt ist. Die Annahme wird häufig durch grundlegende Informationen und Kenntnisse über das betreffende System gestützt. Wird die Annahme nicht erfüllt, so gibt es keine Garantie, dass optimale Schätzungen berechnet werden. [19, S. 15]

Für die Entscheidung, welche Box gesplittet werden soll, werden bei LGO untere Schranken für die Lipschitz-Konstanten  $L$  aus vorangehenden Funktionsauswertungen berechnet [18, S. 300]:

$$L \geq \max_{l,k} \frac{\|f(x_k) - f(x_l)\|}{\|x_k - x_l\|}$$

Der Ansatz hierfür ist, dass aufgrund der Lipschitz-Bedingung ein feines Gitter auf der Box  $X$  konstruiert werden kann, sodass der minimale Funktionswert, der durch die Punkte des Gitters erreicht wird, in einer  $\varepsilon$ -Abschätzung des Optimums liegt [19, S. 16].

## 4 BRANCH-AND-BOUND

Branch-and-Bound ist im Vergleich zu reinem Branching eine Erweiterung, welche mittels Bounding die zu betrachtenden Boxen einschränkt und somit eine höhere Effizienz liefert [9, S. 159].

Zunächst wird eine Zerlegung der Ursprungsbox  $X$  des Problems durchgeführt, um auf den entstehenden Teilboxen lokal optimale Lösungen zu ermitteln [4, S. 179]. Hierbei wird eine Liste aller mit Branching erhaltenden Boxen geführt und für jede Box  $B_i$  eine untere Schranke  $f_{\text{lower}}$  bestimmt, welche den Minimalwert der Zielfunktion in der jeweiligen Box  $B_i$  darstellt [9, S. 159]. Außerdem wird die globale obere Schranke  $f_{\text{upper}}$ , welche den aktuellen Minimalwert der Zielfunktion der besten bisherigen Lösung beschreibt, notiert. Mithilfe der oberen Schranke können in jeder Iteration des Verfahrens diejenigen Boxen aus der Liste entfernt werden, die eine höhere untere Schranke aufweisen, das heißt für die  $f_{\text{lower}} > f_{\text{upper}}$  gilt [9, S. 159]. Die Performance des Algorithmus wird stark von der Wahl der Box beeinflusst, wobei nach Bestensuche (Box mit kleinster Schranke  $f_{\text{lower}}$ ) oder Breitensuche (Wahl der am wenigsten untersuchten Box) vorgegangen werden kann [9, S. 160]. Nach erfolgreicher Durchführung ist die Liste leer und es wurde entweder ein globales Optimum gefunden oder es existiert keine zulässige Lösung [9, S. 160].

## 5 BRANCH-AND-REDUCE

Für eine Verbesserung des Dichtheitstheorems ist eine weitere Einschränkung der zulässigen Domäne hilfreich, da somit irrelevante Gebiete, welche keine global optimale Lösung beinhalten, ausgeschlossen werden können [18, S. 301].

Hierfür wird vom Zusammenspiel folgender unterschiedlicher Bereiche Gebrauch gemacht [18, S. 301]:

- Constraint Propagation
- Konvexe Relaxierung
- Intervallanalyse
- Dualitätsargumente

## 5.1 Constraint Propagation

Constraint Propagation garantiert eine rigorose Reduktion ohne Verlust möglicher Lösungskandidaten. Die Methode basiert auf subdefiniten Berechnungen, das heißt um eine gewünschte Gleichung oder Ungleichung zu lösen wird die Restriktion korrespondierend zum Wert des Ausdrucks durch den Ausdruck propagiert. [14, S. 6]

**Hierzu folgende Aufgabe [14, S. 6]:**

Sei die Ungleichung

$$3x - 2 \geq 5$$

gegeben, finde  $x \in \mathbb{R}$  welche die Ungleichung erfüllen!

**Vorgehen:**

Das zulässige Intervall von Funktionswerten schränkt sich auf  $[5, \infty]$  ein. Eine Propagation durch den Ausdruck  $3x - 2$  berechnet die zulässigen Werte von  $x$  als  $[7/3, \infty]$ . Die Darstellung des Vorgehens erfolgt üblicherweise in einem Berechnungsbaum (siehe Abbildung 4). [14, S. 6]

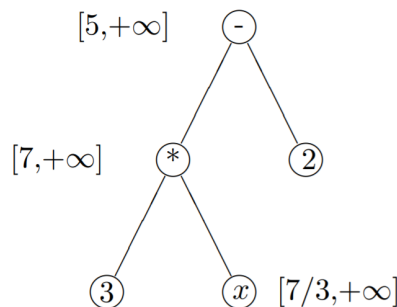


Abbildung 4: Berechnungsbaum [14, S. 6]

Im Kontext globaler Optimierung wird Constraint Propagation verwendet um die Menge von stationären Punkten zu finden, das heißt Lösungen der Gleichung  $f'(x) = 0$ . Die Gleichung wird als Propagated Constraint bezeichnet. Zu Beginn seien die zulässige Menge von Funktionswerten, also das Intervall der Wurzel des Berechnungsbaums, welche  $f'(x)$  repräsentiert, mit  $[0, 0]$  initialisiert und die Wertebereiche der unabhängigen Variablen gesetzt. Anschließend werden den Knoten Intervallwerte zugewiesen, wobei bei fehlenden Informationen  $[-\text{inf}, +\text{inf}]$  verwendet wird. Der Berechnungsbaum kann nun durchlaufen und die Wertebereiche der jeweiligen Knoten aktualisiert werden. Wenn keine weitere Änderung erfolgt, terminiert das Verfahren und man erhält eine Einschränkung der zu betrachtenden Domäne. [14, S. 6]

## 5.2 Konvexe Relaxierung

Konvexe Relaxierung dient der Berechnung guter unterer Schranken für die Zielfunktion auf (Teil-)Mengen, zum Beispiel auf den bei Branching entstehenden Boxen [22, S. 112]. Durch konvexe Relaxierungen können Zusammenhänge von nichtkonvexen und konvexen Mengen und Funktionen hergestellt werden, mit deren Hilfe untere Schranken effizient berechenbar sind [22, S. 120].

Konvex relaxierte Mengen werden wie folgt definiert (vgl. [22, S. 116]):

### 5.1.1 Definition Konvex relaxierte Menge

Es sei  $\emptyset \neq M \subseteq \mathbb{R}^n$ , dann gilt :

- Jede konvexe Menge  $\hat{M} \subseteq \mathbb{R}^n$  mit  $M \subseteq \hat{M}$  heißt **konvexe Relaxierung auf M**.
- Der Durchschnitt aller konvexen Relaxierungen von M

$$\hat{M} := \bigcap \{ \hat{M} \mid \hat{M} \supseteq M, \hat{M} \text{ konvex} \}$$

heißt **konvexe Hülle** von M.

Das Pendant für konvex relaxierte Funktionen lautet folgendermaßen (vgl. [22, S. 116]):

### 5.2.2 Definition Konvex relaxierte Funktion

Es sei  $\emptyset \neq M \subseteq \mathbb{R}^n$  und  $f : M \rightarrow \mathbb{R}$ , dann gilt :

- Jede auf M konvexe Funktion  $\hat{f}$  mit

$$\forall x \in M : \hat{f}(x) \leq f(x)$$

heißt **konvexe Relaxierung von f auf M**.

- Eine konvexe Relaxierung  $\hat{f}$  von f auf M, die für alle anderen konvexen Relaxierungen von f auf M

$$\forall x \in M : \hat{f}(x) \leq \hat{\hat{f}}(x)$$

erfüllt, heißt **konvexe Hüllfunktion** von f auf M.

Der Beweis des nachfolgenden Satzes zeigt, dass eine Konstruktion konvexer Relaxierungen von funktional beschriebenen Mengen möglich ist, während die Übertragung des Verfahrens auf konvexe Hüllen nicht funktioniert. Ziel ist es, Aussagen über nichtkonvexe Optimierungsprobleme zu treffen, indem die Probleme mit konvex relaxierter Zielfunktion und konvex relaxierter zulässiger Menge betrachtet werden. [22, S. 117-118]

### 5.2.3 Satz (vgl. [22, S. 117])

Es seien  $\emptyset \neq M \subseteq \mathbb{R}^n$  konvex (z.B.  $M = \mathbb{R}^n$ ) und  $g_i : M \rightarrow \mathbb{R}, i \in I$ , beliebig und  $S = \{x \in M \mid g_i(x) \leq 0, i \in I\}$  dann gilt:

- Falls für jedes  $i \in I$  die Funktion  $\hat{g}_i$  eine konvexe Relaxierung von  $g_i$  auf M ist, dann ist die Menge

$$\hat{S} = \{x \in M \mid \hat{g}_i(x) \leq 0, i \in I\}$$

eine konvexe Relaxierung von S.

- Selbst wenn für jedes  $i \in I$  die Funktion  $\hat{g}_i$  die konvexe Hüllfunktion von  $g_i$  auf M ist, kann trotzdem

$$\hat{S} \neq \{x \in M \mid \hat{g}_i(x) \leq 0, i \in I\}$$

gelten.

Der Beweis hierfür wird in [22, S. 117-118] erbracht.

Das konvex relaxierte Optimierungsproblem wird, wie folgt, definiert (vgl. [22, S. 118]):

### 5.2.4 Definition Konvex relaxiertes Optimierungsproblem

Für eine nichtleere konvexe Menge  $M \subseteq \mathbb{R}^n$  seien eine beliebige Funktion  $f : M \rightarrow \mathbb{R}$  und eine beliebige Menge  $X \subseteq M$  sowie das Optimierungsproblem

$$P : \min f(x)$$

$$\text{s.t. } x \in X$$

gegeben.

- Die Funktion  $\hat{f}$  sei eine konvexe Relaxierung von f auf M, und  $\hat{X} \subseteq M$  sei eine konvexe Relaxierung von X. Dann heißt

$$\hat{P} : \min \hat{f}(x)$$

$$\text{s.t. } x \in \hat{X}$$

konvexe Relaxierung von P (auf M).

- Die konvexe Hüllfunktion  $\hat{\hat{f}}$  von f auf M existiere, und  $\hat{\hat{M}}$  sei die konvexe Hülle von X. Dann heißt

$$\hat{\hat{P}} : \min \hat{\hat{f}}(x)$$

$$\text{s.t. } x \in \hat{\hat{X}}$$

konvexes Hüllproblem von P (auf M).

Mit den Voraussetzungen der vorangegangenen Definition ergibt sich folgender Satz [22, S. 119]:

### 5.2.5 Satz (vgl. [22, S. 119])

Seien die Voraussetzungen aus Definition 5.2.4 erfüllt, sowie die Optimierungsprobleme P,  $\hat{P}$  und  $\hat{\hat{P}}$  lösbar mit Minimalwerten v,  $\hat{v}$  bzw.  $\hat{\hat{v}}$ , dann gelten:

- Für den Minimalwert  $\hat{v}$  jeder konvexen Relaxierung  $\hat{P}$  von P gilt  $\hat{v} \leq v$ .
- Für den Minimalwert  $\hat{\hat{v}}$  des konvexen Hüllproblems  $\hat{\hat{P}}$  von P und den Minimalwert  $\hat{v}$  jeder konvexen Relaxierung  $\hat{P}$  gilt  $\hat{v} \leq \hat{\hat{v}} \leq v$  (das heißt  $\hat{\hat{v}}$  ist die beste per konvexer Relaxierung erzielbare Unterschranke von v).
- Im Allgemeinen gilt nicht notwendigerweise  $\hat{\hat{v}} = v$ .
- Falls X konvex ist, gilt  $\hat{v} = v$ .
- Falls f linear ist, gilt  $\hat{v} = v$ .

Der Beweis hierfür kann in [22, S. 119] nachgeschlagen werden.

Für die Berechnung unterer Schranken ist die Relation  $\hat{v} \leq v$  wesentlicher Bestandteil. Die Konvexität der Relaxierung  $\hat{P}$  von P lässt eine effiziente Berechenbarkeit mittels Methoden der konvexen Optimierung zu. [22, S. 119]

Da konvexe Optimierungsverfahren im begrenzten Rahmen der Arbeit nicht erfasst werden können, sei für nähere Informationen auf Kapitel zwei von [22] verwiesen.

### 5.3 Intervallarithmetik

Mit Intervallarithmetik bzw. **Intervallanalyse** ist es möglich, konkrete Relaxierungen von Funktionen auf Mengen numerisch zu konstruieren. Hierfür definiert man zunächst Rechenoperationen auf Intervalle und erweitert anschließend elementare Funktionen auf intervallwertige Funktionen. Anschließend werden Verknüpfungen von elementaren Funktionen auf Intervallargumente verallgemeinert. Die Ideen der Intervallarithmetik können dann durch Taylor-Modelle verbessert werden. [22, S. 120-126]

Weitere, für die globale Optimierung wichtige Resultate, sind darüber hinaus [18, S. 303]:

- Durch Intervallanalyse können Nichtlinearitäten auf einfache Art und Weise kontrolliert werden.
- Intervallanalyse erweitert die klassische Analyse durch die Möglichkeit, semilokale Existenz- und Optimalitätsbedingungen, innerhalb eines lokalen Bereichs, um einen Punkt zu berechnen. (Klassische Analyse kann lediglich die Existenz behaupten.)

Die Grundrechenarten auf Intervallen werden folgendermaßen definiert (vgl. [18, S. 304-305] und [22, S. 122-125]):

#### 5.3.1 Definition Grundrechenarten auf Intervallen

Seien  $X = [\underline{x}, \bar{x}]$ ,  $Y = [\underline{y}, \bar{y}]$ , der Vektor  $\mathbf{x} \in \mathbb{R}^n$  und das Skalar  $c \in \mathbb{R}$  mit  $c \geq 0$  gegeben, dann gelten folgende Regeln:

##### Addition:

- $X + Y := [\underline{x} + \underline{y}, \bar{x} + \bar{y}]$
- $\mathbf{x} + \mathbf{y} := [\underline{x} + \underline{y}, \bar{x} + \bar{y}]$

##### Subtraktion:

- $-X := [-\bar{x}, -\underline{x}]$
- $X - Y := [\underline{x} - \bar{y}, \bar{x} - \underline{y}]$

##### Multiplikation:

- $X \cdot Y := [\inf(A), \sup(A)]$ ,  $A = \{\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}\}$
- $c \cdot Y := [c\underline{y}, c\bar{y}]$ , ( $c < 0 \Rightarrow c \cdot Y = [c\bar{y}, c\underline{y}]$ )

##### Division:

- $\frac{1}{X} := [\frac{1}{\bar{x}}, \frac{1}{\underline{x}}]$  mit  $0 \notin X$
- $\frac{X}{Y} := X \cdot \frac{1}{Y}$  mit  $0 \notin Y$

Es sei erwähnt, dass sich nicht alle Rechenregeln auf Intervalle übertragen lassen, da im Allgemeinen die Gleichungen

$$X - X = [0, 0] \quad \text{oder} \quad X/X = [1, 1]$$

mit  $0, 1 \in \mathbb{R}^n$  nicht gelten, sondern nur für den Fall  $\underline{x} = \bar{x}$  erfüllt werden [18, S. 305][22, S. 125].

Sei  $f$  eine elementare Funktion (*exp, sqrt, sin, ...*) und  $X \subseteq \mathbb{R}^n$  ein Intervall so gilt (vgl. [18, S. 304] und [22, S. 121-124]):

$$\text{Bild}(f, X) := [\inf_{\mathbf{x} \in X} (f(\mathbf{x})), \sup_{\mathbf{x} \in X} (f(\mathbf{x}))]$$

Für die Anwendung der Intervallarithmetik ist es erforderlich, dass die vorliegende Funktion faktorisiert ist [22, S. 126]. Eine faktorisierte Funktion wird, wie folgt, beschrieben [22, S. 126]:

#### 5.3.2 Definition Faktorisierte Funktion

Eine Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  heißt faktorisiert, wenn die Funktionsvorschrift von  $f$  sich in endlich viele Elementaroperationen, bestehend aus Grundrechenarten, elementare Funktionen und Kompositionen, zerlegen lässt.

Um Intervalle mit Funktionen abbilden zu können werden Intervallerweiterungen benötigt, welche wie folgt definiert werden (vgl. [22, S. 127]):

#### 5.3.3 Definition Intervallerweiterung

- Für  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  heißt  $F: \mathbb{R}^n \rightarrow \mathbb{R}$  **Intervallerweiterung** von  $f$ , falls  $\forall \mathbf{x} \in \mathbb{R}^n$

$$F([\mathbf{x}, \mathbf{x}]) = [f(\mathbf{x}), f(\mathbf{x})]$$

gilt.

- Für eine faktorisierte Funktion  $k: \mathbb{R}^n \rightarrow \mathbb{R}$  heißt  $F: \mathbb{R}^n \rightarrow \mathbb{R}$ , mit

$$F(X_1, \dots, X_n) := f(X_1, \dots, X_n)$$

**natürliche Intervallerweiterung** von  $f$ .

##### Hinweis:

Im Allgemeinen stimmen  $F(X_1, \dots, X_n)$  und  $\text{Bild}(f, X_1 \times \dots \times X_n)$  nicht überein, insbesondere auch nicht für den Fall  $F(X_1)$  und  $f(X_1)$  [22, S. 127].

##### Hierzu folgendes Beispiel [22, S. 127]:

Sei die Funktion  $f$  gegeben mit

$$f(x) = x - x^2$$

und das Intervall  $X = [0, 1]$ . Dann lässt sich  $F(X)$  unter Verwendung der Rechenoperationen leicht berechnen als

$$F([0, 1]) = [-1, 1],$$

wohingegen sich das Bild von  $X$  unter  $f$  errechnet als

$$\text{Bild}(f, [0, 1]) = [0, \frac{1}{4}].$$

Neben dieser Erkenntnis tritt auch das Phänomen des sogenannten **Abhängigkeitseffekts** in Erscheinung. Jedes Einsetzen einer Variable verhält sich hierbei derart, als würde es sich um unabhängige Variablen handeln, womit die Terme  $X(1 - X)$  und  $X(1 - Y)$  mit Bedingung  $Y = X$  als gleich betrachtet werden. Dadurch würden allerdings Abhängigkeiten wie

- $x = 1 \Rightarrow 1 - x = 0$ , sowie
- $x = \frac{1}{2} \Rightarrow 1 - x = \frac{1}{2}$

vernachlässigt und weitere Fehler produziert. [22, S. 128]

Ogleich dieser Eigenschaften liefert  $F(X)$  eine Obermenge von  $\text{Bild}(f, X)$  womit eine garantierte untere Schranke für den Minimalwert, sowie eine garantierte obere Schranke für den Maximalwert von  $f$  auf  $X$  gefunden werden kann. [22, S. 129]

Der Beweis hierfür erfolgt über die Monotonie von Intervallerweiterungen (vgl. [22, S. 130-131]):

### 5.3.4 Definition Monotone Intervallerweiterung

Eine intervallwertige Funktion  $F : \mathbb{R}^n \Rightarrow \mathbb{R}$  heißt monoton, falls

$$\forall X, Y \in \mathbb{R}^n \text{ mit } X \subseteq Y : F(X) \subseteq F(Y)$$

gilt.

In Vorbereitung des Beweises, dass die Intervallerweiterung eine Schranke liefert, brauchen wir noch folgenden Satz inklusive Korollar, wobei deren Beweise im Buch [22, S. 130-131] nachgelesen werden können:

### 5.3.5 Satz

Die Intervallgrundrechenarten, die intervallwertigen elementaren Funktionen sowie die Komposition von Funktionen sind monoton.

### 5.3.6 Korollar

Für jede faktorisierte Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  ist ihre natürliche Intervallerweiterung  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  monoton.

Somit lässt sich nun zeigen [22, S. 131]:

### 5.3.7 Satz

Für jede faktorisierte Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , ihre natürliche Intervallerweiterung  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  und alle  $X \in \mathbb{R}^n$  gilt

$$\text{Bild}(f, X) \subseteq F(X)$$

**Beweis:**

Für alle  $x \in X$  gilt

$$[f(x), f(x)] = F([x, x]) \stackrel{F \text{ monoton}}{\subseteq} F(X)$$

und somit  $f(x) \in F(x)$ , sowie  $\text{Bild}(f, X) \subseteq F(X)$ .

Da die Schranken aufgrund des Abhängigkeitseffektes relativ grob sein können, kann für verbesserte Schranken auf Taylor-Modelle zurückgegriffen werden [22, S. 131]. Da diese mit stetig differenzierbar allerdings starke Anforderungen stellen, sei hier nur auf [22, S. 131-132] verwiesen.

Nach Constraint Propagation, konvexer Relaxierung und Intervallarithmetik bzw. Intervallanalyse, welche im Vergleich nur geringe Reduktion der Boxen liefern und damit nur geringe Effizienzsteigerungen gewährleisten, werden nachfolgend Dualitätsargumente betrachtet, welche signifikante Reduktionen ermöglichen [18, S. 303].

## 5.4 Dualitätsargumente

Als exemplarisches Beispiel, wie **Dualitätsargumente**, bessere Schranken und somit eine höhere Reduktion liefern, wird die Lagrange-Dualität hinzugezogen. Es kann gezeigt werden, dass die Dualitätslücke, das heißt die Differenz des Minimalwerts des Problems 1 und dem Maximalwert seines Lagrange-Duals, im Grenzfall gegen Null geht [6, S. 347][22, S. 65-66].

Für die Definition der konvexen Hülle wird zunächst der Begriff der Halbstetigkeit von unten von Funktionen benötigt (vgl. [12, S. 10]):

### 5.4.1 Definition Halbstetigkeit von unten

Eine Funktion  $f : X \rightarrow \mathbb{R}$  mit  $X$  topologischer Raum (zum Beispiel  $X \subset \mathbb{R}^n$ ) heißt halbstetig von unten, wenn für jeden Punkt  $x \in X$  und jedes  $a \in \mathbb{R}$  die Bedingung  $f(x) > a$  die Existenz einer Umgebung um den Punkt  $x$  impliziert, in welcher alle Werte von  $f$  größer als  $a$  sind.

Mit dieser Definition kann nun die konvexe Hülle einer Funktion definiert werden (vgl. [6, S. 349]) als:

### 5.4.2 Definition konvexe Hülle

Sei  $X \subset \mathbb{R}^n$  eine nichtleere, kompakte und konvexe Menge sowie  $f : X \rightarrow \mathbb{R}$  eine auf  $X$  von unten halbstetige Funktion. Dann ist die Funktion

$$\phi_{X,f} : X \rightarrow \mathbb{R}, x \mapsto \phi_{X,f}(x)$$

die konvexe Hülle von  $f$  über  $X$ .

Mit folgendem Lemma wird der Zusammenhang zwischen einer Funktion  $f$  und ihrer konvexen Hülle auf einer kompakten Menge aufgezeigt (vgl. [6, S. 350]):

### 5.4.3 Lemma

Sei  $f$  und  $X$  definiert wie in der vorherigen Definition,  $K \subset X$  kompakt und konvex und  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  eine affine Funktion. Dann gilt:

- a)  $m := \min\{f(x) \mid x \in X\} = \min\{\phi_{X,f}(x) \mid x \in X\}$
- b)  $\{y \in X \mid f(y) = m\} \subseteq \{y \in X \mid \psi_{X,f}(y) = m\}$
- c)  $\phi_{X,f}(x) \geq \phi_{K,f}(x), \quad \forall x \in X$
- d)  $\phi_{X,f+g} = \phi_{X,f} + g$ .

Das Lemma ist für spätere Abschätzungen mittels eines konvexifizierten Problems von Bedeutung.

Sei das nicht-konvexe globale Optimierungsproblem (P) von 1 gegeben mit  $X \subset \mathbb{R}^n$  nichtleere, kompakte und konvexe Menge, die Funktion  $f : X \rightarrow \mathbb{R}$ , welche auf  $X$  von unten halbstetig ist, sowie die Funktion  $g : X \rightarrow \mathbb{R}^m$ , welche komponentenweise von unten halbstetig auf  $X$  ist.

Nehme an, dass die zulässige Lösungsmenge gegeben durch

$$S = \{x \in X \mid g(x) \leq 0\}$$

nichtleer ist, sodass eine Lösung existiert.

Für  $u \in \mathbb{R}_+^m$ ,  $x \in X$  definiere die Lagrange-Funktion von (P) als

$$L(x, u) := f(x) + u^T g(x),$$

dann ist das duale Problem (D) gegeben durch

$$\begin{array}{ll} \max & d(u) \\ \text{s.t.} & u \in \mathbb{R}_+^m \end{array} \quad (2)$$

mit  $d : \mathbb{R}_+^m \rightarrow \mathbb{R}$ ,  $u \mapsto d(u) = \min_{x \in X} L(x, u)$ . [6, S. 350-351]

Seien  $\min(P)$  und  $\sup(D)$  die Optimalwerte des primalen und dualen Optimierungsproblems, dann gilt mit

$$(P) \min_{x \in X} \max_{u \in \mathbb{R}_+^m} L(x, u)$$

die schwache Dualität  $\min(\mathbf{P}) \geq \sup(\mathbf{D})$ , wobei  $\Delta$  mit

$$\Delta := \min(\mathbf{P}) - \sup(\mathbf{D}) > 0$$

die Dualitätslücke beschreibt. [6, S. 350-351]

Sind die Restriktionen  $g$  linear, ist es möglich, vom dualen Problem  $(\mathbf{D})$  auf ein konvexifiziertes Problem  $(\hat{\mathbf{P}})$  zu wechseln, welches ein lineares Programm darstellt. Im Fall der Nichtlinearität ist eine Abschätzung der Dualitätslücke über Definitionen fehlender Konvexität einer Funktion möglich. [6, S. 351-352]

Sei folgend  $g$  linear mit  $g(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$  mit  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$ , so erhält man das konvexifizierte Problem  $(\hat{\mathbf{P}})$  durch Ersetzen der Zielfunktion  $f$  durch die konvexe Hülle  $\phi_{X,f}$  (vgl. [6, S. 351-352]):

$$\begin{array}{ll} \min & \phi_{X,f}(\mathbf{x}) \\ \text{s.t.} & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \in X \end{array} \quad (3)$$

Außerdem sei  $(\hat{\mathbf{D}})$  das duale Problem zu  $(\hat{\mathbf{P}})$  mit der Lagrange-Funktion  $\hat{L}(\mathbf{x}, \mathbf{u}) = \phi_{X,f}(\mathbf{x}) + \mathbf{u}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$  gegeben, lässt sich zeigen (vgl. [6, S. 350-351]):

#### 5.4.4 Proposition

Sei  $g(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$ ,  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$  und  $(\hat{\mathbf{P}})$  entstehe durch Ersetzen von  $f$  durch die konvexe Hülle  $\phi_{X,f}$ , dann gilt:

- a)  $\min(\hat{\mathbf{P}}) = \sup(\hat{\mathbf{D}}) = \sup(\mathbf{D})$
- b)  $\Delta = \min(\mathbf{P}) - \min(\hat{\mathbf{P}})$

Sei  $(\hat{\mathbf{P}})$  das Problem, welches aus  $(\mathbf{P})$  entsteht, indem alle Funktionen durch ihre konvexen Hüllen über  $X$  ersetzt werden. Sind  $d(\mathbf{u})$  die korrespondierende Zielfunktion sowie  $(\hat{\mathbf{D}})$  das korrespondierende Optimierungsproblem, so ist die Lagrange-Funktion für jedes konvexe Primalproblem eine konvexe Unterabschätzung der Lagrange-Funktion von  $\mathbf{P}$  über  $X$  für alle  $\mathbf{u} \in \mathbb{R}_+^m$ . Somit gelten [6, S. 352]:

- (i)  $\min(\mathbf{P}) \geq \sup(\mathbf{D}) \geq \sup(\hat{\mathbf{D}})$
- (ii)  $\sup(\hat{\mathbf{D}}) = \inf(\hat{\mathbf{P}})$

Weiterhin ergibt sich folgendes Korollar [6, S. 353]:

#### 5.4.5 Korollar

Sei  $(\mathbf{P})$ ,  $(\mathbf{D})$  und  $(\hat{\mathbf{P}})$ ,  $(\hat{\mathbf{D}})$  wie eben definiert und  $(\hat{\mathbf{P}})$  entstehe aus  $(\mathbf{P})$  durch Ersetzen von  $f$  durch die konvexe Hülle, dann gilt

$$0 \leq \Delta = \inf(\mathbf{P}) - \sup(\mathbf{D}) \leq \inf(\mathbf{P}) - \inf(\hat{\mathbf{P}}).$$

Die gewonnenen Erkenntnisse lassen sich nun auf Branch-Bound-Verfahren übertragen, indem Abschätzungen auf die vorgenommenen Teilungen durchgeführt werden. Hierbei wird gezeigt, dass die Abschätzung, mittels konvexer Hülle, auch bei Teilung erhalten bleibt.

Hierfür zunächst folgendes Lemma [6, S. 353]:

#### 5.4.6 Lemma

Sei  $A_k \subset \mathbb{R}^n$  kompakt und  $\emptyset \neq A_{k+1} \subset A_k$  für alle  $k \in \mathbb{N}$ . Dann gilt

$$\text{conv}(\cap_{k \in \mathbb{N}} A_k) = \cap_{k \in \mathbb{N}} \text{conv}(A_k) \quad (1)$$

wobei  $\text{conv}$  die konvexe Hüllen-Operation darstellt.

Das Lemma zeigt, dass der Schnitt über alle konvexen Hüllen gleich der konvexen Hülle aller Schnitte der kompakten Mengen entspricht. Somit gilt insbesondere (vgl. [6, S. 354]):

#### 5.4.7 Korollar

Für alle  $k \in \mathbb{N}$ , sei  $X_k \subset \mathbb{R}^n$  kompakt, konvex und  $\emptyset \neq X_{k+1} \subset X_k$ . Sei weiterhin  $X := \lim_{k \rightarrow \infty} X_k = \cap_{k \in \mathbb{N}} X_k$  und  $f : X_1 \rightarrow \mathbb{R}$  eine auf  $X_1$  beschränkte von unten halbstetige Funktion.

Dann gilt für die konvexen Hüllen  $\phi_{X_k,f}$  und  $\phi_{X,f}$  von  $f$  über  $X_k$  und  $X$

$$\lim_{k \rightarrow \infty} \phi_{X_k,f}(\mathbf{x}) = \sup_{k \in \mathbb{N}} \phi_{X_k,f}(\mathbf{x}) = \phi_{X,f}(\mathbf{x}) \quad \forall \mathbf{x} \in X.$$

Das Grenzwertverhalten der Folgen von Minima der konvexen Hüllen folgt aus dem allgemeineren Ergebnis (vgl. [6, S. 355]):

#### 5.4.8 Lemma

Für alle  $k \in \mathbb{N}$ , sei  $X_k \subset \mathbb{R}^n$  kompakt und es gelte  $\emptyset \neq X_{k+1} \subset X_k$ . Sei weiterhin  $X := \lim_{k \rightarrow \infty} X_k = \cap_{k \in \mathbb{N}} X_k$  und  $f_k : X_k \rightarrow \mathbb{R}$  eine auf  $X_k$  von unten halbstetige Funktion, welche  $f_{k+1} \geq f_k$  auf  $X_{k+1}$  und  $f_k \leq \mu$  auf  $X_k$ ,  $\forall k$  erfüllt, dann gilt

$$\begin{aligned} \min_{k \in \mathbb{N}} \{ \sup_{\mathbf{x} \in X_k} f_k(\mathbf{x}) \mid \mathbf{x} \in X \} &= \lim_{k \rightarrow \infty} \{ \min_{\mathbf{x} \in X_k} f_k(\mathbf{x}) \mid \mathbf{x} \in X_k \} \\ &= \sup_{k \in \mathbb{N}} \{ \min_{\mathbf{x} \in X_k} f_k(\mathbf{x}) \mid \mathbf{x} \in X_k \}. \end{aligned}$$

Unter Verwendung des vorherigen Korollars ergibt sich für konvexe Hüllen folgendes [6, S. 356]:

#### 5.4.9 Korollar

Mit den Notationen und Folgerungen aus des vorherigen Korollars gilt

$$\min \{ \phi_{X,f} \mid \mathbf{x} \in X \} = \lim_{k \rightarrow \infty} \{ \min \phi_{X_k,f} \mid \mathbf{x} \in X_k \}.$$

Sei nun das Problem 1 gegeben und eine Partition der Box  $X$  wobei  $f$  und  $g$  von unten halbstetige reellwertige Funktionen auf einer kompakten Menge  $X_1 \supset S$  sind. Dann wird für jeder Iteration  $k$  des Branch-and-Bound-Verfahrens für die unteren Schranken  $\beta(X)$  die Ungleichung

$$\beta(X) \leq \min \{ f(\mathbf{x}) \mid \mathbf{x} \in X \cap S \}$$

für jede kompakte konvexe Menge aus  $X \in \mathbb{P}(k)$  erfüllt, wobei  $\mathbb{P}(k)$  eine endliche Partition einer Teilmenge  $X_1$  ist. Es stellt

$$\beta_k := \min \{ \beta(k) \mid X \in \mathbb{P}(k) \}$$

die bisher erreichte untere Schranke für  $\min(\mathbf{P})$  dar.

Für das finale Theorem wird noch folgendes Lemma benötigt [6, S. 357]:



#### 5.4.10 Lemma

Erfüllt eine unendliche ansteigende Folge  $\{X_k\}_{k \in \mathbb{N}}$  von sukzessive verfeinerten Partitionen die Bedingungen

$$K := \bigcap_{k \in \mathbb{N}} X_k \subset S \quad \text{und} \\ \lim_{k \rightarrow \infty} \beta(X_k) = \min\{f(x) \mid x \in X\},$$

so gilt

$$\lim_{k \rightarrow \infty} \beta_k = \min\{f(x) \mid x \in S\}$$

Somit ergibt sich folgendes Theorem (vgl. [6, S. 357-358]), dessen Beweis unter [6, S. 357-358] in Ausführlichkeit nachgelesen werden kann:

#### 5.4.11 Theorem

Sei  $S = \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$  nichtleer und kompakt mit von unten halbstetigen Komponentenfunktionen  $g_i : X_1 \rightarrow \mathbb{R} \quad \forall i$  auf der kompakten, konvexen Menge  $X_1 \supset S$ . Sei  $f : X_1 \rightarrow \mathbb{R}$  von unten halbstetig und beschränkt auf  $X_1$  und sei  $\{X_k\}_{k \in \mathbb{N}}$  eine absteigende Folge von nichtleeren, kompakten, konvexen Mengen in  $\mathbb{R}^n$ , die gegen  $X \subset S$  konvergieren. Außerdem sei für jedes  $k \in \mathbb{N}$ ,

$$v_k = \max_{u \in \mathbb{R}_+^m} \min_{x \in X_k} \{f(x) + u^T g(x)\},$$

und  $(\hat{P})$  entstehe durch das Ersetzen von  $f$  durch die konvexe Hülle, womit gilt

$$(\hat{P}_k) \min\{\phi_{X_k, f}(x) \mid \phi_{X_k, g_i}(x) \leq 0 \quad \forall i, x \in X_k\}.$$

Dann gilt

$$\lim_{k \rightarrow \infty} v_k = \min\{f(x) \mid x \in X\}.$$

Die Lagrange-Dualität liefert somit sehr gute untere Schranken, die zu konvergenten Algorithmen führen und im Gegensatz zum Branch-and-Bound-Ansatz keine oberen Schranken benötigt. In vielen Klassen von Optimierungsproblemen kann die Äquivalenz zwischen dem Lagrange-Dualen Problemen und einem Linearen Programm gezeigt werden, womit eine Lösbarkeit in polynomieller Laufzeit möglich ist. Zu den Problemen zählen unter anderem konkave Minimierung unter umgekehrt konvexen Nebenbedingungen, bilineare Programmierung, Multiojektive Programming, Quadratische Probleme mit quadratischen Nebenbedingungen. [3, S. 103]

Im folgenden Abschnitt werden Anwendungsfälle der Deterministischen Globalen Optimierung präsentiert und ein paar Software-Programme genannt, die für die Lösung von Deterministischen Globalen Optimierung entwickelt wurden.

## 6 ANWENDUNGEN UND SOFTWARE

Die Forschungsgebiete der Globalen Optimierung sind sehr vielfältig und decken unter anderem die Computerchemie, die Molekularbiologie, die Wirtschaftswissenschaften, die Thermodynamik und die Logistik ab [7, S. 23-24]. Insbesondere im Bereich der Chemie wurde in theoretischen und algorithmischen Studien sehr starker Aufwand betrieben [7, S. 27].

Eine Anwendung stellt die Strukturvorhersage von Proteinfaltungen dar. Aufgrund der intra- und intermolekularen Wechselwirkungen können Proteine unterschiedliche Strukturen annehmen, welche die biologische Wirksamkeit beeinflussen. Die Vorhersage der Struktur wird in der Computerchemie und der Molekularbiologie erforscht. Die Strukturvorhersage kann auf die Lösung eines nichtlinearen, zweifach differenzierbaren Optimierungsproblems mit Boxenbeschränkung zurückgeführt werden. [7, S. 11-12]

Ein weiterer Anwendungsfall für Globale Optimierung ist die Strukturvorhersage von Clustern. Cluster sind hierbei Aggregate von Atomen, Ionen oder Molekülen, die so klein sind, dass ein signifikanter Anteil der Einheiten auf ihren Oberflächen vorhanden ist. Die Vorhersage der Struktur hat besondere Bedeutung für die Thermodynamik, die Materialwissenschaft und Katalyse, sowie der Astrophysik. Die Formulierung des Problem verfolgt den Ansatz  $N$  Teilchen zu finden, die mit zentralen Zweikörperkräften wechselwirken und ihre Konfigurationen im dreidimensionalen euklidischen Raum, unter Einbeziehung des globalen Minimum der potentiellen Energie, finden. Das mathematische Modell ist eine zweifach differenzierbare nichtlineare und nichtkonvexe Zielfunktion, die Beschränkungen durch Boxen unterliegt. [7, S. 10-11]

Weitere Beispiele sind zum Beispiel in [7] und [18] zu finden.

Für die Anwendung unterschiedlicher Verfahren sind einige Software-Anwendungen entwickelt worden. Exemplarisch seien hier drei genannt (vgl. [18, S. 288]):

- 1. BARON** steht für Branch-and-Reduce Optimization Navigator und stellt einen allgemeinen Löser für Optimierungsprobleme mit nichtlinearen Restriktionen dar. BARON basiert auf Branching und Box-Reduktionen über Verwendung von konvexen und polyhedralen Relaxierungen und Lagrange-Techniken.
- 2. GlobSol** bzw. Global Solver basiert auf Branch and Bound mit allgemeinen faktorisierten Restriktionen. Die Boxen-Reduktion erfolgt mit Intervallanalyse um keine global optimale Lösung zu verlieren.
- 3. Frontline Interval Global Solver** ist ein Löser, basierend auf Intervall-Methoden und linearen Relaxierungen. Er bietet Interfaces für Visual Basic und Excel.

## 7 DISKUSSION

Ziel der Arbeit war die Betrachtung von deterministischer Globaler Optimierung in Bezug auf lineare und nichtlineare Programmierung, sowie ganzzahlige lineare und nichtlineare Programmierung.

Um diesem Ziel nahezukommen wurden allgemeine Methoden und Ansätze der deterministischen Globalen Optimierung gewählt, welche sich für alle der genannten Probleme eignen und somit keine Einschränkung auf spezifische Problemstellungen darstellen, da die im Verlauf der Recherche betrachteten Journals und Artikel

zum Teil sehr nischenhafte Probleme untersuchen. Eine verständliche Darbietung der zugehörigen Lösungsansätze hätte den Umfang der Arbeit deutlich erhöht, womit ein Fokus auf zentrale Ansätze gelenkt wurde.

Die ausführlichen und mathematischen Darstellungen im Abschnitt Branch-and-Reduce sollen verdeutlichen, dass für den Entwurf eines Verfahrens viele unterschiedliche Bereiche Relevanz haben, und diese auch in Kombination sinnvoll angewendet werden können und sollten. Somit erhält man in der Deterministischen Globalen Optimierung exakte Verfahren, die einen Minimalwert, gegebenenfalls bei sehr langer Laufzeit, errechnen, während die stochastische Globale Optimierung oft auf Heuristiken zurückgreift, um in kürzerer Zeit Ergebnisse zu liefern. Allerdings gibt es bei stochastischen Verfahren keine Garantie, dass es sich um eine gute Annäherung an den Minimalwert handelt.

Naturinspirierte Verfahren und Heuristiken sind ein wichtiges Werkzeug in vielen Bereichen um Abschätzungen zu treffen und effizient Lösungen bzw. Lösungskandidaten zu bestimmen. In der Deterministischen Globalen Optimierung zeigt sich allerdings, dass die Heuristiken gegenüber den exakten Verfahren unterlegen sein können [15, S. 13].

Aktuelle Forschungen und Erkenntnisse sind unter anderem im **Journal of Global Optimization** oder im **Journal of Optimization Theory and Applications** zu finden.

## 8 AUSBLICK

Die Globale Optimierung und somit auch die Deterministische Globale Optimierung ist ein sehr dynamisches Feld, welches viele Anwendungen bietet [17, S. vii]. Durch die Entwicklung leistungsfähiger algorithmischer Ideen und Fortschritten bei der theoretischen Analyse und Bewertung der Komplexität rückt es stärker in den Fokus [17, S. 1]. Neben der Einbindung von naturinspirierten Verfahren und Heuristiken, werden auch neuronale Netzwerke in die Methoden der Deterministischen Globalen Optimierung integriert [20]. Das Feld der Deterministischen Globalen Optimierung wird somit voraussichtlich weiter wachsen und mit der ansteigenden Möglichkeiten der Rechner, insbesondere in Bezug auf Quantencomputer, an Einsatzmöglichkeiten und Relevanz gewinnen [1][2][16]. Es lässt sich zum Beispiel zeigen, dass Quantencomputer Methoden der Globalen Optimierung, welche Lipschitz-Bedingungen nutzen, quadratisch beschleunigen können [2, S. 2].

## LITERATUR

- [1] Akshay Ajagekar, Travis Humble, und Fengqi You. Quantum computing based hybrid solution strategies for large-scale discrete-continuous optimization problems. *Computers & Chemical Engineering*, 132:106630, 2020. DOI: 10.1016/j.compchemeng.2019.106630.
- [2] Cezar-Mihail Alexandru, Ella Bridgett-Tomkinson, Noah Linden, Joseph MacManus, Ashley Montanaro, und Hannah Morris. Quantum speedups of some general-purpose numerical optimisation algorithms. *Quantum Science and Technology*, 5(4):045014, 2020. DOI: 10.1088/2058-9565/abb003.
- [3] Charles Audet, Pierre Hansen, und Giles Savard. *Essays and Surveys in Global Optimization* -. Springer Science & Business Media, Berlin Heidelberg, 2005. ISBN 978-0-387-25569-9. DOI: 10.1007/b135610.
- [4] Dirk Briskorn. *Operations Research - Eine (möglichst) natürlchsprachige und detaillierte Einführung in Modelle und Verfahren*. Springer Berlin Heidelberg, Berlin Heidelberg, 2019. ISBN 978-3-662-60782-4. DOI: 10.1007/978-3-662-60783-1.
- [5] Wolfgang Domschke, Andreas Drexl, Robert Klein, und Armin Scholl. *Einführung in Operations Research*. Springer Gabler, Berlin Heidelberg, 9. Auflage, 2015. ISBN 978-3-662-48216-2. DOI: 10.1007/978-3-662-48216-2.
- [6] M. Dür und R. Horst. Lagrange duality and partitioning techniques in nonconvex global optimization. *Journal of Optimization Theory and Applications*, 95(2):347–369, 1997. DOI: 10.1023/a:1022687222060.
- [7] Christodoulos A. Floudas. *Deterministic Global Optimization - Theory, Methods and Applications*. Springer Science & Business Media, Berlin Heidelberg, 2000. ISBN 978-1-4419-4820-5. DOI: 10.1007/978-1-4757-4949-6.
- [8] Matthias Gerdt und Frank Lempio. *Mathematische Optimierungsverfahren des Operations Research*. Walter de Gruyter, Berlin, 2011. ISBN 978-3-110-24998-9. DOI: 10.1515/9783110249989.
- [9] Eligius M.T. Hendrix und Boglárka G.-Tóth. *Introduction to Nonlinear and Global Optimization*. Springer, Berlin, Heidelberg, 2010. ISBN 978-0-387-88670-1. DOI: 10.1007/978-0-387-88670-1.
- [10] Waltraud Huyer und Arnold Neumaier. Global optimization by multilevel coordinate search. *Journal of Global Optimization*, 14, 12 1998. DOI: 10.1023/A:1008382309369.
- [11] D. R. Jones, C. D. Perttunen, und B. E. Stuckman. Lipschitzian optimization without the lipschitz constant. *Journal of Optimization Theory and Applications*, 79(1):157–181, 1993. DOI: 10.1007/bf00941892.
- [12] Vladimir Kadets. *A Course in Functional Analysis and Measure Theory*. Springer, Berlin, Heidelberg, 2018. ISBN 978-3-319-92004-7. DOI: 10.1007/978-3-319-92004-7.
- [13] Klaus-Peter Kistner. *Optimierungsmethoden - Einführung in die Unternehmensforschung für Wirtschaftswissenschaftler*. Springer-Verlag, Berlin Heidelberg, 2003. ISBN 978-3-7908-0043-2. DOI: 10.1007/978-3-642-57437-5.
- [14] Steffen Kjeller, Pavel Kozine, Kaj Madsen, und Ole Stauning. *Non-linear Global Optimization Using Interval Arithmetic and Constraint Propagation*, pages 45–58. Springer US, Boston, MA, 2007. ISBN 978-0-387-36721-7. DOI: 10.1007/978-0-387-36721-7\_3.
- [15] Dmitri E. Kvasov und Marat S. Mukhametzhonov. Metaheuristic vs. deterministic global optimization algorithms: The univariate case. *Applied Mathematics and Computation*, 318:245–259, 2018. DOI: 10.1016/j.amc.2017.05.014.
- [16] Pedro Lara, Renato Portugal, und Carlile Lavor. A new hybrid classical-quantum algorithm for continuous global optimization problems. *Journal of Global Optimization*, 60, 01 2013. DOI: 10.1007/s10898-013-0112-8.
- [17] Marco Locatelli und Fabio Schoen. *Global Optimization - Theory, Algorithms, and Applications*. SIAM, Philadelphia, 2013. ISBN 978-1-611-97266-5. DOI: 10.1137/1.9781611972672.
- [18] Arnold Neumaier. Complete search in continuous global optimization and constraint satisfaction. *Acta Numerica 2004*, page 271–370, 2004. DOI: 10.1017/cbo9780511569975.004.
- [19] János D. Pintér. *Global Optimization in Action - Continuous and Lipschitz Optimization: Algorithms, Implementations and Applications*. Springer Science & Business Media, Berlin Heidelberg, 2013. ISBN 978-1-475-72502-5. DOI: 10.1007/978-1-4757-2502-5.
- [20] Artur M. Schweidtmann und Alexander Mitsos. Deterministic global optimization with artificial neural networks embedded. *Journal of Optimization Theory and Applications*, 180(3):925–948, 2018. DOI: 10.1007/s10957-018-1396-0.
- [21] Rainer Schwenkert und Yvonne Stry. *Operations Research kompakt - Eine an Beispielen orientierte Einführung*. Springer Gabler, Berlin Heidelberg, 2015. ISBN 978-3-662-48396-1. DOI: 10.1007/978-3-662-48397-8.
- [22] Oliver Stein. *Grundzüge der Globalen Optimierung*. Springer Spektrum, Berlin Heidelberg, 2018. ISBN 978-3-662-55359-6. DOI: 10.1007/978-3-662-55360-2.
- [23] Brigitte Werners. *Grundlagen des Operations Research - Mit Aufgaben und Lösungen*. Springer Berlin Heidelberg, Wiesbaden, 3. Auflage, 2013. ISBN 978-3-642-40101-5. DOI: 10.1007/978-3-642-40102-2.