Masterarbeit

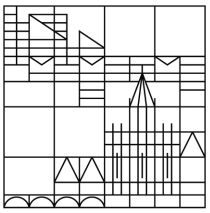
Konvexe Optimierung am Beispiel volumenmaximaler einbeschriebener Rechtecksmengen

vorgelegt von

Marco Daub

an der

Universität Konstanz



Fachbereich Mathematik und Statistik

1. Gutachter Prof. Dr. Stefan Volkwein
2. Gutachter Prof. Dr. Matthias Gerdts

Konstanz, März 2017

Konstanzer Online-Publikations-System (KOPS)
URL: http://nbn-resolving.de/urn:nbn:de:bsz:352-0-420923

Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit mit dem Thema:

Konvexe Optimierung am Beispiel volumenmaximaler einbeschriebener Rechtecksmengen

selbständig verfasst und keine anderen Hilfsmittel als die angegebenen benutzt habe. Die Stellen, die anderen Werken dem Wortlaut oder dem Sinne nach entnommen sind, habe ich in jedem einzelnen Falle durch Angaben der Quelle, auch der benutzten Sekundärliteratur, als Entlehnung kenntlich gemacht.

Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Konstanz, den 1. März 2017

Marco Daub

Inhaltsverzeichnis

1. Einleitung **3**

2. Topologische Grundlagen **7**
2.1. Euklidische Räume 7
2.2. Produkträume 12

3. Maßtheoretischer Volumenbegriff **19**
3.1. Borel-Mengen 19
3.2. Borel-messbare Funktionen 23
3.3. Das Lebesgue-Maß 27

4. Konvexe Theorie **32**
4.1. Konvexe Mengen und Funktionen 32
4.2. Konvexe Optimierungsprobleme 41

5. Das VER-Problem **46**
5.1. Einführung in die Problemstellung 46
5.2. Analytisches Problem 52

6. Das konvexe VER-Problem **63**
6.1. Allgemeine Strukturen 63
6.2. Eindimensionale Strukturen 71

7. Numerische Lösung des konvexen VER-Problems **76**
7.1. Grundlegendes 76
7.2. Beispielprobleme 83

8. Fazit **90**

A. Anhang **92**

Literatur **99**

1. Einleitung

Bei der Entwicklung eines komplexen Produkts, welches sich aus mehreren Komponenten zusammensetzt, müssen oftmals mehrere Anforderungen erfüllt werden. Ob eine Anforderung erfüllt ist, hängt dabei von der Zusammenstellung der einzelnen Komponenten ab. Diese Komponenten werden von verschiedenen Produktionsbereichen bereitgestellt. Anstatt nun ein optimales Design mit spezifischen Komponenten vorzugeben, ist es der Anspruch, jedem Produktionsbereich maximale Variabilität in der Realisierbarkeit ihrer Komponenten zu gewährleisten. Dies kann als anschauliches mathematisches Problem formuliert werden: Gesucht ist ein kartesisches Produkt mit maximalem Volumen, d.h. eine spezielle Menge zulässiger Designs, sodass jedes dieser Designs alle Anforderungen erfüllt. Mit anderen Worten, eine volumenmaximale einbeschriebene Rechtecksmenge.

Diese Problemstellung wurde unter der Voraussetzung, dass sich das kartesische Produkt, also die Rechtecksmenge, aus Intervallen zusammensetzt, bereits in [9] und [24] untersucht und Algorithmen zur numerischen Lösung vorgeschlagen. Der in [24] vorgestellte Algorithmus basiert darauf, eine Lösung durch zufällige Betrachtung zulässiger und nicht zulässiger Designs zu finden und der in [9] vorgestellte auf der Voraussetzung, dass den Anforderungen lineare Funktionen zugrunde liegen.
In dieser Arbeit werden Rechtecksmengen verwendet, bei denen die Dimensionen der Mengen, aus welchen sie sich zusammensetzen, beliebig sind. Zudem sind auch Anforderungen erlaubt, welche durch sogenannte additiv zerlegbare Funktionen beschrieben werden können. Auf der einen Seite steigt dadurch die Komplexität des mathematischen Problems. Auf der anderen Seite kann die Anwendbarkeit, vor allem gegenüber [9], erheblich erweitert werden. Die Auflockerung der Struktur des kartesischen Produkts ermöglicht nämlich einzelnen Produktionsbereichen, welche mehrere Komponenten gleichzeitig betrachten, weitere Möglichkeiten ihre einzelnen Komponenten zu kombinieren. Außerdem wird durch die additiv zerlegbaren Funktionen, im Vergleich zu den linearen, eine größere Funktionenklasse zur Beschreibung der Anforderungen zugelassen.

Da für dieses Vorhaben noch keine Literatur vorhanden ist, ist es das Ziel der Arbeit, zunächst eine saubere mathematische Formulierung des Problems zu schaffen, um dieses anschließend untersuchen und lösen zu können. Dazu sind Grundlagen der Topologie und Maßtheorie notwendig, auf deren relevantesten Inhalte hierfür im Folgenden eingegangen wird. Insbesondere kann gezeigt werden, dass unter etwas stärkeren Voraussetzungen ein konvexes Optimierungsproblem vorliegt, weshalb der

Fokus der Arbeit stark auf die konvexe Theorie gerichtet ist. Für diesen Fall wird dann genau beschrieben, wie das mathematische Problem numerisch gelöst werden kann. Die Aussagen können dabei nahezu direkt auf das allgemein untersuchte Problem übertragen werden.

Kapitel 1 stellt die Einleitung dar und gibt einen Überblick über die verwendeten Notationen.

Kapitel 2 stellt Elemente der Topologie vor, welche für den weiteren Verlauf der Arbeit von Bedeutung sind und legt die dafür verwendeten Notationen fest.

Kapitel 3 behandelt die ebenfalls für die Arbeit entscheidende, maßtheoretische Grundlagen. Insbesondere werden messbare Mengen eingeführt, denen dann über das Lebesgue-Maß ein Volumen zugeordnet werden kann.

Kapitel 4 betrachtet zunächst konvexe Mengen und Funktionen und anschließend Optimierungsprobleme, denen konvexe Funktionen zugrunde liegen.

Kapitel 5 führt das volumenmaximale-einbeschriebene-Rechtecksmengen-Problem ein und untersucht wichtige Eigenschaften des Problems.

Kapitel 6 betrachtet das VER-Problem für Anforderungen, welche durch konvexe Funktionen beschrieben werden und zeigt, dass dies ein konvexes Optimierungsproblem darstellt. Außerdem wird eine alternative Formulierung für einen Spezialfall des VER-Problems hergeleitet.

Kapitel 7 stellt vor, wie das konvexe VER-Problem mit einer ein- oder zweidimensionaler Struktur numerisch mit der Software Matlab® gelöst werden kann.

Kapitel 8 stellt das Fazit dar und gibt einen Ausblick.

Notationen

$\frac{\mathrm{d}}{\mathrm{d}x} f(x)$	Ableitung
$\forall$	für alle
$\Leftrightarrow$	Äquivalenz
$\lvert \cdot \rvert$	Betrag bzw. Betrag bei Multiindexnotation
$\square$	Beweisende
$\mathcal{B}(X^d)$	Borel-σ-Algebra
$\cap$	Durchschnitt
$\in$	Element von
$\notin$	nicht Element von
$\exists$	es existiert
$(a_i)_{i \in \mathbb{N}}$	Folge
$\lim_{i \to \infty}$	Grenzwert
$\Rightarrow$	Implikation
inf	Infimum
$\int$	Integral
$[a, b]$	abgeschlossenes Intervall
$(a, b]$	halboffenes Intervall
(a, b)	offenes Intervall
$\times$	kartesisches Produkt
$\setminus$	Komplement
A^{c}	Komplement von A
$\circ$	Komposition
$B(x_0, r)$	offene Kugel
o, $\mathcal{O}$	Landausymbole
$L(\gamma)$	Länge von γ
λ^d	Lebesgue-Maß
$\emptyset$	leere Menge
lim inf	limes inferior
lim sup	limes superior
log	Logarithmus zur Basis e
∇	Nabla (Gradient)
$\lVert \cdot \rVert$	Norm
$\mathcal{P}(A)$	Potenzmenge von A

$\prod_{i=1}^{d}$	Produkt bzw. mehrfaches kartesisches Produkt
$\bigotimes_{k=1}^{n} \mu^k(A)$	Produktmaß
$\bigotimes_{k=1}^{n} \mathcal{B}(X^k)$	Produkt-σ-Algebra
$\partial\Omega$	Rand von Ω
$\langle \cdot, \cdot \rangle$	Skalarprodukt
$\sigma(\mathcal{E})$	von $\mathcal{E}$ erzeugte σ- Algebra
$\sum_{i=1}^{d}$	Summe
sup	Supremum
$\subset$	Teilmenge von
$\subsetneq$	echte Teilmenge von
a^{T}	transponiertes Tupel zu a
$f^{-1}(V)$	Urbild von V
$\cup$	Vereinigung
$\dot{\cup}$	disjunkte Vereinigung
$\mathbb{Q}$	rationale Zahlen
$\mathbb{R}$	reelle Zahlen

2. Topologische Grundlagen

In diesem Kapitel werden Elemente der Topologie vorgestellt, welche für den weiteren Verlauf der Arbeit von Bedeutung sind und die dafür verwendeten Notationen festgelegt.

2.1. Euklidische Räume

Im Folgenden sei X eine nichtleere Menge, die mit entsprechenden Verknüpfungen einen Vektorraum über $\mathbb{R}$ darstellt, siehe [7] für eine Definition. Der Vektorraum wird der Einfachheit halber ebenfalls mit X bezeichnet.

Definition 2.1. *Eine Abbildung*

$$\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$$

wird ***Skalarprodukt*** *auf X genannt, falls gilt*

(i) $\forall x, y, z \in X \ \forall \alpha, \beta \in \mathbb{R} : \ \langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$,

(ii) $\forall x, y \in X : \ \langle x, y \rangle = \langle y, x \rangle$,

(iii) $\forall x \in X : \ \langle x, x \rangle \geq 0$ *und* $(\langle x, x \rangle = 0 \Rightarrow x = 0)$.

Das Paar $(X, \langle \cdot, \cdot \rangle)$ heißt dann ***euklidischer Raum****.*

Beispiel 2.2. Der Standardvertreter für einen euklidischen Vektorraum ist $X = \mathbb{R}^d$ mit Standard-Skalarprodukt. Dabei ist $\mathbb{R}^d$ ein Vektorraum über $\mathbb{R}$ mit festem $d \in \mathbb{N}$, dessen Elemente $x \in \mathbb{R}^d$ geordnete d-Tupel sind, die auch Vektoren genannt werden. Diese werden im Folgenden in der Regel als Zeilenvektor, d.h. $x = (x_1, \ldots, x_d)$ mit $x_i \in \mathbb{R}$ für $i = 1, \ldots, d$ notiert. Dennoch wird $x \in \mathbb{R}^d$ an manchen Stellen auch als Spaltenvektor aufgefasst, insbesondere bei der Beschreibung von linearen Funktionen. Das Standard-Skalarprodukt im $\mathbb{R}^d$ ist gegeben über

$$\langle x, y \rangle_{\mathbb{R}^d} \ := \ \sum_{i=1}^{d} x_i y_i \tag{2.1}$$

für $x, y \in \mathbb{R}^d$. Nachfolgend steht der Begriff „$\mathbb{R}^d$“ stets für den euklidischen Raum $(\mathbb{R}^d, \langle \cdot, \cdot \rangle_{\mathbb{R}^d})$.

Neben dem Skalarprodukt können auf Vektorräumen auch Normen betrachtet werden.

Normierte Räume

Definition 2.3. *Eine Abbildung*

$$\|\cdot\| : X \to [0, \infty)$$

wird ***Norm*** *auf X genannt, falls gilt*

(i) $\forall x \in X : \ \|x\| = 0 \ \Leftrightarrow \ x = 0,$

(ii) $\forall \alpha \in \mathbb{R}, \ x \in X : \ \|\alpha x\| = |\alpha| \|x\|,$

(iii) $\forall x, y \in X : \ \|x + y\| \leq \|x\| + \|y\|.$

Das Paar $(X, \|\cdot\|)$ heißt ***normierter Raum****.*

Sei $(X, \langle \cdot, \cdot \rangle)$ ein euklidischer Raum, dann ist die Abbildung

$$\|\cdot\| : X \to [0, \infty), \ x \mapsto \|x\| := \sqrt{\langle x, x \rangle}$$

eine Norm auf X, vgl. [7]. Somit ist jeder euklidische Raum auch ein normierter Raum. Im $\mathbb{R}^d$ ist die euklidische Norm oder 2-Norm eines Vektors $x \in \mathbb{R}^d$ definiert als

$$\|x\|_2 \ := \ \sqrt{\sum_{i=1}^{d} |x_i|^2},$$

was gerade der induzierten Norm durch das Standard-Skalarprodukt entspricht. Ein weiteres Beispiel für eine Norm im $\mathbb{R}^d$ ist die ∞-Norm, die definiert ist als

$$\|x\|_\infty \ := \ \max_{i=1,\ldots,d} |x_i|$$

für $x \in \mathbb{R}^d$.

Nun werden wichtige Eigenschaften von Mengen in normierten Räumen eingeführt.

Definition 2.4. *Seien $(X, \|\cdot\|)$ ein normierter Raum, $x_0 \in X$ und $r \in (0, \infty)$.*

(i) Eine ***offene Kugel um x_0 mit Radius r*** *ist definiert als*

$$B(x_0, r) \ := \ \{x \in X \mid \|x - x_0\| < r\}.$$

(ii) Eine Teilmenge $U \subset X$ heißt ***offen****, falls*

$$\forall x \in U \ \exists r \in (0, \infty) : \ B(x, r) \subset U.$$

(iii) Eine Teilmenge $U \subset X$ heißt ***abgeschlossen****, falls U^c offen ist.*

Es gilt, dass in normierten Räumen die beliebige Vereinigung offener Mengen offen ist, ebenso wie der Durchschnitt endlich vieler offener Mengen. Die Beweise hierzu finden sich beispielsweise in [22].
Da $\mathbb{Q}$ dicht in $\mathbb{R}$ ist, ist außerdem nachweisbar, dass jede offene Menge in $\mathbb{R}$ als Vereinigung offener Intervalle darstellbar ist, was ebenfalls in [22] begründet ist. Diese Eigenschaft lässt sich sofort auf offene Mengen im $\mathbb{R}^d$ erweitern. Hierbei sei bemerkt, dass die Eigenschaft, ob eine Menge im $\mathbb{R}^d$ offen ist oder nicht, nicht von der gewählten Norm abhängt, weil alle Normen auf dem $\mathbb{R}^d$ äquivalent sind.
Weitere bedeutsame Mengen in normierten Räumen sind zum Beispiel **kompakte Mengen**, siehe [7] für eine allgemeine Definition. Im $\mathbb{R}^d$ ist eine Menge genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Bemerkung 2.5. Generell lassen sich noch allgemeinere Räume als normierte Räume betrachten, wie zum Beispiel metrische oder topologische Räume, siehe [7] für weitere Informationen. Viele der hier für normierte Räume eingeführten Definitionen und Eigenschaften lassen sich direkt auf die soeben genannten Räume übertragen bzw. verallgemeinern.
Jeder normierte Raum $(X, \|\cdot\|)$ ist mit der Abbildung

$$d : X \times X \to [0, \infty), \ (x, y) \mapsto d(x, y) := \|x - y\|$$

ein metrischer Raum. Ferner bildet das Mengensystem der offenen Mengen

$$\mathcal{T} \ := \ \{U \subset X \mid U \text{ ist offen in } X\}$$

eine Topologie auf dem normierten Raum. Im euklidischen Raum wird die so definierte Topologie **euklidische Topologie** genannt.

Stetige und halbstetige Funktionen

Für Abbildungen zwischen normierten Räumen ist die Stetigkeit, welche beispielsweise in [13] definiert ist, eine bedeutsame Eigenschaft. Zum Nachweis der Stetigkeit einer Abbildung lässt sich oftmals folgende Äquivalenz verwenden.

Satz 2.6. *Seien $(X, \|\cdot\|)$ und $(Y, \|\cdot\|')$ normierte Räume und $f : X \to Y$ eine Abbildung. Dann sind die folgenden Aussagen äquivalent:*

(i) f ist stetig.

(ii) $\forall V \subset Y$, V offen: $f^{-1}(V) \subset X$ offen.

(iii) $\forall V \subset Y$, V abgeschlossen: $f^{-1}(V) \subset X$ abgeschlossen.

Beweis. Siehe [13]. □

Stetige Funktionen haben zum Beispiel die Eigenschaft, dass die Komposition stetiger Funktionen stetig ist. Für $Y = \mathbb{R}$ sind konstante Funktionen sowie endliche Summen bzw. Produkte stetiger Funktionen stetig. Außerdem nimmt jede stetige Funktion auf einer kompakten Menge ihr Maximum bzw. Minimum an. Für $Y = \mathbb{R}^m$, $m \in \mathbb{N}$, ist $f := (f_1, \ldots, f_m) : X \to \mathbb{R}^m$ genau dann stetig, wenn jede Komponente $f_j : X \to \mathbb{R}$, $j = 1, \ldots, m$, stetig ist, vgl. [14].

Beispiel 2.7. Bekannte Vertreter stetiger Funktionen sind beispielsweise Polynome $p : \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$. Um diese definieren zu können, wird zunächst die Multiindex-Schreibweise eingeführt. Ein Multiindex ist ein Element $\alpha \in \mathbb{N}_0^d$ mit Tupelschreibweise $\alpha = (\alpha_1, \ldots, \alpha_d)$. Für $x \in \mathbb{R}^d$ und $\alpha \in \mathbb{N}_0^d$ sind definiert

$$\begin{aligned} x^\alpha &:= \prod_{i=1}^{n} x_i^{\alpha_i}, \\ |\alpha| &:= \sum_{i=1}^{n} \alpha_i, \end{aligned}$$

wobei $0^0 := 1$ Konvention ist. Für jedes $r \in \mathbb{N}_0$ gibt es endlich viele Multiindizes $\alpha \in \mathbb{N}_0^d$ mit $|\alpha| \leq r$. Ist zu jedem Multiindex außerdem ein $a_\alpha \in \mathbb{R}$ gegeben mit $a_\alpha \neq 0$ für mindestens ein $\alpha \in \mathbb{N}_0^d$, so wird die Abbildung

$$p : \mathbb{R}^d \to \mathbb{R},\ x \mapsto p(x) = \sum_{|\alpha| \leq r} a_\alpha x^\alpha$$

Polynom vom Grad r genannt.
Jedes Polynom ist als Summe von Produkten stetiger Funktionen stetig. Dem liegt zugrunde, dass Koordinatenprojektionen stetig sind, worauf in Kapitel 2.2 nochmals genauer eingegangen wird.

Eine schwächere Eigenschaft von Funktionen als die Stetigkeit ist die Halbstetigkeit.

Definition 2.8. *Seien $(X, \|\cdot\|)$ ein normierter Raum und $f : X \to [-\infty, \infty]$ eine Funktion. Diese heißt*

*(i) nach **unten halbstetig**, wenn für jedes $\alpha \in \mathbb{R}$ die Menge $\{x \in X \mid f(x) \leq \alpha\}$ abgeschlossen bzw. die Menge $\{x \in X \mid f(x) > \alpha\}$ offen ist und*

*(ii) nach **oben halbstetig**, wenn für jedes $\alpha \in \mathbb{R}$ die Menge $\{x \in X \mid f(x) \geq \alpha\}$ abgeschlossen bzw. die Menge $\{x \in X \mid f(x) < \alpha\}$ offen ist.*

Es ist sofort ersichtlich, dass eine Funktion f genau dann unten halbstetig ist, wenn $-f$ nach oben halbstetig ist. Außerdem gilt, dass Funktionen $f : X \to \mathbb{R}$ genau dann stetig sind, wenn sie sowohl nach unten als auch nach oben halbstetig sind. Auf einer kompakten Menge nimmt jede nach unten halbstetige Funktion ihr Minimum und jede nach oben halbstetige Funktion ihr Maximum an. Diese Aussagen sind alle in [1] begründet bzw. bewiesen.
Zusätzlich wird dort gezeigt, dass die untere Halbstetigkeit äquivalent ist zur Gültigkeit von

$$\liminf_{y \to x} f(y) \quad \geq \quad f(x)$$

mit $\liminf_{y \to x} f(y) := \lim_{\delta \searrow 0} \left(\inf_{y \in B(x,\delta)} f(y) \right)$ für alle $x \in X$. Entsprechend ist die obere Halbstetigkeit äquivalent zu

$$\limsup_{y \to x} f(y) \quad \leq \quad f(x)$$

mit $\limsup_{y \to x} f(y) := \lim_{\delta \searrow 0} \left(\sup_{y \in B(x,\delta)} f(y) \right)$ für alle $x \in X$.

Beispiel 2.9. Sei $(X, \|\cdot\|)$ ein normierter Raum. Die **charakteristische Funktion von** $A \subset X$ ist definiert über

$$\chi_A : X \to \mathbb{R},\ x \mapsto \chi_A(x) := \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

Diese ist nach unten halbstetig, falls A offen ist und nach oben halbstetig, falls A abgeschlossen ist.

Eine weitere Eigenschaft von Funktionen, welche später ebenfalls aufgegriffen wird, ist die Monotonie. Diese wird nun für eine Funktion $f : D \to \mathbb{R}$ mit $D \subset \mathbb{R}^d$ definiert.

Definition 2.10. *Seien $D \subset \mathbb{R}^d$ und $f : D \to \mathbb{R}$ eine Funktion. Dann heißt f* ***monoton wachsend****, wenn für alle $x, y \in D$ mit $x \leq y$ gerade $f(x) \leq f(y)$ gilt. Entsprechend heißt f* ***monoton fallend****, wenn für alle $x, y \in D$ mit $x \leq y$ gerade $f(x) \geq f(y)$ gilt. Das Ungleichungssystem $x \leq y$ ist dabei stets komponentenweise zu verstehen.*

2.2. Produkträume

Nun werden kartesische Produkte von Mengen betrachtet.

Kartesische Produkte

Definition 2.11. *Das* ***kartesische Produkt*** *$A \times B$ der Mengen A und B ist definiert als die Menge aller geordneten Paare (a, b), für deren erste Komponente $a \in A$ und für deren zweite Komponente $b \in B$ gilt, d.h.*

$$A \times B \quad := \quad \{(a, b) \mid a \in A, b \in B\}.$$

Das kartesische Produkt besitzt unter anderem die nachfolgend aufgeführten Eigenschaften.

Satz 2.12. *Seien A, B, C und D Mengen. Für das kartesische Produkt gelten die Eigenschaften*

(i) $A \times (B \cap C) = (A \times B) \cap (A \times C)$ *bzw.* $A \times (B \cup C) = (A \times B) \cup (A \times C)$,

(ii) $A \times B = \emptyset \Leftrightarrow A = \emptyset$ *oder* $B = \emptyset$,

(iii) $A \times B \subset C \times D \Leftrightarrow A \subset C$ *und* $B \subset D$,

(iv) $(A \times B)^{\mathrm{c}} = (A^{\mathrm{c}} \times B^{c}) \cup (A^{\mathrm{c}} \times B) \cup (A \times B^{\mathrm{c}})$,

(v) $(A \cap B) \times (C \cap D) = (A \times C) \cap (B \times D)$.

Beweis. Siehe Anhang A. □

Bemerkung 2.13. Das kartesische Produkt ist im Allgemeinen weder kommutativ noch assoziativ. Für paarweise disjunkte Mengen $A := \{a\}$, $B := \{b\}$ und $C := \{c\}$ gilt beispielsweise nämlich

$$\begin{aligned} A \times B &= \{a\} \times \{b\} = \{(a,b)\} \\ &\neq \{(b,a)\} = \{b\} \times \{a\} = B \times A, \\ (A \times B) \times C &= (\{a\} \times \{b\}) \times \{b\} = \{((a,b),c)\} \\ &\neq \{(a,(b,c))\} = \{a\} \times (\{b\} \times \{c\}) = A \times (B \times C)\,. \end{aligned}$$

Obwohl das kartesische Produkt generell nicht kommutativ ist, ähneln sich die Mengen $A \times B$ und $B \times A$ insofern, dass je ein Element der einen Menge bis auf Umordnung einem Element der andern entspricht. Die Zuordnung dieser Elemente definiert eine Bijektion, weshalb $A \times B$ mit $B \times A$ identifiziert werden kann. Analog existiert solch eine kanonische Bijektion für die Mengen $(A \times B) \times C$ und $A \times (B \times C)$. Bevor auf eine solche Identifizierung genauer eingegangen wird, wird zunächst das kartesische Produkt auf endlich viele Mengen erweitert.

Definition 2.14. *Das **mehrfache kartesische Produkt** $A_1 \times \ldots \times A_d$ der Mengen $A_1, \ldots, A_d$, $d \in \mathbb{N}$ fest, ist definiert als die Menge aller geordneter d-Tupel $(a_1, \ldots, a_d)$, wobei $a_i \in A_i$ für $i = 1, \ldots, d$. Also gilt*

$$A_1 \times \ldots \times A_d := \{(a_1, \ldots, a_d) \mid a_i \in A_i,\ i = 1, \ldots, d\}.$$

Mithilfe des Produktzeichens kann das mehrfache kartesische Produkt durch

$$\prod_{i=1}^{d} A_i := A_1 \times \ldots \times A_d.$$

notiert werden. Dies ist nicht zu verwechseln mit dem Produkt von Skalaren, für welches das Produktzeichen „$\prod$“ in dieser Arbeit ebenfalls verwendet wird.
Die Eigenschaften des kartesischen Produkts zweier Mengen lassen sich direkt auf das mehrfache kartesische Produkt übertragen.

Allgemeines zu Produkträumen

In der folgenden Definition werden nun normierte Räume betrachtet, die auf dem mehrfachen kartesischen Produkt der Grundmengen von normierten Räumen beru-

hen. Dabei sei bemerkt, dass der verwendete, hochgestellte Index nicht mit einem Exponenten zu verwechseln ist.

Definition 2.15. *Seien $n \in \mathbb{N}$ und $(X^k, \|\cdot\|_{X^k})$, $k = 1, \ldots, n$, normierte Räume. Ferner sei $X := \prod_{k=1}^{n} X^k$ das n-fache kartesische Produkt der Grundmengen X^k, welches mit komponentenweiser Addition und Skalarmultiplikation einen Vektorraum über $\mathbb{R}$ darstellt. Dieser Vektorraum X beinhaltet alle n-Tupel $x = (x^1, \ldots, x^n)$ mit $x^k \in X^k$, $k = 1, \ldots, n$.*

(i) Wird X mit der Norm versehen, die für alle $x \in X$ definiert ist über

$$\|x\| \quad := \quad \sum_{k=1}^{n} \|x^k\|_{X^k}, \tag{2.2}$$

*so heißt der normierte Raum $(X, \|\cdot\|)$ **Produktraum** der Räume $(X^k, \|\cdot\|_{X^k})$, $k = 1, \ldots, n$.*
Für den Fall, dass $(X^k, \|\cdot\|_{X^k})$, $k = 1, \ldots, n$, sogar euklidische Räume mit Skalarprodukt $\langle\cdot,\cdot\rangle_{X^k}$ sind, ist auf X über

$$\langle x, y\rangle_X \quad := \quad \sum_{k=1}^{n} \langle x, y\rangle_{X^k} \tag{2.3}$$

für alle $x, y \in X$ ein Skalarprodukt definiert. Es sei bemerkt, dass hierbei die von dem Skalarprodukt (2.3) induzierte Norm in der Regel nicht mehr mit der Norm (2.2) des Produktraums $(X, \|\cdot\|)$ übereinstimmt.

*(ii) Die **kanonische Projektion**, welche von X auf X^k abbildet, ist definiert durch*

$$\pi^k : X \to X^k, \; x \mapsto x^k,$$

für $k \in \{1, \ldots, n\}$.

*(iii) Die **kanonische Einbettung** von X^k nach X ist definiert durch*

$$\iota^k : X^k \to X, \; x^k \mapsto (\tilde{x}^1, \ldots, \tilde{x}^{k-1}, x^k, \tilde{x}^{k+1}, \ldots, \tilde{x}^n)$$

für $k \in \{1, \ldots, n\}$ und feste $\tilde{x}^{\tilde{k}} \in X^{\tilde{k}}$ mit $\tilde{k} = 1, \ldots, n$, $\tilde{k} \neq k$.

Dass das in (2.3) definierte Skalarprodukt bzw. die in (2.2) definierte Norm auch tatsächlich ein Skalarprodukt bzw. eine Norm auf X darstellen, lässt sich unter Ver-

wendung der entsprechenden Definitionen direkt nachrechnen. Gleiches gilt für den Nachweis der Vektorraumeigenschaften von X.

Bemerkung 2.16. Falls $U^k \in X^k$ für $k \in \{1, \ldots, n\}$ offen ist, so ist das Urbild $\pi^{-1}(U^k) = X^1 \times \ldots \times X^{k-1} \times U^k \times X^{k+1} \times \ldots \times X^n$ ebenfalls offen in X. Denn für jedes $x \in X$ mit $x^k \in U^k$ existiert eine offene Kugel in X^k um x^k mit Radius r, die vollständig in U^k enthalten ist. Daher ist auch die offene Kugel in X um x mit Radius r vollständig in $\pi^{-1}(U^k)$ enthalten. Dies sagt außerdem aus, dass π^k für $k \in \{1, \ldots, n\}$ stetig ist. In der Literatur ist es gängig den umgekehrten Weg zu gehen und die Stetigkeit der kanonischen Projektion zu fordern, um so beispielsweise eine Topologie auf einem Produkt von topologischen Räumen zu erhalten.
Ebenso ist das Urbild $(\iota^k)^{-1}(U) = \{x^k \in X^k \mid (x^1, \ldots, x^n) \in U\}$ einer offenen Menge $U \in X$ für $k \in \{1, \ldots, n\}$ und feste $x^{\tilde{k}} \in X^{\tilde{k}}$ mit $\tilde{k} = 1, \ldots, n$ und $\tilde{k} \neq k$ offen, was die Stetigkeit von ι^k, impliziert.

Auf dem Produktraum $(X, \|\cdot\|)$ haben Abbildungen, die sich aus Funktionen zusammensetzen, welche jeweils auf einer der Mengen X^k definiert sind, spezielle Eigenschaften.

Definition 2.17. *Seien $n \in \mathbb{N}$ und $(X^k, \|\cdot\|^k)$, $k = 1, \ldots, n$, normierte Räume mit Produktraum $(X, \|\cdot\|)$. Eine Funktion $f = (f_1, \ldots, f_m) : X \to \mathbb{R}^m$, für welche Funktionen $f^k = (f_1^k, \ldots, f_m^k) : X^k \to \mathbb{R}^m$, $k = 1, \ldots, n$, existieren mit*

$$f(x) \quad = \quad \sum_{k=1}^{n} f^k(x^k)$$

*für alle $x \in X$, wird **additiv zerlegbar** genannt.*

Diese Form der Zerlegung ist bis auf konstante Terme eindeutig.

Satz 2.18. *Eine additiv zerlegbare Funktion $f : X \to \mathbb{R}^m$ ist genau dann stetig, wenn die Funktionen $f^k : X^k \to \mathbb{R}^m$ für $k = 1, \ldots, n$ stetig sind.*

Beweis. Da eine Funktion, die nach $\mathbb{R}^m$ abbildet, genau dann stetig ist, wenn jede ihrer m Komponenten stetig ist, genügt es die Äquivalenz für die j-te Komponente zu zeigen, $j \in \{1, \ldots, m\}$.
„$\Leftarrow$“: Seien $f_j^k : X^k \to \mathbb{R}$, $k = 1, \ldots, n$, stetig, so gilt für alle $x \in X$ bereits $f_j(x) = \sum_{k=1}^{n} (f_j^k \circ \pi^k)(x)$ und f_j ist als Komposition und Summe stetiger Funktionen stetig.

„⇒“: Sei f_j stetig. Für feste $\tilde{x}^{\tilde{k}} \in X^{\tilde{k}}$ mit $\tilde{k} = 1, \ldots, n$, $\tilde{k} \neq k$, gilt

$$f_j^k(x^k) \quad = \quad (f_j \circ \iota^k)(x^k) - \sum_{\substack{\tilde{k}=1 \\ \tilde{k} \neq k}}^{n} f_j^{\tilde{k}}(\tilde{x}^{\tilde{k}})$$

für alle $x^k \in X^k$. Da konstante Funktionen f_j und ι^k stetig sind und Kompositionen sowie Summen stetiger Funktionen stetig bleiben, ist auch f_j^k, $k = 1, \ldots, n$, stetig.

□

Im weiteren Verlauf der Arbeit werden weitere Eigenschaften von additiv zerlegbaren Funktionen eingeführt. Zunächst soll jedoch der $\mathbb{R}^d$ als Produktraum genauer untersucht und darin enthaltene Teilmengen betrachtet werden.

Der Produktraum $\mathbb{R}^d$

Beispiel 2.19. Sei $d \in \mathbb{N}$ fest.

(i) Das d-fache kartesische Produkt der reellen Zahlen ergibt die Menge

$$\prod_{i=1}^{d} \mathbb{R} \quad = \quad \{(x_1, \ldots, x_d) \mid x_1, \ldots, x_d \in \mathbb{R}\}.$$

Da die Elemente von $\prod_{i=1}^{d} \mathbb{R}$ und $\mathbb{R}^d$ identisch sind, gilt $\prod_{i=1}^{d} \mathbb{R} = \mathbb{R}^d$. Als euklidische Räume betrachtet stimmen das in (2.1) definierte Standard-Skalarprodukt auf dem $\mathbb{R}^d$ und das in (2.3) definierte Skalarprodukt überein.

(ii) Das d-fache kartesische Produkt der Intervalle $[a_i, b_i] \subset \mathbb{R}$ mit $a_i < b_i$ für $i = 1, \ldots, d$ ergibt die Menge

$$\prod_{i=1}^{d} [a_i, b_i] \;=\; \{(x_1, \ldots, x_d) \in \mathbb{R}^d \mid a_i \leq x_i \leq b_i \text{ für } i = 1, \ldots, d\} \;\subset\; \mathbb{R}^d.$$

Diese wird aufgrund ihrer geometrischen Darstellung auch als d-dimensionales Rechteck bzw. Hyperrechteck bezeichnet.

Aufgrund der Nicht-Assoziativität von mehrfachen kartesischen Produkten ist gene-

rell

$$\begin{aligned}\mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^n} &= (\underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{d^1\text{-mal}}) \times \ldots \times (\underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{d^n\text{-mal}}) \\ &\neq \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{d\text{-mal}} = \mathbb{R}^d\end{aligned}$$

für festes $n \in \mathbb{N}$ und feste $d^k \in \mathbb{N}$, $k = 1, \ldots, n$, mit $d := \sum_{k=1}^{n} d^k$. Allerdings existiert wie oben erwähnt eine kanonische Bijektion $\mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^n} \to \mathbb{R}^d,\ (x^1, \ldots, x^n) \mapsto (x_1, \ldots, x_d)$ mit $x^k = (x^k_1, \ldots, x^k_{d^k}) = \left(x_{i^{k-1}+1}, \ldots, x_{i^k}\right) \in \mathbb{R}^{d^k}$ und $i^k := \sum_{\tilde{k}=1}^{k} d^{\tilde{k}}$ für $k = 1, \ldots, n$, welche die Mengen $\mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_n}$ und $\mathbb{R}^d$ miteinander identifiziert. Da außerdem

$$\sum_{k=1}^{n} \langle x^k, y^k \rangle_{\mathbb{R}^{d^k}} = \sum_{k=1}^{n} \sum_{\tilde{\imath} = i^{k-1}+1}^{i^k} x_{\tilde{\imath}} y_{\tilde{\imath}} = \sum_{i=1}^{d} x_i y_i = \langle x, y \rangle_{\mathbb{R}^d}$$

gilt, sind die euklidischen Vektorräume $\mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^n}$ und $\mathbb{R}^d$ isomorph. Deshalb kann ein n-Tupel aus $\mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^n}$ mit einem d-Tupel aus $\mathbb{R}^d$ gleichgesetzt werden, sowie $\mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^n} = \mathbb{R}^d$ in der Betrachtung als euklidische Räume. Diese Identifikation wird im Folgenden für die gesamte Arbeit vorausgesetzt. Insgesamt gilt dann

$$x = \left(x^1, \ldots, x^n\right) = (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{2.4}$$

Entsprechend ist $\prod_{k=1}^{n} A^k \subseteq \mathbb{R}^d$ für Mengen $A^k \subseteq \mathbb{R}^{d^k}$, $k = 1, \ldots, n$.

Insbesondere als Teilmengen des $\mathbb{R}^d$ mit $d \leq 3$ lassen sich die Elemente von mehrfachen kartesischen Produkten geometrisch als Punktmenge darstellen. Dazu werden die Komponenten des d-Tupels $(x_1, \ldots, x_d)$ im kartesischen Koordinatensystem als Koordinaten eines Punktes aufgefasst.
Beispielsweise lassen sich alle Paare der Menge $A = [1, 3] \times [1, 2] \subset \mathbb{R}^2$ als Punkte innerhalb eines Rechtecks visualisieren, wie in Abbildung 2.1 zu erkennen ist. Ebenso können alle Tripel der Menge $B = \{(x_1, x_2) \in \mathbb{R}^2 \mid (x_1)^2 + (x_2)^2 \leq 1\} \times [-1, 1] \subset \mathbb{R}^3$ als Punkte innerhalb eines Zylinders gedeutet werden, siehe Abbildung 2.2.

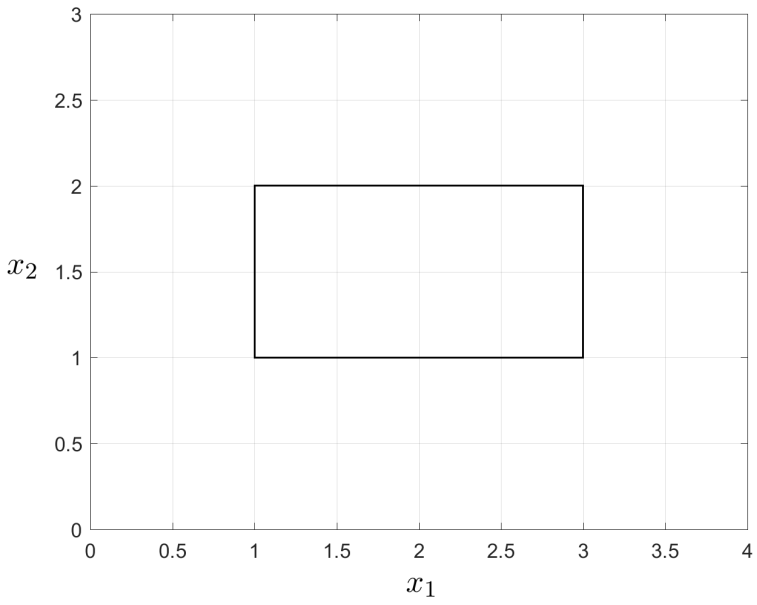


Abbildung 2.1: Darstellung der Menge A (schwarz umrandet)

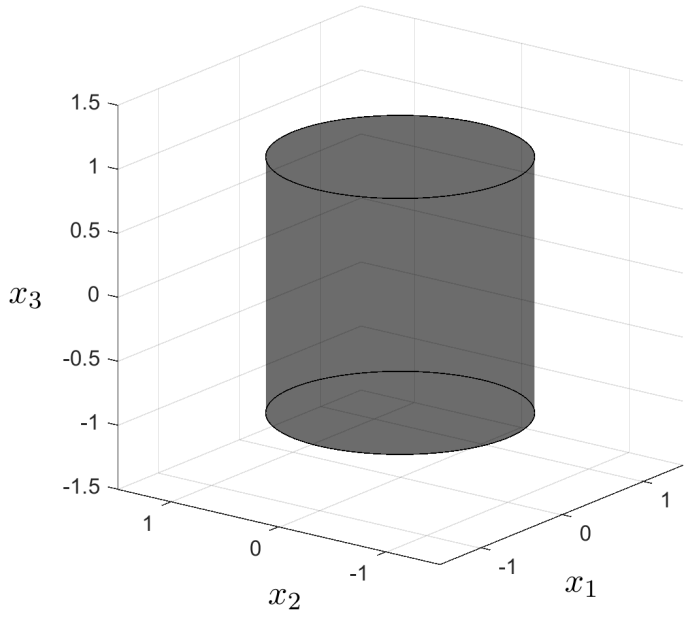


Abbildung 2.2: Darstellung der Menge B (transparent schwarz umrandet)

3. Maßtheoretischer Volumenbegriff

Um das Volumen von Teilmengen des $\mathbb{R}^d$ zu bestimmen, wird oftmals das Lebesgue-Maß verwendet. Zunächst werden dazu messbare Mengen und Funktionen eingeführt und anschließend der Begriff eines Maßes.

3.1. Borel-Mengen

In diesem Kapitel sei X stets eine nichtleere Menge, die nicht notwendigerweise ein Vektorraum über $\mathbb{R}$ darstellt.

Allgemeines zu σ-Algebren

Definition 3.1. *Sei $\mathcal{A} \subset \mathcal{P}(X)$. Dann heißt $\mathcal{A}$ eine **σ-Algebra über** X, falls gilt*

(i) $\emptyset \in \mathcal{A}$,

(ii) $A \in \mathcal{A} \Rightarrow A^c := \{x \in X : x \notin A\} \in \mathcal{A}$,

(iii) $A_j \in \mathcal{A}$, $j \in \mathbb{N} \Rightarrow \bigcup\limits_{j\in\mathbb{N}} A_j \in \mathcal{A}$.

In diesem Fall wird das Paar $(X, \mathcal{A})$ Messraum genannt und die Mengen $A \in \mathcal{A}$ heißen $\mathcal{A}$-messbar.

Aus der Definition einer σ-Algebra $\mathcal{A}$ kann direkt gefolgert werden, dass auch

$$\begin{aligned} X = (\emptyset)^{\mathrm{c}} &\in \mathcal{A}, \\ \bigcap_{j\in\mathbb{N}} A_j = \left(\bigcup_{j\in\mathbb{N}} (A_j)^{\mathrm{c}} \right)^{\mathrm{c}} &\in \mathcal{A}, \\ A_1 \backslash A_2 = A_1 \cap A_2^{\mathrm{c}} &\in \mathcal{A} \end{aligned}$$

für $A_j \in \mathcal{A}$, $j \in \mathbb{N}$ gilt.

Für eine abzählbare Menge X wird eine σ-Algebra oftmals als Potenzmenge von X gewählt. Für eine überabzählbare Menge führt dies zu Schwierigkeiten bei der Einführung eines Maßbegriffs, weshalb σ-Algebren in der Regel durch Erzeugendensysteme konstruiert werden.

Definition 3.2. *Seien $\mathcal{E} \subset \mathcal{P}(X)$ ein Mengensystem und $\sigma(\mathcal{E})$ die kleinste σ-Algebra, die $\mathcal{E}$ umfasst, d.h. für jede σ-Algebra $\mathcal{A}$ über X mit $\mathcal{E} \subset \mathcal{A}$ gilt auch $\sigma(\mathcal{E}) \subset \mathcal{A}$. Dann*

wird $\mathcal{E}$ auch Erzeugendensystem oder kurz ***Erzeuger der σ-Algebra*** *$\sigma(\mathcal{E})$ über X genannt und $\sigma(\mathcal{E})$ „die von $\mathcal{E}$ erzeugte σ-Algebra“.*

Satz 3.3. *Für beliebiges $\mathcal{E} \subset \mathcal{P}(X)$ ist die von $\mathcal{E}$ erzeugte σ-Algebra gegeben als der Durchschnitt aller $\mathcal{E}$ umfassenden σ-Algebren über X, genauer*

$$\sigma(\mathcal{E}) \quad = \quad \bigcap \{\mathcal{A} \mid \mathcal{A} \text{ ist } \sigma\text{-Algebra über } X \text{ und } \mathcal{E} \subset \mathcal{A}\}.$$

Beweis. Siehe Anhang A. □

Satz 3.4. *Seien $\mathcal{E}_1$ ein Erzeuger der σ-Algebra $\mathcal{A}_1$ und $\mathcal{E}_2$ ein Erzeuger der σ-Algebra $\mathcal{A}_2$. Falls $\mathcal{E}_1 \subset \mathcal{A}_2$ und $\mathcal{E}_2 \subset \mathcal{A}_1$, dann gilt $\mathcal{A}_1 = \mathcal{A}_2$.*

Beweis. Siehe Anhang A. □

Borel-σ-Algebren

Der Standardvertreter einer σ-Algebra ist die Borel-σ-Algebra über dem $\mathbb{R}^d$ auf welcher auch das Hauptaugenmerk dieser Arbeit gelegt ist. Dabei sei bemerkt, dass Borel-σ-Algebren generell auch für beliebige topologische Räume formuliert werden können.

Definition 3.5. *Seien $(X, \|\cdot\|)$ ein normierter Raum und $\mathcal{T}$ das System der offenen Teilmengen von X. Die* ***Borel-σ-Algebra*** *über X ist die von den offenen Teilmengen erzeugte σ-Algebra, genauer*

$$\mathcal{B}(X) \quad := \quad \sigma(\mathcal{T}).$$

Die Elemente von $\mathcal{B}(X)$ heißen ***Borel-Mengen****.*

Die Borel-σ-Algebra über dem $\mathbb{R}^d$ wird entsprechend mit $\mathcal{B}(\mathbb{R}^d)$ bezeichnet. Zum besseren Verständnis dieser Borel-σ-Algebra werden nun verschiedene Erzeuger vorgestellt. Dazu wird für $a, b \in \mathbb{R}^d$ mit $a < b$ das d-dimensionale halboffene Rechteck über

$$(a, b] \; := \; \Big\{ x \in \mathbb{R}^d \mid a_i < x_i \leq b_i, \; i = 1, \ldots, d \Big\} \; = \; \prod_{i=1}^{d} (a_i, b_i] \tag{3.1}$$

festgelegt.

Satz 3.6. *Die Mengensysteme*

(i) $\mathcal{T}^d = \{U \subset \mathbb{R}^d \mid U \text{ ist offen in } \mathbb{R}^d\}$,

(ii) $\mathcal{C}^d = \{U \subset \mathbb{R}^d \mid U \text{ ist abgeschlossen in } \mathbb{R}^d\}$,

(iii) $\mathcal{I}^d = \{(a, b] \mid a, b \in \mathbb{R}^d, a < b\}$,

(iv) $\mathcal{I}^d_\infty = \{(a, \infty) \mid a \in \mathbb{R}^d\}$

sind Erzeuger der Borel-σ-Algebra $\mathcal{B}(\mathbb{R}^d)$, wobei (a, ∞) für $a \in \mathbb{R}^d$ analog zu (3.1) erklärt ist.

Beweis. Siehe Anhang A. □

Bemerkung 3.7. Die Borel-σ-Algebra im $\mathbb{R}^d$ enthält nicht jede Teilmenge des euklidischen Raums, d.h. $\mathcal{B}(\mathbb{R}^d) \neq \mathcal{P}(\mathbb{R}^d)$. Allerdings enthält diese so gut wie alle Teilmengen von $\mathbb{R}^d$, die für die Anwendung benötigt werden. Daher ist es in der Regel ausreichend, ausschließlich Borel-Mengen als Teilmengen des $\mathbb{R}^d$ zu betrachten. Darüber hinaus kann auf $\mathcal{P}(\mathbb{R}^d)$ kein Maß definiert werden, das einem Volumenbegriff gerecht wird, vgl. [18].

Produkt-σ-Algebren

Analog zur Norm der Produkträumen aus Kapitel 2.2 kann eine Borel-σ-Algebra auch über ein n-faches Produkt von Borel-σ-Algebren dargestellt werden. Dazu werden zunächst die sogenannten Rechtecksmengen eingeführt, welche für diese Arbeit von zentraler Bedeutung sind.

Definition 3.8. *Seien $n \in \mathbb{N}$ fest und $(X^k, \|\cdot\|_{X^k})$, $k = 1, \ldots, n$, normierte Räume. Teilmengen Ω von $X := \prod_{k=1}^{n} X^k$ der Gestalt*

$$\Omega = \prod_{k=1}^{n} \Omega^k \text{ mit } \Omega^k \in \mathcal{B}(X^k) \tag{3.2}$$

werden ***Rechtecksmengen*** *genannt.*

Definition 3.9. *Seien $n \in \mathbb{N}$ fest und $(X^k, \|\cdot\|^k)$, $k = 1, \ldots, n$, normierte Räume. Die vom System aller Rechtecksmengen erzeugte σ-Algebra wird* ***Produkt-σ-Algebra***

der $\left(\mathcal{B}(X^k)\right)_{k\in\{1,\dots,n\}}$ *über dem Produktraum genannt und ist definiert als*

$$\bigotimes_{k=1}^{n} \mathcal{B}(X^k) := \sigma\left(\left\{\prod_{k=1}^{n} \Omega^k \mid \Omega^k \in \mathcal{B}(X^k),\ k = 1,\dots,n\right\}\right).$$

Bemerkung 3.10. In der Literatur ist es üblich Produkt-σ-Algebren allgemein über sogenannte **Zylindermengen** zu erzeugen. Eine Zylindermenge ist definiert als Urbild der Menge $\Omega^k \in \mathcal{B}(X^k)$, $k \in \{1,\dots,n\}$ unter der aus (2.4) bekannten Projektionsabbildung π^k und besitzt die Darstellung

$$(\pi^k)^{-1}(\Omega^k) = X^1 \times \dots \times X^{k-1} \times \Omega^k \times X^{k+1} \times \dots \times X^n,$$

vgl. Bemerkung 2.16. Aus dieser Darstellung folgt $(\pi^k)^{-1}(\Omega^k) \in \bigotimes_{k=1}^{n} \mathcal{B}(X^k)$. Mit $\prod_{k=1}^{n} \Omega^k = \bigcap_{k=1}^{n} (\pi^k)^{-1}(\Omega^k)$ und Satz 3.4 kann nun direkt eingesehen werden, dass beide Vorgehensweisen zur selben Produkt-σ-Algebra führen.

Im $\mathbb{R}^d$ kann die Borel-σ-Algebra $\mathcal{B}(\mathbb{R}^d)$ wahlweise als von den in Satz 3.6 definierten Mengensystemen erzeugte σ-Algebra oder als Produkt-σ-Algebra, vgl. Definition 3.9, aufgefasst werden. Dies zeigt der folgende Satz.

Satz 3.11. *Seien* $n \in \mathbb{N}$ *und* $d^k \in \mathbb{N}$ *fest für* $k = 1,\dots,n$ *mit* $d := \sum_{k=1}^{n} d^k$. *Dann gilt*

$$\bigotimes_{k=1}^{n} \mathcal{B}(\mathbb{R}^{d^k}) = \mathcal{B}(\mathbb{R}^d).$$

Beweis. Siehe Anhang A. □

Bemerkung 3.12. Aufgrund der Definition der Produkt-σ-Algebra ist jede Rechtecksmenge $\prod_{k=1}^{n} \Omega^k$ mit $\Omega^k \in \mathbb{R}^{d^k}$ nach Satz 3.11 eine Borel-Menge in $\mathbb{R}^d$. Umgekehrt ist jedoch nicht jede Borel-Menge in $\mathbb{R}^d$ eine Rechtecksmenge, wie das folgende Beispiel zeigt. Seien $d = 2$ und $n = 2$, d.h. $d^1 = d^2 = 1$. Die Menge $\Omega := \{x \in \mathbb{R}^2 \mid \|x\|_2 \leq 1\}$ ist abgeschlossen und somit gilt $\Omega \in \mathcal{B}(\mathbb{R}^d)$. Angenommen es existiert eine Rechtecksmenge $\Omega = \Omega^1 \times \Omega^2$ mit $\Omega^1, \Omega^2 \in \mathcal{B}(\mathbb{R})$, dann müsste wegen $(1,0) \in \Omega$ und $(0,1) \in \Omega$ auch $0,1 \in \Omega^1$ und $0,1 \in \Omega^2$ gelten. Dies bedeutet aber auch, dass $(1,1) \in \Omega$ sein müsste, was einen Widerspruch darstellt.

3.2. Borel-messbare Funktionen

Für eine Abbildung $f : X \to Y$ ist das Urbild einer σ-Algebra $\mathcal{S} \subset \mathcal{P}(Y)$ mit dem Urbild einer Menge $B \subset Y$ definiert über $f^{-1}(\mathcal{S}) := \{f^{-1}(B) \mid B \in \mathcal{S}\}$.

Allgemeines zu messbaren Funktionen

Definition 3.13. *Seien $(X, \mathcal{A})$ und $(Y, \mathcal{S})$ Messräume. Eine Funktion $f : X \to Y$ heißt **messbar**, genauer $\mathcal{A}$-$\mathcal{S}$-messbar, falls Urbilder messbarer Mengen messbar sind, genauer falls $f^{-1}(\mathcal{S}) \subset \mathcal{A}$.*

Gilt $(X, \mathcal{A}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)$ und $(Y, \mathcal{S}) = (\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m)$, so wird eine $\mathcal{A}$-$\mathcal{S}$-messbare Funktion auch Borel-messbar genannt. Da in der Regel bekannt ist, um welche σ-Algebren es sich handelt, werden diese im weiteren Verlauf der Arbeit nicht immer explizit erwähnt.

Bei der Untersuchung der Messbarkeit von Funktionen reicht es aus, sich auf ein Erzeugendensystem zu beschränken, wie der folgende Satz zeigt.

Satz 3.14. *Seien $(X, \mathcal{A})$ und $(Y, \mathcal{S})$ Messräume und $\mathcal{E}$ ein Erzeuger von $\mathcal{S}$, also $\mathcal{S} = \sigma(\mathcal{E})$. Dann ist $f : X \to Y$ genau dann messbar, falls*

$$f^{-1}(\mathcal{E}) \subset \mathcal{A}. \tag{3.3}$$

Beweis. Wegen

$$\begin{aligned} f^{-1}\left(\bigcup_{j\in\mathbb{N}} B_j\right) &= \left\{x \in X \mid f(x) \in \bigcup_{j\in\mathbb{N}} B_j\right\} = \bigcup_{j\in\mathbb{N}} \{x \in X \mid f(x) \in B_j\} \\ &= \bigcup_{j\in\mathbb{N}} f^{-1}(B_j), \\ f^{-1}(B^{\mathrm{c}}) &= \{x \in X \mid f(x) \in B^{\mathrm{c}}\} = X \backslash \{x \in X \mid f(x) \in B\} \\ &= X \backslash f^{-1}(B) = \left(f^{-1}(B)\right)^{\mathrm{c}} \end{aligned}$$

für $B, B_j \in \mathcal{P}(Y)$ und $f^{-1}(Y) = X$ ist das Mengensystem

$$S' := \{B \in \mathcal{P}(Y) \mid f^{-1}(B) \in \mathcal{A}\}$$

eine σ-Algebra. Nach Definition 3.13 ist f genau dann messbar, wenn $\mathcal{S} = \sigma(\mathcal{E}) \subset \mathcal{S}'$. Dies ist äquivalent zu $\mathcal{E} \subset \mathcal{S}'$, da $\mathcal{S}$ die kleinste σ-Algebra ist, welche $\mathcal{E}$ enthält. Dies ist wiederum der Fall, falls Bedingung (3.3) erfüllt ist. □

Hierbei sei bemerkt, dass der Beweis dieses Satzes den Beweis des Satzes 3.11 teilweise verallgemeinert. Für Abbildungen $f : X \to \mathbb{R}$ lässt sich aus Satz 3.14 direkt die nachstehende Äquivalenz ableiten.

Korollar 3.15. *Seien $(X, \mathcal{A})$ ein Messraum und $f : X \to \mathbb{R}$ eine Funktion. Dann sind die folgenden Aussagen äquivalent:*

(i) f ist $\mathcal{A} - \mathcal{B}(\mathbb{R})$-messbar.

(ii) Für alle $\alpha \in \mathbb{R}$ gilt $\{x \in X \mid f(x) > \alpha\} \in \mathcal{A}$.

(iii) Für alle $\alpha \in \mathbb{R}$ gilt $\{x \in X \mid f(x) \geq \alpha\} \in \mathcal{A}$.

(iv) Für alle $\alpha \in \mathbb{R}$ gilt $\{x \in X \mid f(x) < \alpha\} \in \mathcal{A}$.

(v) Für alle $\alpha \in \mathbb{R}$ gilt $\{x \in X \mid f(x) \leq \alpha\} \in \mathcal{A}$.

Beweis. Es gilt $\{x \in X \mid f(x) > \alpha\} = f^{-1}((\alpha, \infty))$. Da nach Satz 3.6 das Mengensystem $\mathcal{I}^d_\infty = \{(\alpha, \infty) \mid \alpha \in \mathbb{R}\}$ ein Erzeuger der Borel-σ-Algebra $\mathcal{B}(\mathbb{R})$ ist, folgt mit Satz 3.14 die Äquivalenz von (i) und (ii). Die Äquivalenz von (ii) - (v) ergibt sich aus

$$\begin{aligned}
\{x \in X \mid f(x) \geq \alpha\} &= \bigcap_{k \in \mathbb{N}} \{x \in X \mid f(x) > \alpha - \frac{1}{k}\}, \\
\{x \in X \mid f(x) < \alpha\} &= \{x \in X \mid f(x) \geq \alpha\}^{\mathrm{c}}, \\
\{x \in X \mid f(x) \leq \alpha\} &= \bigcap_{k \in \mathbb{N}} \{x \in X \mid f(x) < \alpha + \frac{1}{k}\}, \\
\{x \in X \mid f(x) > \alpha\} &= \{x \in X \mid f(x) \leq \alpha\}^{\mathrm{c}}.
\end{aligned}$$

□

Mengen der Gestalt $\{x \in X \mid f(x) \leq \alpha\}$ werden im weiteren Verlauf der Arbeit noch des Öfteren auftauchen.

Definition 3.16. *Sei $f : X \to \mathbb{R}$ eine Funktion. Die Menge*

$$\{x \in X \mid f(x) \leq \alpha\} = f^{-1}((-\infty, \alpha])$$

*wird α-**Subniveaumenge** von f genannt.*

Bevor einige elementare Eigenschaften von messbaren Funktionen vorgestellt und bewiesen werden, werden zunächst Beispiele messbarer Funktionen aufgeführt.

Beispiel 3.17. (i) Seien $(X, \|\cdot\|)$ und $(Y, \|\cdot\|')$ normierte Räume und $f : X \to Y$ eine stetige Abbildung, dann ist f gerade $\mathcal{B}(X)$-$\mathcal{B}(Y)$-messbar. Nach Definition der Stetigkeit, vgl. Satz 2.6, sind nämlich Urbilder offener Mengen offen und offene Mengen sind Erzeuger von Borel-σ-Algebren. Mit Satz 3.14 folgt somit sofort die Messbarkeit von f.
Insbesondere sind damit die in Definition 2.15 definierten kanonischen Projektionen und Einbettungen messbar.

(ii) Sei $(X, \mathcal{A})$ ein Messraum. Die charakteristische Funktion χ_A, definiert wie in Beispiel 2.9, ist messbar, falls $A \in \mathcal{A}$. Dies gilt, da die Urbilder von χ_A nur die Mengen $\emptyset$, A, A^c und X sein können, welche nur dann in $\mathcal{A}$ enthalten sind, wenn $A \in \mathcal{A}$ gilt.

Eigenschaften messbarer Funktionen

Die folgenden Eigenschaften sind bereits von den stetigen Funktionen bekannt und lassen sich auf messbare Funktionen erweitern.

Satz 3.18. *Seien $(X, \mathcal{A})$, $(Y_1, \mathcal{S}_1)$ und $(Y_2, \mathcal{S}_2)$ Messräume und seien $f : X \to Y_1$ und $g : Y_1 \to Y_2$ messbar, so ist auch $g \circ f : X \to Y_2$ messbar.*

Beweis. Für beliebiges $B \in \mathcal{S}_2$ ist $g^{-1}(B) \in \mathcal{S}_1$ und weiter $f^{-1}(g^{-1}(B)) \in \mathcal{A}$. Insgesamt gilt also $(g \circ f)^{-1}(B) \in \mathcal{A}$. □

Satz 3.19. *Seien $(X, \mathcal{A})$ ein Messraum und $f := (f_1, \ldots, f_m) : X \to \mathbb{R}^m$. Die Funktion f ist genau dann $\mathcal{A}$-$\mathcal{B}(\mathbb{R}^m)$-messbar, wenn die Funktionen $f_j : X \to \mathbb{R}$, $j = 1, \ldots, m$, $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-messbar sind.*

Beweis. Analog zu Abbildung (2.4) sei $\pi_j : \mathbb{R}^m \to \mathbb{R}$, $j \in \{1, \ldots, m\}$, die Projektion auf die j-te Komponente. Wie in Beispiel 3.17 beschrieben ist π_j deshalb Borel-messbar.
„⇒“ Sei f messbar, so ist mit Satz 3.18 auch $\pi_j \circ f$ für $j = 1, \ldots, m$ messbar.
„⇐“ Seien $f_j : X \to \mathbb{R}$ für $j = 1, \ldots, m$ nun $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-messbar, d.h. für alle $B_j \in \mathcal{B}(\mathbb{R})$ gilt $(\pi_j \circ f)^{-1}(B) \in \mathcal{A}$. Da nach Bemerkung 3.10 und Satz 3.11 die Mengen $\pi_j^{-1}(B)$ die Borel-σ-Algebra $\mathcal{B}(\mathbb{R}^m)$ erzeugen, folgt mit Satz 3.14 die Messbarkeit von f. □

Aus den Sätzen 3.18 und 3.19 kann hergeleitet werden, dass die Summe von messbaren Funktionen, welche nach $\mathbb{R}$ abbilden, ebenfalls messbar ist.

Satz 3.20. *Seien $(X, \mathcal{A})$ ein Messraum, $f, g : X \to \mathbb{R}$ $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-messbare Funktionen und $\alpha, \beta \in \mathbb{R}$. Dann ist auch $\alpha f + \beta g$ sowie $f \cdot g$ messbar.*

Beweis. Da f und g messbar sind, ist mit Satz 3.19 auch die Abbildung

$$h_1 : X \to \mathbb{R}^2,\ x \mapsto (f(x), g(x))$$

messbar. Ferner ist die Abbildung

$$h_2 : \mathbb{R}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto \alpha x_1 + \beta x_2$$

stetig, vgl. Beispiel 2.7 und damit ebenfalls messbar. Daraus ergibt sich mit Satz 3.18 die Messbarkeit von $h_2 \circ h_1 = \alpha f + \beta g$. Entsprechend kann die Messbarkeit von $f \cdot g$ mit $h_2 : \mathbb{R}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto x_1 x_2$ gezeigt werden. □

Satz 3.21. *Seien $n \in \mathbb{N}$ und $(X^k, \|\cdot\|_{X^k})$ normierte Räume, die mit den Borel-σ-Algebren $\mathcal{B}(X^k)$ für $k = 1, \ldots, n$ Messräume darstellen. Der Produktraum dieser Räume ist $(X, \|\cdot\|)$ mit Borel-σ-Algebra $\mathcal{B}(X)$. Falls eine Funktion $f : X \to \mathbb{R}^m$ nach Definition 2.17 additiv zerlegbar ist mit $f(x) = \sum_{k=1}^{n} f^k(x^k)$ für $x \in X$ und $f^k : X^k \to \mathbb{R}^m$, so ist f genau dann $\mathcal{B}(X)$-$\mathcal{B}(\mathbb{R}^m)$-messbar, wenn die Funktionen f^k $\mathcal{B}(X^k)$-$\mathcal{B}(\mathbb{R}^m)$-messbar sind, $k = 1, \ldots, n$.*

Beweis. Da nach Satz 3.19 eine Funktion, die nach $\mathbb{R}^m$ abbildet, genau dann messbar ist, wenn jede ihrer m Komponenten messbar ist, genügt es, die Äquivalenz für die j-te Komponente zu zeigen mit $j \in \{1, \ldots, m\}$.
„⇐“: Seien $f_j^k : X^k \to \mathbb{R}$ für $k = 1, \ldots, n$ messbar, so gilt für alle $x \in X$ gerade $f_j(x) = \sum_{k=1}^{n} (f_j^k \circ \pi^k)(x)$ und f_j ist als Komposition und Summe messbarer Funktionen nach den Sätzen 3.18 und 3.20 messbar.
„⇒“: Sei f_j messbar. Mit der kanonischen Einbettung, welche in Definition 2.15 eingeführt wurde und stetig und somit messbar ist, gilt für $k \in \{1, \ldots, n\}$ und feste $\tilde{x}^{\tilde{k}} \in X^{\tilde{k}}$ mit $\tilde{k} = 1, \ldots, n$, $\tilde{k} \neq k$, nun

$$(f_j \circ \iota^k)(x^k) \quad = \quad f_j^k(x^k) + \sum_{\substack{\tilde{k}=1 \\ \tilde{k} \neq k}}^{n} f_j^{\tilde{k}}\big(\tilde{x}^{\tilde{k}}\big)$$

für alle $x^k \in X^k$. Somit ist das Urbild von $(b, \infty) \in \mathcal{I}_\infty^1$, $b \in \mathbb{R}$ unter f_j^k darstellbar

als

$$(f_j^k)^{-1}((b,\infty)) \quad = \quad (f_j \circ \iota^k)^{-1}\big((\tilde{b},\infty)\big),$$

wobei $\tilde{b} = b + \sum_{\substack{\tilde{k}=1 \\ \tilde{k}\neq k}}^{n} f_j^{\tilde{k}}(\tilde{x}^{\tilde{k}}) \in \mathbb{R}$. Aufgrund der Messbarkeit von f_j und ι^k ist dann $(f_j^k)^{-1}((b,\infty)) \in \mathcal{B}(X^k)$ und somit auch f_j^k wegen Satz 3.14 für $k = 1, \ldots, n$ messbar. □

3.3. Das Lebesgue-Maß

Ein Maß ordnet einer messbaren Menge einen nicht-negativen Wert zu, wie zum Beispiel das Volumen der Menge.

Allgemeines zu Maßen

Definition 3.22. *Sei $(X, \mathcal{A})$ ein Messraum. Eine Abbildung $\mu : \mathcal{A} \to [0,\infty]$ heißt **Maß auf** $\mathcal{A}$, wenn gilt*

(i) $\mu(\emptyset) = 0$,

(ii) *μ ist σ-additiv, d.h. für jede Folge paarweiser disjunkter Mengen $(A_i)_{i\in\mathbb{N}} \subset \mathcal{A}$ gilt*

$$\mu\left(\dot{\bigcup_{i\in\mathbb{N}}} A_i\right) \quad = \quad \sum_{i\in\mathbb{N}} \mu(A_i).$$

Das Tripel $(X, \mathcal{A}, \mu)$ wird dann Maßraum genannt.
*Ein Maß μ auf $\mathcal{A}$ heißt **σ-endlich**, falls es eine Folge $(A_i)_{i\in\mathbb{N}} \subset \mathcal{A}$ mit $\bigcup_{i\in\mathbb{N}} A_i = X$ und $\mu(A_i) < \infty$, $i \in \mathbb{N}$ gibt.*

In einem Maßraum $(X, \mathcal{A}, \mu)$ werden Mengen $A \in \mathcal{A}$, welche sich bzgl. des Maßes nicht von der leeren Menge unterscheiden, also $\mu(A) = 0$ gilt, Nullmengen genannt. Aussagen, die für alle Elemente von X bis auf die Elemente einer Nullmenge gelten, werden „fast überall“ gültig genannt.

Im folgenden Satz sind Eigenschaften des Maßes aufgeführt, die sich direkt aus der σ-Additivität ergeben. Für die Mengen $A, A_1, A_2, \ldots \in \mathcal{A}$ werden dazu die Schreib-

weisen $A_i \uparrow A$, falls $A_1 \subset A_2 \subset \ldots$ und $A = \bigcup_{i\in\mathbb{N}} A_i \in \mathcal{A}$ gilt und $A_i \downarrow A$, falls $A_1 \supset A_2 \supset \ldots$ und $A = \bigcap_{i\in\mathbb{N}} A_i \in \mathcal{A}$ gilt, verwendet.

Satz 3.23. *Seien $(X, \mathcal{A}, \mu)$ ein Maßraum und $A_i \in \mathcal{A}$, $i \in \mathbb{N}$. Dann gelten*

(i) Subtraktivität, d.h. für $A_1 \subset A_2$ und $\mu(A_1) < \infty$ ist $\mu(A_2 \backslash A_1) = \mu(A_2) - \mu(A_1)$,

(ii) Monotonie, d.h. für $A_1 \subset A_2$ ist $\mu(A_1) \leq \mu(A_2)$,

(iii) σ-Stetigkeit, d.h. für $A_i \uparrow A$ gilt $\lim_{i\to\infty} \mu(A_i) = \mu(A)$ und für $A_i \downarrow A$ mit $\mu(A_1) < \infty$ ebenfalls $\lim_{i\to\infty} \mu(A_i) = \mu(A)$.

Beweis. Siehe Anhang A. □

Ein Maß, das einem d-dimensionalen Intervall $(a, b] = \prod_{i=1}^{d} (a_i, b_i] \subset \mathbb{R}^d$ sein Volumen zuordnet, ist das Lebesgue-Maß. Der folgende Satz besagt dessen Eindeutigkeit und Existenz als „Volumenmaß“, weshalb die nachfolgenden Betrachtungen weitestgehend auf das Lebesgue-Maß beschränkt werden.

Satz und Definition 3.24. *Für festes $d \in \mathbb{N}$ existiert genau ein Maß*

$$\lambda^d : \mathcal{B}(\mathbb{R}^d) \to [0, \infty],$$

sodass für jedes d-dimensionale Intervall $(a, b] = \prod_{i=1}^{d} (a_i, b_i] \subset \mathbb{R}^d$ gerade

$$\lambda^d((a, b]) \;=\; \prod_{i=1}^{d} (b_i - a_i)$$

*gilt. Das Maß λ^d heißt dann das **d-dimensionale Lebesgue-Maß**.*

Beweis. Siehe [18]. □

Bemerkung 3.25. (i) Für jeden Punkt $x \in \mathbb{R}^d$ ist die Menge $\{x\}$ eine Nullmenge. Da $\prod_{i=1}^{d} (x_i - \frac{1}{k}, x_i] \downarrow \{x\}$ gilt, ist nach Satz 3.23 $\lambda^d(\{x\}) = \lim_{k\to\infty} \left(\frac{1}{k}\right)^d = 0$.

(ii) Für alle $a, b \in \mathbb{R}^d$ mit $a \leq b$ gilt $\lambda^d((a, b)) = \lambda^d([a, b)) = \lambda^d((a, b]) = \lambda^d([a, b]) = \prod_{i=1}^{d} (b_i - a_i)$.

(iii) Für jede abzählbare Teilmenge $A \subset \mathbb{R}^d$ gilt $\mu(A) = 0$. Insbesondere ist auch $\lambda^1(\mathbb{Q}) = 0$.

(iv) Das Lebesgue-Maß ist σ-endlich wegen $\mathbb{R}^d = \bigcup_{N\in\mathbb{N}} \prod_{i=1}^{d} [-N, N]$.

In einem Maßraum $(X, \mathcal{A}, \mu)$ kann nun für $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-messbare Funktionen $f : X \to \mathbb{R}$ ein Integralbegriff eingeführt werden. Das detaillierte Vorgehen findet sich beispielsweise in [7]. Die Idee dabei ist es, f in einen positiven und negativen Anteil zu zerlegen, d.h. $f = f_+ - f_-$ mit $f_+(x) := \max\{f(x), 0\}$ und $f_-(x) := -\min\{0, f(x)\}$ für $x \in \mathbb{R}^d$ und durch eine Folge von Stufenfunktionen anzunähern.
Für eine nicht-negative Stufenfunktion $s := \sum_{j=1}^{m} c_j \chi_{A_j}$ mit $c_j \in [0, \infty)$ und $A_j \in \mathcal{A}$ ist das Integral bzgl. μ über

$$\int s \,\mathrm{d}\mu \;:=\; \sum_{j=1}^{m} c_j \mu(A_j) \in [0, \infty]$$

definiert. Zu $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-messbaren Funktionen f mit $f \geq 0$, also $f(x) \geq 0$ für alle $x \in \mathbb{R}^d$, gibt es eine Folge monoton wachsender Stufenfunktionen $(s_j)_{j\in\mathbb{N}}$, welche punktweise gegen f konvergiert, d.h. $\lim_{k\to\infty} s_j(x) = f(x)$ für $x \in X$. Das Integral von f bzgl. μ ist dann über

$$\int f \,\mathrm{d}\mu \;:=\; \lim_{j\to\infty} \int s_j \,\mathrm{d}\mu \in [0, \infty]$$

festgelegt. Insgesamt ergibt sich dann für das Integral allgemeiner messbarer Funktionen

$$\int f \,\mathrm{d}\mu \;:=\; \int f_+ \,\mathrm{d}\mu - \int f_- \,\mathrm{d}\mu,$$

sofern nicht beide Integrale den Wert ∞ annehmen. Dies bedeutet, dass das Integral nicht immer existieren muss. Eine andere Darstellung des Integrals ist beispielsweise $\int f(x) \,\mathrm{d}\mu(x) := \int f \,\mathrm{d}\mu$. Ist zusätzlich der Integrationsbereich auf $A \in \mathcal{A}$ beschränkt, so wird $\int_A f \,\mathrm{d}\mu$ geschrieben.
Falls $\mu = \lambda^d$ gilt, so wird auch die Schreibweise $\int_A f(x) \,\mathrm{d}x := \int_A f(x) \,\mathrm{d}\lambda^d(x)$ verwendet. Dies ist daher motiviert, dass für beschränkte Funktionen $f : \mathbb{R}^d \to \mathbb{R}$ mit kompaktem Träger aus der Riemann-Integrierbarkeit bereits die Lebesgue-Integrierbarkeit folgt, siehe [7].

Beispiel 3.26. Die einfachste Stufenfunktion ist die charakteristische Funktion χ_A, $A \in \mathcal{A}$, welche bereits in der Definition der Stufenfunktion selbst verwendet und in Beispiel 2.9 eingeführt wurde. Im Maßraum $(X, \mathcal{A}, \mu)$ gilt für ihr Integral

$$\int \chi_A \, \mathrm{d}\mu \quad = \quad \mu(A)$$

und für $(X, \mathcal{A}, \mu) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda^d)$ gilt

$$\int \chi_A(x) \, \mathrm{d}x \quad = \quad \lambda^d(A).$$

Analog zu den Produkträumen und der Produkt-σ-Algebra ist es nun das Ziel, aus den Maßräumen $(X^k, \mathcal{A}^k, \mu^k)$, $k = 1, \ldots, n$, einen Produktmaßraum zu bilden. Zunächst wird dazu festgelegt, welche Eigenschaft ein Produktmaß erfüllen muss.

Definition 3.27. *Seien $(X^k, \mathcal{A}^k, \mu^k)$ für $k = 1, \ldots, n$ Maßräume. Ein Maß*

$$\bigotimes_{k=1}^{n} \mu^k : \bigotimes_{k=1}^{n} \mathcal{A}^k \to [0, \infty]$$

*wird **Produktmaß** der μ^k genannt, falls für alle $A^k \in \mathcal{A}^k$, $k = 1, \ldots, n$*

$$\bigotimes_{k=1}^{n} \mu^k(A^1 \times \ldots \times A^n) \quad = \quad \prod_{k=1}^{n} \mu^k(A^k)$$

gilt, wobei $0 \cdot \infty := 0$ verwendet wird.

Das folgende Resultat ist zunächst für zwei Maßräume formuliert.

Satz 3.28. *Seien $(X^k, \mathcal{A}^k, \mu^k)$, $k = 1, 2$, σ-endliche Maßräume. Dann existiert ein eindeutiges Produktmaß $\bigotimes_{k=1}^{2} \mu^k$ auf $\bigotimes_{k=1}^{2} \mathcal{A}^k$. Für alle $A \in \bigotimes_{k=1}^{2} \mathcal{A}^k$ gilt außerdem*

$$\begin{aligned} \bigotimes_{k=1}^{2} \mu^k(A) &= \int \left(\int \chi_A(x^1, x^2) \, \mathrm{d}\mu^1(x^1) \right) \mathrm{d}\mu^2(x^2) \\ &= \int \left(\int \chi_A(x^1, x^2) \, \mathrm{d}\mu^2(x^2) \right) \mathrm{d}\mu^1(x^1). \end{aligned}$$

Beweis. Siehe [8]. □

Dieser Satz kann sofort auf n Maßräume erweitert werden, vgl. ebenfalls [8]. Speziell

für das Lebesgue-Maß folgt daraus

$$\lambda^d = \bigotimes_{k=1}^{n} \lambda^{d^k} \tag{3.4}$$

für feste $d^k \in \mathbb{N}$, $k = 1, \ldots, n$ und $d := \sum_{k=1}^{n} d^k$.

4. Konvexe Theorie

In diesem Kapitel werden zunächst konvexe Mengen und Funktionen untersucht. Anschließend werden Optimierungsprobleme betrachtet, denen konvexe Funktionen zugrunde liegen.

4.1. Konvexe Mengen und Funktionen

Die nachfolgenden Betrachtungen zur Konvexität sind auf den $\mathbb{R}^d$ beschränkt.

Konvexe Mengen

Definition 4.1. *Eine Menge $C \subset \mathbb{R}^d$ wird **konvex** genannt, wenn die Verbindungsstrecke zweier beliebiger Punkte in C wieder in C liegt. D.h. also, falls für $x, y \in C$ und $\theta \in [0,1]$ gerade*

$$\theta x + (1-\theta)y \in C$$

gilt.

Beispiel 4.2. Ein Beispiel für eine konvexe Menge ist die Lösungsmenge eines linearen Gleichungssystems $C = \{x \in \mathbb{R}^d \mid Ax = b\}$, wobei $A \in \mathbb{R}^{m \times d}$ und $b \in \mathbb{R}^m$. Denn sind $x, y \in C$, d.h. $Ax = b$ und $Ay = b$, so gilt für beliebiges $\theta \in [0,1]$ nun

$$\begin{aligned} A(\theta x + (1-\theta)y) &= \theta Ax + (1-\theta)Ay \\ &= \theta b + (1-\theta)b \\ &= b. \end{aligned}$$

Deshalb ist auch $\theta x + (1-\theta)y \in C$ und C konvex.

Auch die leere Menge $\emptyset$, die Menge, welche einen einzelnen Punkt $x \in \mathbb{R}^d$ enthält und der gesamte $\mathbb{R}^d$ sind Beispiele konvexer Teilmengen des $\mathbb{R}^d$.

Der folgende Satz stellt zwei wichtige Eigenschaften von konvexen Mengen vor.

Satz 4.3. *(i) Seien $C_j \subset \mathbb{R}^d$, $j = 1, \ldots, m$, $m \in \mathbb{N}$ konvexe Mengen, dann ist auch $C := \bigcap_{j=1}^{m} C_j$ eine konvexe Menge.*

(ii) Seien $n \in \mathbb{N}$ und $d^k \in \mathbb{N}$ fest für $k = 1, \dots, n$ mit $d := \sum_{k=1}^{n} d^k$. Ferner seien $C^k \subset \mathbb{R}^{d^k}$ nichtleer. Die Menge $C := \prod_{k=1}^{n} C^k \subset \mathbb{R}^d$ ist genau dann konvex, wenn die Mengen C^k, $k = 1, \dots, n$, konvex sind.

Beweis. Siehe Anhang A. □

Mit diesem Satz kann gefolgert werden, dass d-dimensionale Rechtecke im $\mathbb{R}^d$, bekannt aus Beispiel 2.19, konvex sind, da jedes Intervall in den reellen Zahlen offensichtlich konvex ist.

Konvexe Funktionen

Neben konvexen Mengen können auch konvexe Funktionen betrachtet und deren Eigenschaften untersucht werden.

Definition 4.4. *(i) Sei $C \subset \mathbb{R}^d$ konvex. Eine Funktion $f : C \to \mathbb{R}$ heißt **konvex**, falls für alle $x, y \in C$ mit $x \neq y$ und für alle $\theta \in (0, 1)$ gilt, dass*

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta) f(y). \tag{4.1}$$

*(ii) Gilt in der Ungleichung (4.1) sogar „<“, so wird f **streng konvex** genannt.*

*(iii) Falls $-f$ konvex ist, so heißt f **konkav**.*

*(iv) Analog heißt f **strikt konkav**, falls $-f$ strikt konvex ist.*

Die Ungleichung (4.1) kann gleichermaßen für $\theta \in [0, 1]$ formuliert werden. Da diese für $\theta = 0$ bzw. $\theta = 1$ jedoch stets erfüllt ist, werden die beiden Fälle in der Literatur oftmals weggelassen. Geometrisch bedeutet die Ungleichung (4.1), dass die Verbindungsstrecke zwischen $(x, f(x))$ und $(y, f(y))$ nie unterhalb des Graphen von f liegt.

Anknüpfend an Definition 4.4 wird eine Funktion $f = (f_1, \dots, f_m)^{\mathrm{T}} : C \to \mathbb{R}^m$ genau dann als konvex bezeichnet, wenn jede ihrer Komponenten f_j, $j = 1, \dots, m$, eine konvexe Funktion darstellt. Für derartige Funktionen kann die Ungleichung (4.1) als Ungleichungssystem auch komponentenweise verstanden werden. Entsprechendes gilt für strikt konvex, konkav und strikt konkav.
Die nachfolgenden Betrachtungen werden zunächst auf Funktionen, die nach $\mathbb{R}$ abbilden, beschränkt. Jedoch können diese in der Regel sofort auf Funktionen mit Wertebereich $\mathbb{R}^m$ erweitert werden.

Wichtige Eigenschaften von konvexen Funktionen sind, dass jede Minimalstelle global ist und dass konvexe Funktionen für offenes C stetig sind, vgl. [2]. Für f strikt konvex ist eine globale Minimalstelle, sofern sie existiert, sogar eindeutig.
Falls konvexe Funktionen zusätzlich differenzierbar sind, können Bedingungen erster und zweiter Ordnung hergeleitet werden.

Bedingungen erster und zweiter Ordnung

Satz 4.5. *Seien $C \subset \mathbb{R}^d$ offen und konvex und $f : C \to \mathbb{R}$ differenzierbar. So gilt, dass f genau dann konvex ist, wenn für alle $x, y \in C$ mit $x \neq y$ die Ungleichung*

$$f(x) + \nabla f(x)^{\mathrm{T}}(y - x) \quad \leq \quad f(y) \tag{4.2}$$

erfüllt ist.

Beweis. Siehe Anhang A. □

Ist in Satz 4.5 die Funktion f strikt konvex, so kann gezeigt werden, dass für Ungleichung (4.2) sogar „<“ gilt.

An dieser Stelle soll kurz an die Taylor-Formel am Beispiel einer zweimal differenzierbaren Funktion $f : C \to \mathbb{R}$ mit offenem und konvexen $C \subset \mathbb{R}^d$ erinnert werden. In [7] findet sich eine allgemeine Einführung in die Thematik. Seien $x \in C$, $h \in \mathbb{R}^d$ mit $x + h \in C$, so gibt es nach dem Satz von Taylor ein $\theta \in (0, 1)$, sodass

$$f(x + h) \quad = \quad f(x) + \nabla f(x)^{\mathrm{T}} h + \frac{1}{2} h^{\mathrm{T}} \nabla^2 f(x + \theta h) h \tag{4.3}$$

gilt. Die ersten beiden Summanden stellen hier das Taylorpolynom von Ordnung 1 dar.
Die Ungleichung (4.2) besagt deshalb, dass die Funktionswerte, welche das Taylorpolynom erster Ordnung von f um einen beliebigen Punkt $x \in C$ annimmt, stets kleiner als die von f am selben Punkt sind. Aus lokalen Informationen von f, d.h. dem Funktionswert und dem Gradienten an einer bestimmten Stelle $x \in C$, ergeben sich also globale Informationen. Insbesondere gilt $f(y) \geq f(x)$ für alle $y \in C$, falls $\nabla f(x)$ der Nullvektor ist. Dies verdeutlicht nochmals, dass jede lokale Minimalstelle auch eine globale ist.

Satz 4.6. *Seien $C \subset \mathbb{R}^d$ offen und konvex und $f : \mathbb{R}^d \to \mathbb{R}$ zweimal differenzierbar. So ist f genau dann konvex, wenn $\nabla^2 f(x)$ für alle $x \in C$ positiv semidefinit ist.*

Beweis. Siehe Anhang A. □

Analog zum Beweis von Satz 4.6 kann darauf geschlossen werden, dass f strikt konvex ist, falls $\nabla^2 f(x)$ für alle $x \in C \subset \mathbb{R}^d$ positiv definit ist. Die Rückrichtung gilt jedoch nicht, wie das Beispiel $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^4$ zeigt.
Die Bedingungen erster und zweiter Ordnung lassen sich direkt auf konkave Funktionen übertragen. Stellvertretend dafür sei die Bedingung erwähnt, dass $f : \mathbb{R}^d \to \mathbb{R}$ genau dann konkav ist, wenn $\nabla^2 f(x)$ negativ semidefinit für alle $x \in C$ ist. Dabei muss wie oben beschrieben stets $C \subset \mathbb{R}^d$ offen und konvex sein.

Neben der Definition von konvexen Funktionen ist auch die Bedingung zweiter Ordnung aus Satz 4.6 zum Nachweis der Konvexität von Funktionen hilfreich.

Beispiel 4.7. (i) Jede lineare Funktion $f : \mathbb{R}^d \to \mathbb{R}^m$ ist konvex. Aus Beispiel 4.2 kann direkt gefolgert werden, dass für lineare Funktionen in Ungleichung (4.1) Gleichheit gilt.

(ii) Die quadratische Funktion $f : \mathbb{R}^d \to \mathbb{R}$, gegeben über

$$f(x) \;=\; \frac{1}{2} x^{\mathrm{T}} P x + q^{\mathrm{T}} + r$$

für alle $x \in \mathbb{R}^d$ mit $P \in \mathbb{R}^{d \times d}$, $q \in \mathbb{R}^d$ und $r \in \mathbb{R}$, ist konvex, falls P positiv semidefinit ist, da $\nabla^2 f(x) = P$ gilt.

(iii) Jede Norm $\|\cdot\|$ auf $\mathbb{R}^d$ ist konvex. Denn für $\theta \in (0,1)$ gilt

$$\|\theta x + (1-\theta) y\| \;\leq\; \|\theta x\| + \|(1-\theta) y\| \;\leq\; \theta \|x\| + (1-\theta)\|y\|.$$

Weitere Eigenschaften von konvexen Funktionen

Im Folgenden werden weitere Eigenschaften von konvexen Funktionen untersucht. Diese werden hier auf den Definitionsbereich $C = \mathbb{R}^d$ reduziert.

Satz 4.8. *Die α-Subniveaumenge $C_\alpha := \{x \in \mathbb{R}^d \mid f(x) \leq \alpha\}$, bekannt aus Definition 3.16, einer konvexen Funktion $f : \mathbb{R}^d \to \mathbb{R}$ ist eine Borel-Menge und konvex.*

Beweis. Da $f : \mathbb{R}^d \to \mathbb{R}$ konvex ist, ist f auch stetig und damit Borel-messbar. Somit ist C_α nach Korollar 3.15 eine Borel-Menge.
Seien nun $x, y \in C_\alpha$, dann gelten $f(x) \leq \alpha$ und $f(y) \leq \alpha$. Daher ist für $\theta \in (0,1)$ stets $f(\theta x + (1-\theta) y) \leq \alpha$ erfüllt, woraus $\theta x + (1-\theta) y \in C_\alpha$ folgt, d.h. C_α ist konvex. □

Umgekehrt werden Funktionen $f : \mathbb{R}^d \to \mathbb{R}$, deren α-Subniveaumengen für alle $\alpha \in \mathbb{R}$ konvex sind, als **quasikonvex** bezeichnet. Offensichtlich ist jede konvexe Funktion somit quasikonvex, aber nicht jede quasikonvexe Funktion ist konvex, vgl. [4].

Satz 4.9. *Seien $f_j : \mathbb{R}^d \to \mathbb{R}$, $j = 1, \ldots, m$, konvexe Funktionen. Dann gilt*

(i) die Funktion $f := \sum_{j=1}^{m} f_j$ ist konvex,

(ii) das punktweise Maximum $\max_{j\in\{1,\ldots,m\}} f_j$ ist konvex.

Beweis. Für $f_j : \mathbb{R}^d \to \mathbb{R}$, $j = 1, \ldots, m$, konvex, $x, y \in \mathbb{R}^d$ und $\theta \in (0, 1)$ gilt

$$\begin{aligned} f(\theta x + (1-\theta)y) &= \sum_{j=1}^{m} f_j(\theta x + (1-\theta)y) \\ &\leq \sum_{j=1}^{m} (\theta f_j(x) + (1-\theta) f_j(y)) \\ &= \theta f(x) + (1-\theta) f(y). \end{aligned}$$

und

$$\begin{aligned} \max_{j\in\{1,\ldots,m\}} f_j(\theta x + (1-\theta)y) &\leq \max_{j\in\{1,\ldots,m\}} (\theta f_j(x) + (1-\theta) f(y)) \\ &\leq \theta \max_{j\in\{1,\ldots,m\}} f_j(x) + (1-\theta) \max_{j\in\{1,\ldots,m\}} f_j(y). \end{aligned}$$

Also sind f und $\max_{j\in\{1,\ldots,m\}} f_j$ konvex. □

Satz 4.10. *Seien $n \in \mathbb{N}$ und $d^k \in \mathbb{N}$ für $k = 1, \ldots, n$ fest mit $d := \sum_{k=1}^{n} d^k$. Falls eine Funktion $f : \mathbb{R}^d \to \mathbb{R}^m$ nach Definition 2.17 additiv zerlegbar ist mit $f(x) = \sum_{k=1}^{n} f^k(x^k)$ für $x \in \mathbb{R}^d$ und $f^k : \mathbb{R}^{d^k} \to \mathbb{R}^m$, so ist f genau dann konvex, wenn die Funktionen f^k für $k = 1, \ldots, n$ konvex sind.*

Beweis. Da eine Funktion, die nach $\mathbb{R}^m$ abbildet, genau dann konvex ist, wenn jede ihrer m Komponenten konvex ist, genügt es die Äquivalenz für die j-te Komponente zu zeigen mit $j \in \{1, \ldots, m\}$.
„⇐“: Seien f_j^k, $k = 1, \ldots, n$, konvex, so ist wegen Satz 4.9 mit $f_j^k \circ \pi^k(x) = f_j^k(x^k)$ für $x \in \mathbb{R}^d$ auch f_j konvex.

„⇒“: Seien f_j konvex und $k \in \{1, \dots, n\}$ beliebig. Mit der kanonischen Einbettung ι^k für feste $\tilde{x}^{\tilde{k}} \in \mathbb{R}^{d^{\tilde{k}}}$ mit $\tilde{k} = 1, \dots, n$, $\tilde{k} \neq k$, kann jedem $x^k \in \mathbb{R}^{d^k}$ ein $x \in \mathbb{R}^d$ zugeordnet werden. Für $x^k, y^k \in \mathbb{R}^{d^k}$ und $\theta \in (0,1)$ ist dann

$$\begin{aligned} f_j^k(\theta x^k + (1-\theta)y^k) &= f_j(\theta \iota^k(x^k) + (1-\theta)\iota^k(y^k)) - \sum_{\substack{\tilde{k}=1 \\ \tilde{k}\neq k}}^{n} f_j^{\tilde{k}}(\tilde{x}^{\tilde{k}}) \\ &\leq \theta f_j(\iota^k(x^k)) + (1-\theta) f_j(\iota^k(y^k)) - \sum_{\substack{\tilde{k}=1 \\ \tilde{k}\neq k}}^{n} f_j^{\tilde{k}}(\tilde{x}^{\tilde{k}}) \\ &= \theta f_j^k(x^k) + (1-\theta) f_j^k(y^k). \end{aligned}$$

□

Bemerkung 4.11. Für quasikonvexe Funktionen gilt der Satz 4.10 nicht. Beispielsweise ist für die quasikonvexen Funktionen $f^k : \mathbb{R} \to \mathbb{R}, x_k \mapsto (x_k)^3$, $k = 1, 2$, die (-1)-Subniveaumenge der Funktion f, definiert über $f(x_1, x_2) := f^1(x_1) + f^2(x_2)$ für alle $x_1, x_2 \in \mathbb{R}$, wegen $f(-1, 0) = f(0, -1) = -1$ und $f(-0{,}5, -0{,}5) > -1$ nicht konvex. Also ist f nicht quasikonvex.

Fortsetzung konvexer Funktionen

Für $C \subsetneq \mathbb{R}^d$ ist es möglich den Definitionsbereich von konvexen Funktionen zu erweitern.

Definition 4.12. *Sei $C \subset \mathbb{R}^d$ konvex. Falls $f : C \to \mathbb{R}$ eine konvexe Funktion ist, so heißt*

$$\tilde{f} : \mathbb{R}^d \to (-\infty, \infty],\ x \mapsto \begin{cases} f(x) & \text{falls } x \in C, \\ \infty & \text{sonst} \end{cases}$$

*die **werteerweiterte Fortsetzung von** f.*

Der Definitionsbereich der ursprünglichen Funktion f kann hierbei über die Menge $C = \left\{x \in \mathbb{R}^d \mid \tilde{f}(x) < \infty\right\}$ beschrieben werden.
Für alle $x, y \in \mathbb{R}^d$ und alle $\theta \in (0,1)$ gilt für die werteerweiterte Fortsetzung einer konvexen Funktion

$$\tilde{f}(\theta x + (1-\theta)y) \leq \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y). \tag{4.4}$$

Denn für $x, y \in C$ stimmt die Ungleichung (4.4) mit Ungleichung (4.1) überein. Falls $x \notin C$ oder $y \notin C$ nimmt die rechte Seite von (4.4) den Wert ∞ an und die Ungleichung ist immer erfüllt.

In diesem Sinne kann die Definition 4.4 auf Funktionen $\tilde{f} : \mathbb{R}^d \to (-\infty, \infty]$ erweitert werden. So wird $\tilde{f}$ konvex genannt, falls die Ungleichung (4.4) erfüllt ist. Dies impliziert nämlich, dass $\widetilde{C} := \{x \in \mathbb{R}^d \mid \tilde{f}(x) < \infty\}$ konvex ist und für alle $x, y \in \widetilde{C}$ und alle $\theta \in (0, 1)$ die Ungleichung (4.1) gilt. Insgesamt können solche Funktionen $\tilde{f}$ also immer als eine werteerweiterte Fortsetzung einer konvexen Funktion aufgefasst werden. Eine wichtige Eigenschaft, welche die werteerweiterte Fortsetzung von der ihr zugrunde liegenden konvexen Funktion übernimmt, ist die Globalität jeder Minimalstelle.

Gleichermaßen können konkave Funktionen fortgesetzt werden, indem Variablen, die nicht in C liegen, der Wert $-\infty$ zugeordnet wird. Hierbei ist für alle $x, y \in \mathbb{R}^d$ und alle $\theta \in (0, 1)$ die Ungleichung

$$\tilde{f}(\theta x + (1-\theta)y) \quad \geq \quad \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y), \tag{4.5}$$

erfüllt.

Logarithmisch-konkave und logarithmisch-konvexe Funktionen

Für die konvexe Optimierung sind logarithmisch-konvexe bzw. logarithmisch-konkave Funktionen oftmals hilfreich. Dazu sei bemerkt, dass während der gesamten Arbeit unter „Logarithmus“ stets der natürliche Logarithmus zu verstehen ist.

Definition 4.13. *Eine Funktion $f : \mathbb{R}^d \to (0, \infty)$, auch notiert über $f > 0$, wird **logarithmisch-konkav** genannt, falls $\log f$ konkav ist. Entsprechend wird f **logarithmisch-konvex** genannt, falls $\log f$ konvex ist.*

Durch Anwenden der Logarithmen-Gesetze, vgl. [7], ergibt sich, dass $f > 0$ genau dann logarithmisch-konvex ist, wenn $1/f$ logarithmisch konkav ist.

Die logarithmische Konkavität kann auch auf Funktionen $f : \mathbb{R}^d \to [0, \infty)$, d.h. $f \geq 0$, erweitert werden. Im Sinne der werteerweiterten Fortsetzung sei

$$\tilde{f} : \mathbb{R}^d \to [-\infty, \infty), \; x \mapsto \begin{cases} \log f(x) & \text{falls } f(x) > 0, \\ -\infty & \text{falls } f(x) = 0. \end{cases}$$

Falls $\tilde{f}$ die Bedingung (4.5) erfüllt, also $\tilde{f}$ konkav ist, so wird f logarithmisch-konkav

genannt.

Die logarithmische Konkavität kann auch ohne Logarithmus ausgedrückt werden.

Satz 4.14. *Eine Funktion $f : \mathbb{R}^d \to [0, \infty)$ ist genau dann logarithmisch-konkav, wenn für alle $x, y \in \mathbb{R}^d$ und für alle $\theta \in (0, 1)$ gerade*

$$f(\theta x + (1-\theta)y) \geq f(x)^\theta f(y)^{1-\theta} \tag{4.6}$$

gilt.

Beweis. Sei $f > 0$. Dann sind für $x, y \in \mathbb{R}^d$ und $\theta \in (0, 1)$ die Aussagen

$$\begin{aligned}
\log f(\theta x + (1-\theta)y) &\geq \theta \log f(x) + (1-\theta) \log f(y), \\
\log f(\theta x + (1-\theta)y) &\geq \log\left(f(x)^\theta\right) + \log\left(f(y)^{1-\theta}\right), \\
\log f(\theta x + (1-\theta)y) &\geq \log\left(f(x)^\theta f(y)^{1-\theta}\right), \\
f(\theta x + (1-\theta)y) &\geq f(x)^\theta f(y)^{(1-\theta)}
\end{aligned}$$

aufgrund der Logarithmen-Gesetze und der Monotonie der Logarithmusfunktion äquivalent. O.B.d.A. gelte nun $f(x) = 0$ für $x \in \mathbb{R}^d$. Dann ist sowohl die Ungleichung (4.5) wie oben beschrieben als auch die Ungleichung (4.6) wegen $f \geq 0$ stets erfüllt. □

Beispiel 4.15. Die charakteristische Funktion von $A \subset \mathbb{R}^d$ ist logarithmisch-konkav, falls A konvex ist. Denn für $x \notin A$ oder $y \notin A$ ist die Ungleichung (4.6) stets erfüllt und für $x, y \in A$ gilt $\theta x + (1-\theta)y \in A$, weshalb (4.6) ebenfalls erfüllt ist.

Nun wird auf eine wichtige Eigenschaft von logarithmisch-konvexen Funktionen eingegangen. Dazu wird die Prékopa-Leindler Ungleichung benötigt.

Satz 4.16. *Seien $f, g, h : \mathbb{R}^d \to [0, \infty)$ Borel-messbare Funktionen, welche die Bedingung*

$$h(\theta x + (1-\theta)y) \geq f(x)^\theta g(y)^{1-\theta} \tag{4.7}$$

für alle $x, y \in \mathbb{R}^d$ und alle $\theta \in (0, 1)$ erfüllen. Dann gilt

$$\int_{\mathbb{R}^d} h(x)\,\mathrm{d}x \geq \left(\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x\right)^\theta \left(\int_{\mathbb{R}^d} g(x)\,\mathrm{d}x\right)^{1-\theta}. \tag{4.8}$$

Diese Ungleichung wird auch ***Prékopa-Leindler Ungleichung*** *genannt.*

Beweis. Siehe [20] und [21]. □

Korollar 4.17. *Seien $A \in \mathcal{B}(\mathbb{R}^d)$ konvex, $f : \mathbb{R}^d \times \mathbb{R}^m \to [0,\infty)$ logarithmisch-konkav und $f(\cdot, y)$ Borel-messbar für alle $y \in \mathbb{R}^m$. Dann ist die Funktion*

$$g : \mathbb{R}^m \to [0,\infty],\ y \mapsto g(y) := \int_A f(x,y)\, dx \tag{4.9}$$

ebenfalls eine logarithmisch-konkave Funktion.

Beweis. Seien $y', y'' \in \mathbb{R}^m$ und $\theta \in (0,1)$. Ferner seien

$$f_1 : \mathbb{R}^d \to [0,\infty),\ x \mapsto \begin{cases} f(x,y') & \textit{falls } x \in A, \\ 0 & \text{sonst}, \end{cases}$$

$$f_2 : \mathbb{R}^d \to [0,\infty),\ x \mapsto \begin{cases} f(x,y'') & \textit{falls } y \in A, \\ 0 & \text{sonst}, \end{cases}$$

$$f_3 : \mathbb{R}^d \to [0,\infty),\ x \mapsto \begin{cases} f(x,\theta y' + (1-\theta)y'') & \textit{falls } x \in A, \\ 0 & \text{sonst}. \end{cases}$$

Da A konvex ist und f logarithmisch-konkav, gilt

$$f_3(\theta x' + (1-\theta)x'') \geq f_1(x')^\theta f_2(x'')^{1-\theta}$$

für alle $x', x'' \in \mathbb{R}^d$. Mit Satz 4.16 folgt schließlich

$$\begin{aligned} g(\theta y' + (1-\theta)y'') &= \int_A f(x,\theta y' + (1-\theta)y'')\,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} f_3(x)\,\mathrm{d}x \\ &\geq \left(\int_{\mathbb{R}^d} f_1(x)\,\mathrm{d}x\right)^\theta \left(\int_{\mathbb{R}^d} f_2(x)\,\mathrm{d}x\right)^{1-\theta} \\ &= \left(\int_A f(x,y')\,\mathrm{d}x\right)^\theta \left(\int_A f(x,y'')\,\mathrm{d}x\right)^{1-\theta} \\ &= g(y')^\theta g(y'')^{1-\theta}, \end{aligned}$$

für alle $y', y'' \in \mathbb{R}^m$ und alle $\theta \in (0,1)$. D.h. g ist logarithmisch-konkav. □

4.2. Konvexe Optimierungsprobleme

Bevor genauer auf konvexe Optimierungsprobleme eingegangen wird, werden zunächst noch Grundlagen zu allgemeinen Optimierungsproblemen besprochen.

Grundlegendes

Sei $D \subset \mathbb{R}^d$. Bei einem Optimierungsproblem ist eine Minimal- bzw. Maximalstelle einer **Zielfunktion** $J : D \to \mathbb{R}$ auf einer **zulässigen Menge** $\mathcal{F} \subset D$ gesucht. Da das Maximieren von J für $x \in \mathcal{F}$ dem Minimieren von $-J$ für $x \in \mathcal{F}$ entspricht, wird in diesem Kapitel stellvertretend nur das Minimierungsproblem betrachtet. In der Literatur wird dieses generell über

$$\text{minimiere } J(x) \text{ u.d.N. } x \in \mathcal{F} \quad \text{bzw.} \quad \underset{x \in \mathcal{F}}{\text{minimiere}}\ J(x)$$

notiert, wobei die Abkürzung „u.d.N." für „unter der Nebenbedingung" steht. Falls $\mathcal{F} = D = \mathbb{R}^d$ gilt, so handelt es sich um eine unrestringierte Optimierungsaufgabe. Häufig kann die zulässige Menge über ein System von Gleichungen und Ungleichungen definiert werden, was in dieser Arbeit stets der Fall sein soll. Die zulässige Menge $\mathcal{F}$ kann dann über

$$\mathcal{F} \quad = \quad \{x \in D \mid e(x) = e_\mathrm{c},\ f(x) \leq f_\mathrm{c}\}$$

beschrieben werden, wobei $e = (e_1, \ldots, e_p)^\mathrm{T} : D \to \mathbb{R}^p$, $e_\mathrm{c} \in \mathbb{R}^p$ mit $p \in \mathbb{N}$, $p \leq d$ und $f = (f_1, \ldots, f_m)^\mathrm{T} : D \to \mathbb{R}^m$, $f_\mathrm{c} \in \mathbb{R}^m$ mit $m \in \mathbb{N}$ gelten. Das Gleichungssystem $e(x) = e_\mathrm{c}$ sowie das Ungleichungssystem $f(x) \leq f_\mathrm{c}$ müssen dabei für alle $x \in \mathcal{F}$ komponentenweise erfüllt sein. Da auch $D \subset \mathbb{R}^d$ über ein System von Gleichungen und Ungleichungen festgelegt werden kann, ist es ohne Weiteres möglich $D = \mathbb{R}^d$ zu wählen und die Nebenbedingungen entsprechend zu ergänzen. Diese Eigenschaft wird später noch verwendet. An dieser Stelle sei außerdem bemerkt, dass die zulässige Menge $\mathcal{F}$ gleichermaßen auch nur mit einer bzw. ohne eine der Bedingungen $e(x) = e_\mathrm{c}$ und $f(x) \leq f_\mathrm{c}$ formuliert werden kann.
Im Folgenden wird ein Minimierungsproblem daher immer in der Form

$$\begin{aligned} &\underset{x \in D}{\text{minimiere}} && J(x) \\ &\text{u.d.N.} && e(x) = e_\mathrm{c},\ f(x) \leq f_\mathrm{c} \end{aligned} \tag{4.10}$$

geschrieben. In der nun folgenden Definition wird der Begriff einer Lösung eines Minimierungsproblems eingeführt.

Definition 4.18. *(i) Ein Punkt $x^* \in \mathcal{F}$ heißt **lokale Lösung** des Minimierungsproblems (4.10), falls ein $\delta \in (0, \infty)$ existiert mit $J(x^*) \leq J(x)$ für alle $x \in \mathcal{F} \cap B(x^*, \delta)$.*

*(ii) Ein Punkt $x^* \in \mathcal{F}$ heißt **globale** bzw. **optimale Lösung** des Minimierungsproblems (4.10), falls $J(x^*) \leq J(x)$ für alle $x \in \mathcal{F}$.*

Eine globale Lösung existiert beispielsweise unter der Voraussetzung, dass $\mathcal{F}$ kompakt und J stetig bzw. nach unten halbstetig ist, vgl. Kapitel 2.1. Ferner ist eine globale Lösung eindeutig, falls in Definition 4.18 sogar $J(x^*) < J(x)$ für alle $x \in \mathcal{F}$ gilt. Handelt es sich bei $\mathcal{F}$ um eine konvexe Menge und bei J um eine konvexe Funktion, so sind alle lokalen Lösungen von (4.10) auch globale Lösungen, vgl. Kapitel 4.1.
Die zulässige Menge ist für eine konvexe Menge D, eine lineare Funktion e und eine konvexe Funktion f mit Beispiel 4.2, Satz 4.3 und 4.8 konvex. Dies motiviert die folgende Definition.

Definition 4.19. *Seien die Menge $D \subset \mathbb{R}^d$ konvex, $e : D \to \mathbb{R}^p$ eine lineare Funktion und $f : D \to \mathbb{R}^m$, $J : D \to \mathbb{R}$ konvexe Funktionen. Dann wird das Problem (4.10) **konvexes Optimierungsproblem** bzw. konvexes Minimierungsproblem genannt. Mit $e(x) = Ex$, $E \in \mathbb{R}^{p \times d}$ für $x \in D$ kann das konvexe Optimierungsproblem in der Form*

$$\begin{aligned} &\underset{x \in D}{\text{minimiere}} && J(x) \\ &\text{u.d.N.} && Ex = e_{\mathrm{c}},\ f(x) \leq f_{\mathrm{c}} \end{aligned} \tag{4.11}$$

dargestellt werden.

Der Vorteil eines konvexen Optimierungsproblems gegenüber dem allgemeinen ist es, nach lokalen Lösungen suchen zu können, um eine optimale Lösung zu erhalten.

Beispiel 4.20. Seien $f = (f_1, \ldots, f_m)^{\mathrm{T}} : \mathbb{R}^d \to \mathbb{R}^m$ eine konvexe Funktion und $x^{\mathrm{l}}, x^{\mathrm{u}} \in \mathbb{R}^d$ mit $x^{\mathrm{l}} < x^{\mathrm{u}}$. Das Problem

$$\begin{aligned} &\underset{x \in \mathbb{R}^d}{\text{minimiere}} && \|f(x)\|_\infty \\ &\text{u.d.N.} && -x \leq x^{\mathrm{l}},\ x \leq x^{\mathrm{u}} \end{aligned} \tag{4.12}$$

ist ein konvexes Optimierungsproblem. Unter Verwendung von Satz 4.9 bzw. Beispiel 4.7 ist dies sofort ersichtlich.
Selbst wenn die Funktionen f_j, $j = 1, \ldots, m$, differenzierbar sind, ist in der Regel nicht gegeben, dass die Zielfunktion des Beispielproblems (4.12) ebenso differenzierbar ist. Hier ist es allerdings möglich das Problem (4.12) auf ein äquivalentes, differenzierbares Problem zu transformieren, welches

$$\begin{array}{ll} \underset{(\alpha,x)\in\mathbb{R}\times\mathbb{R}^d}{\text{minimiere}} & \alpha \\ \text{u.d.N.} & -x \leq x^{\mathrm{l}},\ x \leq x^{\mathrm{u}},\ f_j(x) \leq \alpha,\ j = 1, \ldots, m \end{array}$$

lautet, vgl. [12].

Optimalitätskriterien

Für die Bestimmung einer Lösung eines Optimierungsproblems sind Optimalitätskriterien von großer Bedeutung. Dazu wird zunächst wieder das allgemeine Optimierungsproblem (4.10) betrachtet. Im Folgenden sei $D = \mathbb{R}^d$. Falls J, f und g stetig differenzierbar sind, können solche Optimalitätskriterien mithilfe der Lagrange-Funktion hergeleitet werden.

Definition 4.21. *Die Funktion $L : \mathbb{R}^d \times \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}$, definiert über*

$$L(x, \lambda, \mu) \quad := \quad J(x) + \sum_{i=1}^{p} \lambda_i (e_i(x) - e_{\mathrm{c},i}) + \sum_{j=1}^{m} \mu_j (f_j(x) - f_{\mathrm{c},j})$$

*für $(x, \lambda, \mu) \in \mathbb{R}^d \times \mathbb{R}^p \times \mathbb{R}^m$, wird **Lagrange-Funktion** des Optimierungsproblems (4.10) genannt, wobei $\lambda_i, i = 1, \ldots, p$ die Komponenten des Vektors $\lambda \in \mathbb{R}^p$ und $\mu_i, i = 1, \ldots, m$, die Komponenten des Vektors $\mu \in \mathbb{R}^m$ sind.*

Definition 4.22. *Gegeben sei das Optimierungsproblem (4.10) mit stetig differenzierbaren Funktionen J, e und f.*

(i) Die Bedingungen

$$\begin{aligned} \nabla_x L(x, \lambda, \mu) &= 0, && (4.13) \\ e(x) &= e_{\mathrm{c}}, && (4.14) \\ f(x) &\leq f_{\mathrm{c}}, && (4.15) \\ \mu &\geq 0, && (4.16) \\ \mu^{\mathrm{T}}(f(x) - f_{\mathrm{c}}) &= 0 && (4.17) \end{aligned}$$

heißen ***KKT-Bedingungen*** *des Optimierungsproblems (4.10), wobei*

$$\nabla_x L(x, \lambda, \mu) \quad = \quad \nabla f(x) + \sum_{i=1}^{p} \lambda_i \nabla e_i(x) + \sum_{j=1}^{m} \mu_i \nabla f_j(x)$$

der Gradient der Lagrange-Funktion L bezüglich der Variablen x ist.

(ii) Ein Vektor $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^d \times \mathbb{R}^p \times \mathbb{R}^m$, welcher die KKT-Bedingungen erfüllt, heißt ***KKT-Punkt*** *des Optimierungsproblems (4.10). Dabei werden die Komponenten von $\lambda^* \in \mathbb{R}^p$ und $\mu^* \in \mathbb{R}^m$ auch* ***Lagrange-Multiplikatoren*** *genannt.*

Für den Fall, dass es sich um eine unrestringierte Optimierungsaufgabe handelt, so reduzieren sich die KKT-Bedingungen auf die Bedingung $\nabla f(x) = 0$, also gerade die notwendige Optimalitätsbedingung erster Ordnung, vgl. [10]. Unter gewissen Regularitätsbedingungen an die zulässige Menge, d.h. an Funktionen f bzw. g, kann gezeigt werden, dass die KKT-Bedingungen auch eine notwendige Optimalitätsbedingung erster Ordnung an das Problem (4.10) stellen, siehe [11].
Für den Fall, dass die Nebenbedingungen durch lineare Funktionen dargestellt werden, sind keine besonderen Regularitätsbedingungen nötig und es gilt der nachstehende Satz.

Satz 4.23. *Sei $x^* \in \mathcal{F}$ eine lokale Lösung des Optimierungsproblems (4.10) mit linearen Funktionen e und f. Dann existieren $\lambda^* \in \mathbb{R}^p$, $\mu^* \in \mathbb{R}^m$, sodass der Vektor (x^*, λ^*, μ^*) die KKT-Bedingungen (4.13)-(4.17) erfüllt.*

Beweis. Siehe [11]. □

Hinreichende Optimalitätskriterien des Optimierungsproblems (4.10) sind generell von zweiter Ordnung, vgl. [11]. Beim konvexen Problem können sogar hinreichende Optimalitätskriterien erster Ordnung erzielt werden, vgl. Satz 4.25. Zudem gibt es hierbei eine einprägsame Regularitätsbedingung.

Definition 4.24. *Falls es einen Vektor $\tilde{x} \in \mathbb{R}^d$ mit $E\tilde{x} = e_c$ und $f(\tilde{x}) < f_c$ gibt, so genügt das konvexe Optimierungsproblem (4.11) den* ***Regularitätsbedingungen von Slater****.*

Zusammen mit den KKT-Bedingungen stellt die Regularitätsbedingungen von Slater eine notwendige Optimalitätsbedingung erster Ordnung dar.

Satz 4.25. *Seien $x^* \in \mathcal{F}$ eine optimale Lösung des konvexen Optimierungsproblems (4.11) und die Slater-Bedingung erfüllt. Dann existieren $\lambda^* \in \mathbb{R}^p$, $\mu^* \in \mathbb{R}^m$, sodass der Vektor (x^*, λ^*, μ^*) die KKT-Bedingungen (4.13)-(4.17) erfüllt.*

Beweis. Siehe [11]. □

Erfüllt umgekehrt $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^d \times \mathbb{R}^p \times \mathbb{R}^m$ die KKT-Bedingung, so ist x^* eine optimale Lösung von (4.10), was ebenfalls in [11] gezeigt wird. Die Gültigkeit der KKT-Bedingungen ist somit schon ein hinreichendes Optimalitätskriterium bei einem konvexen Optimierungsproblem.

Im Folgenden werden für Minimierungsprobleme auch Zielfunktionen der Gestalt $J : \mathbb{R}^d \to (-\infty, \infty]$ zugelassen. Für $\widetilde{D} := \{x \in \mathcal{F} \mid J(x) < \infty\}$ gelten die in diesem Kapitel getroffenen Aussagen analog. Falls $\widetilde{D} = \emptyset$, so wird jedes $x \in \mathcal{F}$ als optimale Lösung bezeichnet.

5. Das VER-Problem

In diesem Kapitel wird nun das volumenmaximale-einbeschriebene-Rechtecksmengen-Problem, kurz **VER-Problem**, vorgestellt. Dabei werden wichtige Eigenschaften des Problems untersucht. Ziel des VER-Problems ist es, eine bezüglich des Volumens maximale Teilmenge einer zulässigen Menge zu finden, die sich als kartesisches Produkt darstellen lässt.

5.1. Einführung in die Problemstellung

Zu Beginn werden einige grundlegende Begriffe geklärt, die helfen, das VER-Problem greifbar zu machen. Wo es möglich ist, wird sich dabei an die in [9] und [24] eingeführten Notationen gehalten. An entsprechender Stelle werden diese dann erweitert und ergänzt.

Grundlegendes

Ein Vektor im $x \in \mathbb{R}^d$ wird als **Design** bezeichnet, welches sich aus $d \in \mathbb{N}$ Designkomponenten zusammensetzt. Die Designkomponenten werden dann mit x_i, $i = 1, \ldots, d$, bezeichnet und es gilt $x = (x_1, \ldots, x_d)$.
Jedem Design $x \in \mathbb{R}^d$ kann über eine Funktion $f = (f_1, \ldots, f_m)^{\mathrm{T}} : \mathbb{R}^d \to \mathbb{R}^m$ ein **Output** $f(x)$ zugeordnet werden, weshalb f auch **Outputfunktion** heißt. Ein Output besteht aus m Outputkomponenten $f_j(x)$, $j = 1, \ldots, m$, mit $m \in \mathbb{N}$.

In der Anwendung sind in der Regel nur nach oben und unten beschränkte Designkomponenten realisierbar. Seien $x_i^{\mathrm{lb}} \in \mathbb{R}$ die unteren Grenzen der Designkomponenten und $x_i^{\mathrm{ub}} \in \mathbb{R}$ die oberen Grenzen mit $x_i^{\mathrm{lb}} < x_i^{\mathrm{ub}}$ für $i = 1, \ldots, d$. Damit ist $x_i^{\mathrm{lb}} \leq x_i \leq x_i^{\mathrm{ub}}$ bzw. $x_i \in [x_i^{\mathrm{lb}}, x_i^{\mathrm{ub}}]$ für $i = 1, \ldots, d$. Die Menge aller realisierbaren Designs ergibt sich dann über das d-fache kartesische Produkt

$$\Omega_{\mathrm{ds}} \quad := \quad \prod_{i=1}^{d} [x_i^{\mathrm{lb}}, x_i^{\mathrm{ub}}] \quad \subset \quad \mathbb{R}^d. \tag{5.1}$$

Diese Menge wird auch **Designmenge** genannt.

Des Weiteren sind aus technischer Sicht meist nur Designs sinnvoll, deren Outputkomponenten jeweils einen kritischen Wert nicht über- bzw. unterschreiten. Da jede nach unten beschränkte Outputkomponente durch Multiplikation mit -1 in eine nach oben beschränkte Komponente umgewandelt werden kann, ist die Annahme zulässig, dass für alle $x \in \mathbb{R}^d$ die Ungleichungen $f_j(x) \leq f_{\mathrm{c},j}$, $f_{\mathrm{c},j} \in \mathbb{R}$, $j = 1, \ldots, m$,

erfüllt sein müssen. Das m-Tupel, welches die Grenzen $f_{c,j}$, $j = 1, \ldots, m$, vereint, wird **Outputgrenze** f_c genannt und es gilt $f_c = (f_{c,1}, \ldots, f_{c,m})^T \in \mathbb{R}^m$.
Insgesamt können diese Bedingungen über das Ungleichungssystem

$$f(x) \leq f_c \tag{5.2}$$

zusammengefasst werden. Dieses muss für alle $x \in \mathbb{R}^d$ komponentenweise gültig sein, um die Anforderungen an den Output zu erfüllen.

Die Menge aller Designs $x \in \mathbb{R}^d$, welche in der Designmenge liegen und das Ungleichungssystem (5.2) erfüllen, wird als **zulässige Designmenge** Ω_{f_c} bezeichnet und es gilt

$$\Omega_{f_c} \quad := \quad \{x \in \Omega_{ds} \mid f(x) \leq f_c\} \quad \subset \quad \Omega_{ds}. \tag{5.3}$$

Beispiel 5.1. Sei die Designmenge über $\Omega_{ds} = [0,1]^2$ und die Outputfunktion über $f : \mathbb{R}^2 \to \mathbb{R},\ x \mapsto x_1 + x_2$ mit Outputgrenze $f_c = 1$ gegeben. In Abbildung 5.1 ist die zulässige Designmenge Ω_{f_c} hierfür visualisiert.

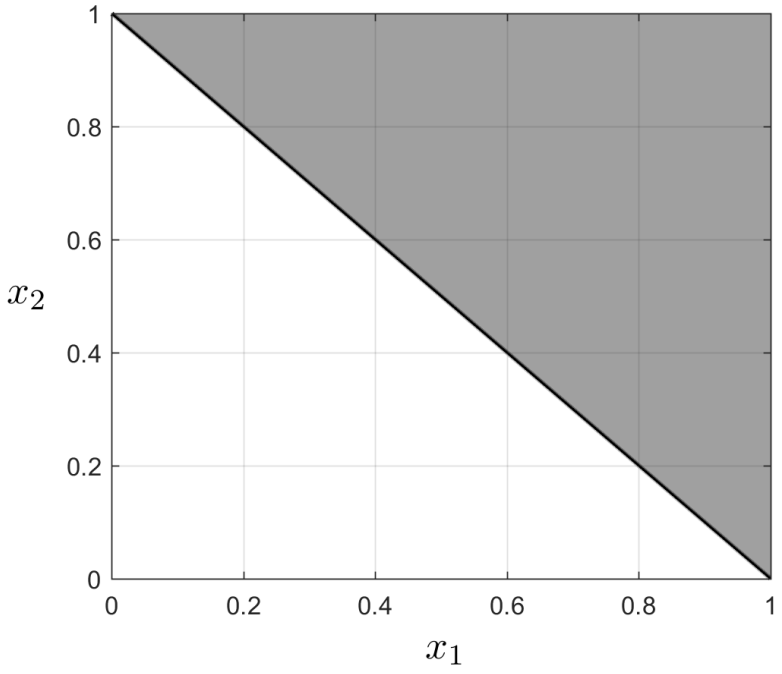


Abbildung 5.1: Zulässige Designmenge Ω_{f_c} (weiß)

Bemerkung 5.2. (i) Die Designmenge Ω_{ds} ist als kartesisches Produkt abgeschlossener Mengen eine Borel-Menge mit Lebesgue-Maß

$$\lambda^d(\Omega_{\mathrm{ds}}) \;=\; \prod_{i=1}^{d}(x_i^{\mathrm{ub}} - x_i^{\mathrm{lb}}) \;<\; \infty. \tag{5.4}$$

(ii) Für die zulässige Menge gilt

$$\begin{aligned}\Omega_{f_{\mathrm{c}}} &= \bigcap_{j=1}^{m}\{x \in \Omega_{\mathrm{ds}} \mid f_j(x) \le f_{\mathrm{c},j}\}\\ &= \bigcap_{j=1}^{m}\Big(\Omega_{\mathrm{ds}} \cap \{x \in \mathbb{R}^d \mid f_j(x) \le f_{\mathrm{c},j}\}\Big)\\ &= \Omega_{\mathrm{ds}} \cap \bigcap_{j=1}^{m}\{x \in \mathbb{R}^d \mid f_j(x) \le f_{\mathrm{c},j}\}. \end{aligned}\tag{5.5}$$

Falls die Outputfunktion $f : \mathbb{R}^d \to \mathbb{R}^m$ Borel-messbar ist, sind die Mengen $\{x \in \mathbb{R}^d \mid f_j(x) \le f_{\mathrm{c},j}\}$, $j = 1,\dots,m$, nach Korollar 3.15 Borel-messbar. Somit ist die zulässige Designmenge als Schnitt von Borel-Mengen ebenfalls eine Borel-Menge. Das Lebesgue-Maß von $\Omega_{f_{\mathrm{c}}}$ ist wegen $\Omega_{f_{\mathrm{c}}} \subset \Omega_{\mathrm{ds}}$ endlich und es gilt

$$\lambda^d(\Omega_{f_{\mathrm{c}}}) \;\le\; \lambda^d(\Omega_{\mathrm{ds}}) \;<\; \infty.$$

Bevor das VER-Problem formuliert werden kann, müssen noch die Dimensionen der Mengen festgelegt werden, aus denen sich das mehrfache kartesische Produkt zusammensetzt, welches das Problem lösen soll. Dazu ist die folgende Definition hilfreich.

Definition 5.3. *Seien $d, n \in \mathbb{N}$ fest und $s := (d^1,\dots,d^n) \in \mathbb{N}^n$. Falls $\sum_{k=1}^{n} d^k = d$ gilt bzw. im Sinne der Multiindexschreibweise, welche in Beispiel 2.7 eingeführt wurde, $|s| = d$, so wird das n-Tupel s **Struktur** genannt.*

Für gegebenes s kann ein n-faches kartesisches Produkt aus Mengen $\Omega^k \in B(\mathbb{R}^{d^k})$, $k = 1,\dots,n$, gebildet werden, für das dann $\prod_{k=1}^{n}\Omega^k \subset \mathbb{R}^d$ gilt. Die Menge $\Omega := \prod_{k=1}^{n}\Omega^k$ wird daher auch als n-faches kartesisches Produkt mit Struktur s bezeichnet. Dabei ist die in Kapitel 2.2 beschriebene Identifikation von $\prod_{k=1}^{n}\mathbb{R}^{d^k}$ und $\mathbb{R}^d$ zu beachten. Insbesondere sei an dieser Stelle auch nochmals an die möglichen Darstellungen eines

Designs $x \in \mathbb{R}^d$ erinnert, welche in Gleichung (2.4) aufgelistet sind. Für das Lebesgue-Maß von Ω gilt wegen Gleichung (3.4) gerade

$$\lambda^d(\Omega) \quad = \quad \prod_{k=1}^{n} \lambda^{d^k}(\Omega^k).$$

In dieser Arbeit wird zusätzlich $d^{k-1} \geq d^k$ für $k = 2, \ldots, n$ vorausgesetzt, da die Indizes der Designkomponenten beliebig umgeordnet werden können.

Für jede Struktur s ist die Designmenge ein mehrfaches kartesisches Produkt mit Struktur s. Es sei

$$\Omega_{\mathrm{ds}}^k \quad := \quad \prod_{i=i^{k-1}+1}^{i^k} [x_i^{\mathrm{lb}}, x_i^{\mathrm{ub}}] \ \subset \ \mathbb{R}^{d^k} \tag{5.6}$$

für $k = 1, \ldots, n$ mit $i^k = \sum_{\tilde{k}=1}^{k} d^{\tilde{k}}$. Dann gilt $\Omega_{\mathrm{ds}} = \prod_{k=1}^{n} \Omega_{\mathrm{ds}}^k$.
Analog dazu wird die Outputfunktion f additiv zerlegbar mit Struktur s genannt, falls Funktionen $f^k = (f_1^k, \ldots, f_m^k) : \mathbb{R}^{d^k} \to \mathbb{R}^m$, $k = 1, \ldots, n$, existieren, sodass $f(x) = \sum_{k=1}^{n} f^k(x^k)$ für alle $x \in \mathbb{R}^d$ erfüllt ist, vgl. Definition 2.17. Hierbei sei bemerkt, dass es zu einer Outputfunktion in der Regel mehrere Möglichkeiten gibt die Struktur s zu wählen. So sind beispielsweise für die Funktion $f : \mathbb{R}^3 \to \mathbb{R},\ x \mapsto x_1 + x_2 + x_3$ die Strukturen $s = (1, 1, 1)$, $s = (2, 1)$ und $s = 3$ möglich.

Falls die Einträge der Struktur s alle den Wert 1 besitzen, so wird auch von einer eindimensionalen Struktur gesprochen. Dies ist erfüllt sofern $d^1 = 1$ gilt, da dann für $k = 2, \ldots, n$ sofort $d^k = 1$ folgt. Ist der maximale Eintrag der Struktur s vom Wert 2, wobei mindestens ein Eintrag diesen Wert auch annimmt, so handelt es sich um eine zweidimensionale Struktur. Hierbei ist $d^1 = 2$ notwendige und hinreichende Bedingung.

Nun kann das VER-Problem als Optimierungsproblem aufgestellt werden. Gesucht wird eine bezüglich des Volumens maximale Teilmenge der zulässigen Designmenge, welche ein kartesisches Produkt mit Struktur s darstellt.
Unter Verwendung der zulässigen Designmenge $\Omega_{f_{\mathrm{c}}}$ aus Beispiel 5.1 und $s = (1, 1)$ bedeutet dies, dass ein volumenmaximales Rechteck gesucht ist, welches in $\Omega_{f_{\mathrm{c}}}$ enthalten ist. Kandidaten möglicher Rechtecke sind in Abbildung 5.2 veranschaulicht.

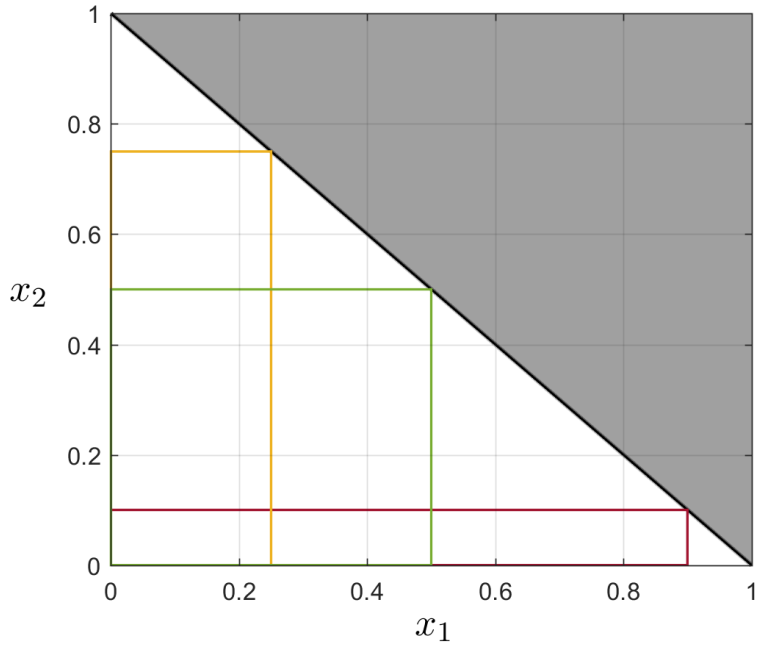


Abbildung 5.2: Kandidaten für ein volumenmaximales Rechteck in Ω_{f_c} (gelb, grün und rot umrandet)

Abstraktes Problem

Definition 5.4. *Seien die Designmenge $\Omega_{\mathrm{ds}} \subset \mathbb{R}^d$, die Struktur $s \in \mathbb{N}^n$ sowie die Borel-messbare Outputfunktion $f : \mathbb{R}^d \to \mathbb{R}^m$ mit Outputgrenze $f_c \in \mathbb{R}^m$ gegeben. Das **abstrakte VER-Problem** lautet*

$$\begin{aligned} &\underset{\Omega^k \in \mathcal{B}(\mathbb{R}^{d^k}),\, k=1,\dots,n}{\text{maximiere}} && \lambda^d\left(\prod_{k=1}^{n} \Omega^k\right) \\ &\text{u.d.N.} && \prod_{k=1}^{n} \Omega^k \subset \Omega_{f_c}, \end{aligned} \tag{5.7}$$

wobei die zulässige Designmenge Ω_{f_c} wie in (5.3) definiert ist.

Die Zielfunktion des abstrakten VER-Problems wird im Folgenden auch mit $\widetilde{J}_{\mathrm{ab}}$ bezeichnet und es gilt

$$\widetilde{J}_{\mathrm{ab}} : \prod_{k=1}^{n} \mathcal{B}(\mathbb{R}^{d^k}) \to [0,\infty),\ (\Omega^1,\dots,\Omega^n) \mapsto \lambda^d\left(\prod_{k=1}^{n} \Omega^k\right). \tag{5.8}$$

Die Menge $\prod_{k=1}^{n} \Omega^k$ liegt dabei bekanntermaßen in $\mathcal{B}(\mathbb{R}^d)$ und ihr Lebesgue-Maß ist

als Teilmenge von Ω_{f_c} endlich.

Eine optimale Lösung des abstrakten VER-Problems muss, sofern sie existiert, die Eigenschaft der folgenden Definition erfüllen.

Definition 5.5. *Die Mengen $\Omega^{k,*} \in \mathcal{B}(\mathbb{R}^{d^k})$, $k = 1, \ldots, n$, mit $\prod_{k=1}^{n} \Omega^{k,*} \subset \Omega_{f_c}$ werden **optimale Lösung** des abstrakten VER-Problems (5.7) genannt, wenn die Ungleichung*

$$\lambda^d \left(\prod_{k=1}^{n} \Omega^{k,*} \right) \geq \lambda^d \left(\prod_{k=1}^{n} \Omega^{k} \right) \tag{5.9}$$

bzw. $\widetilde{J}_{\mathrm{ab}}(\Omega^{1,*}, \ldots, \Omega^{n,*}) \geq \widetilde{J}_{\mathrm{ab}}(\Omega^1, \ldots, \Omega^n)$ für alle $\Omega^k \in \mathcal{B}(\mathbb{R}^{d^k})$, $k = 1, \ldots, n$, mit der Eigenschaft $\prod_{k=1}^{n} \Omega^k \subset \Omega_{f_c}$ erfüllt ist.

Eine optimale Lösung des Problems (5.7) maximiert auch die Komposition $\log \circ \widetilde{J}_{\mathrm{ab}}$ mit $\log(0) := -\infty$, da der Logarithmus monoton wachsend auf $(0, \infty)$ ist. Ferner kann, wie bereits in Kapitel 4.2 beschrieben, durch Änderung des Vorzeichens der Zielfunktion jedes Maximierungsproblem in ein Minimierungsproblem überführt werden. Deshalb minimieren $\Omega^{k,*} \in \mathcal{B}(\mathbb{R}^{d^k})$, $k = 1, \ldots, n$, die Funktion $J_{\mathrm{ab}} := -\log \circ \widetilde{J}_{\mathrm{ab}}$. Unter den Voraussetzungen von Definition 5.4 lautet das abstrakte VER-Problem dann

$$\begin{aligned} &\underset{\Omega^k \in \mathcal{B}(\mathbb{R}^{d^k}),\, k=1,\ldots,n}{\text{minimiere}} && -\log\left(\lambda^d \left(\prod_{k=1}^{n} \Omega^k \right) \right) \\ &\text{u.d.N.} && \prod_{k=1}^{n} \Omega^k \subset \Omega_{f_c}. \end{aligned} \tag{5.10}$$

Die Komposition von $-\log$ und der Zielfunktion ermöglicht es später ein konvexes Optimierungsproblem zu erhalten, wie im Laufe der Arbeit ersichtlich wird. Stellvertretend für das abstrakte VER-Problem wird daher nur noch Bezug auf das Optimierungsproblem (5.10) genommen.

Das abstrakte Problem ist aus mathematischer Sicht nicht einfach zu lösen, da Mengen der Form $\Omega^k \in \mathcal{B}(\mathbb{R}^{d^k})$, $k = 1, \ldots, n$, in einem Algorithmus nur schwer zu handhaben sind. Ziel ist es deshalb, diese Mengen analytisch über Variablen zu beschreiben, welche Ungleichungs- bzw. Gleichungsbedingungen einhalten müssen, um $\prod_{k=1}^{n} \Omega^k \subset \Omega_{f_c}$ zu gewährleisten. Falls die Outputfunktion f additiv zerlegbar mit Struktur s ist, so ist das abstrakte Problem als ein solches analytisches Problem

formulierbar, wie im Folgenden gezeigt wird.

5.2. Analytisches Problem

Definition 5.6. *Seien die Designmenge $\Omega_{\mathrm{ds}} \subset \mathbb{R}^d$ sowie die Borel-messbare Outputfunktion $f : \mathbb{R}^d \to \mathbb{R}^m$ mit Outputgrenze $f_{\mathrm{c}} \in \mathbb{R}^m$ gegeben. Ferner sei f additiv zerlegbar mit Struktur $s \in \mathbb{N}^n$. Das* ***analytische VER-Problem*** *lautet dann*

$$\begin{aligned} \underset{f_{\mathrm{c}}^k \in \mathbb{R}^m,\ k=1,\dots,n}{\text{minimiere}} \quad & -\log\left(\lambda^d\left(\prod_{k=1}^{n} \Omega_{f_{\mathrm{c}}^k}^k\right)\right) \\ \text{u.d.N.} \quad & \sum_{k=1}^{n} f_{\mathrm{c}}^k = f_{\mathrm{c}}, \end{aligned} \tag{5.11}$$

wobei $\Omega_{f_{\mathrm{c}}^k}^k := \{x^k \in \Omega_{\mathrm{ds}}^k \mid f^k(x^k) \leq f_{\mathrm{c}}^k\}$.

Die Mengen $\Omega_{f_{\mathrm{c}}^k}^k$ in Definition 5.6 sind dabei stets als Funktionswert der Funktionen $\Omega_{(\cdot)}^k : \mathbb{R}^m \to \mathcal{B}(\mathbb{R}^{d^k}),\ f_{\mathrm{c}}^k \mapsto \Omega_{f_{\mathrm{c}}^k}^k,\ k = 1, \dots, n$, zu verstehen. Aufgrund Satz 3.21 und Bemerkung 5.2 liegen $\Omega_{f_{\mathrm{c}}^k}^k$ für $f_{\mathrm{c}}^k \in \mathbb{R}^m$ in $\mathcal{B}(\mathbb{R}^{d^k})$, $k = 1, \dots, n$. Somit gilt auch $\prod_{k=1}^{n} \Omega_{f_{\mathrm{c}}^k}^k \in \mathcal{B}(\mathbb{R}^d)$. Insgesamt konnte dadurch eine analytische Beschreibung der Mengen $\Omega^k \in \mathcal{B}(\mathbb{R}^{d^k})$, $k = 1, \dots, n$, über Vektoren im $\mathbb{R}^m$ wie gewünscht erreicht werden, was in Satz 5.8 nochmals deutlicher wird.
Für Zielfunktion des analytischen VER-Problems, welche auch mit J_{an} bezeichnet wird, gilt deshalb

$$J_{\mathrm{an}} : \prod_{k=1}^{n} \mathbb{R}^m \to (-\infty, \infty],\ (f_{\mathrm{c}}^1, \dots, f_{\mathrm{c}}^n) \mapsto -\log\left(\lambda^d\left(\prod_{k=1}^{n} \Omega_{f_{\mathrm{c}}^k}^k\right)\right).$$

Eine optimale Lösung des analytischen VER-Problems erfüllt, sofern sie existiert, die Eigenschaft der folgenden Definition. Diese Definition ist identisch mit der Definition einer globalen Lösung eines allgemeinen Minimierungsproblems, siehe Definition 4.18. Analog könnten lokale Lösungen des analytischen Problems formuliert werden.

Definition 5.7. *Die Vektoren $f_{\mathrm{c}}^{k,*} \in \mathbb{R}^m$, $k = 1, \dots, n$, mit $\sum_{k=1}^{n} f_{\mathrm{c}}^{k,*} = f_{\mathrm{c}}$ werden* ***optimale Lösung*** *des analytischen VER-Problems (5.11) genannt, wenn die Ungleichung*

$$-\log\left(\lambda^d\left(\prod_{k=1}^{n} \Omega_{f_{\mathrm{c}}^{k,*}}^k\right)\right) \quad \leq \quad -\log\left(\lambda^d\left(\prod_{k=1}^{n} \Omega_{f_{\mathrm{c}}^k}^k\right)\right)$$

bzw. $J_{\text{an}}(f_{\text{c}}^{1,*}, \ldots, f_{\text{c}}^{n,*}) \leq J_{\text{an}}(f_{\text{c}}^{1}, \ldots, f_{\text{c}}^{n})$ *für alle* $f_{\text{c}}^{k} \in \mathbb{R}^m$, $k = 1, \ldots, n$, *mit der Eigenschaft* $\sum_{k=1}^{n} f_{\text{c}}^{k} = f_{\text{c}}$ *erfüllt ist.*

Der Hauptunterschied zwischen dem abstrakten und dem analytischen VER-Problem ist, dass das abstrakte Problem allgemeiner formuliert ist und somit für beliebige Strukturen und beliebige Borel-messbare Funktionen gilt. Beim analytischen VER-Problem sind die Auswahlmöglichkeiten für die Struktur bereits durch f festgelegt. Im schlechtesten Fall ist nur $s = d$ mit $n = 1$ möglich und $f_{\text{c}} \in \mathbb{R}^m$ ist die einzige zulässige Lösung des zugehörigen Problems (5.11).
Unter der Voraussetzung, dass die Funktion f mit Struktur $s \in \mathbb{N}^n$ additiv zerlegbar ist, kann gezeigt werden, dass die zur optimalen Lösung des analytischen Problems zugehörigen Mengen $\Omega^k_{f_{\text{c}}^{k,*}}$, $k = 1, \ldots, n$, eine optimale Lösung des abstrakten Problems sind.

Satz 5.8. *Wenn die Vektoren* $f_{\text{c}}^{k,*} \in \mathbb{R}^m$, $k = 1, \ldots, n$, *mit* $\sum_{k=1}^{n} f_{\text{c}}^{k,*} = f_{\text{c}}$ *eine optimale Lösung des analytischen VER-Problems (5.11) sind, dann sind die Mengen* $\Omega^k_{f_{\text{c}}^{k,*}}$, $k = 1, \ldots, n$, *eine optimale Lösung des abstrakten VER-Problems (5.10).*

Beweis. Sei $f_{\text{c}}^{k,*} \in \mathbb{R}^m$, $k = 1, \ldots, n$, eine optimale Lösung des analytischen VER-Problems (5.11) mit $\sum_{k=1}^{n} f_{\text{c}}^{k,*} = f_{\text{c}}$. Angenommen es existieren Mengen $\Omega^k \in \mathcal{B}(\mathbb{R}^{d^k})$, $k = 1, \ldots, n$, mit $\prod_{k=1}^{n} \Omega^k \subset \Omega_{f_{\text{c}}}$ und

$$-\log\left(\lambda^d\left(\prod_{k=1}^{n}\Omega^k\right)\right) \;\leq\; -\log\left(\lambda^d\left(\prod_{k=1}^{n}\Omega^k_{f_{\text{c}}^{k,*}}\right)\right). \tag{5.12}$$

Dann gilt für alle $x \in \prod_{k=1}^{n} \Omega^k$ gerade $f(x) \leq f_{\text{c}}$. Da f additiv zerlegbar ist und $\prod_{k=1}^{n} \Omega^k \subset \Omega_{\text{ds}}$ beschränkt ist, existieren $f_{\text{c}}^k \in \mathbb{R}^m$, $k = 1, \ldots, n$, mit $f^k(x^k) \leq f_{\text{c}}^k$ für alle $x^k \in \Omega^k$ und $\sum_{k=1}^{n} f_{\text{c}}^k \leq f_{\text{c}}$. Aufgrund der Tatsache, dass durch komponentenweises Vergrößern einzelner f_{c}^k, der Wert der Zielfunktion niemals ansteigen kann, vgl. den folgenden Satz 5.9 über die Monotonie der Zielfunktion, darf $\sum_{k=1}^{n} f_{\text{c}}^k = f_{\text{c}}$ angenommen werden. Somit ist Ungleichung (5.12) äquivalent zu

$$-\log\left(\lambda^d\left(\prod_{k=1}^{n}\Omega^k_{f_{\text{c}}^k}\right)\right) \;\leq\; -\log\left(\lambda^d\left(\prod_{k=1}^{n}\Omega^k_{f_{\text{c}}^{k,*}}\right)\right),$$

was einen Widerspruch zur Optimalität der Vektoren $f_{\mathrm{c}}^{k,*} \in \mathbb{R}^m$, $k = 1, \ldots, n$, darstellt. □

Eigenschaften der Zielfunktion

Um die Eigenschaften des analytischen VER-Problems zu untersuchen, ist es oftmals hilfreich die Vektoren $f_{\mathrm{c}}^k \in \mathbb{R}^m$, $k = 1, \ldots, n$, in einem $\mathbb{R}^{mn}$-Vektor zusammenzufassen. Analog zu den Betrachtungen in Kapitel 2.2 ist dann $\hat{f}_{\mathrm{c}} := (f_{\mathrm{c}}^1, \ldots, f_{\mathrm{c}}^n)^{\mathrm{T}} \in \mathbb{R}^{mn}$. Diese Darstellung ist auch bei der Berechnung von numerischen Lösungen der VER-Probleme von Bedeutung.

Satz 5.9. *Die Zielfunktion J_{an} des analytischen VER-Problems (5.11) ist*

(i) monoton fallend,

(ii) nach unten halbstetig.

Beweis. (i) Für die Monotonie von J_{an} auf dem $\mathbb{R}^{mn}$ ist zu zeigen, dass für alle $\hat{f}_{\mathrm{c}}, \hat{g}_{\mathrm{c}} \in \mathbb{R}^{mn}$ mit $\hat{f}_{\mathrm{c}} \leq \hat{g}_{\mathrm{c}}$ gerade $J_{\mathrm{an}}(\hat{f}_{\mathrm{c}}) \geq J_{\mathrm{an}}(\hat{g}_{\mathrm{c}})$ erfüllt ist. Seien $f_{\mathrm{c},j}^k \leq g_{\mathrm{c},j}^k$, $j = 1, \ldots, m$, $k = 1, \ldots, n$, dann gelten

$$\{x^k \in \Omega_{\mathrm{ds}}^k \mid f_j^k(x^k) \leq f_{\mathrm{c},j}^k\} \subset \{x^k \in \Omega_{\mathrm{ds}}^k \mid f_j^k(x^k) \leq g_{\mathrm{c},j}^k\}$$

und

$$\begin{aligned} \Omega_{f_{\mathrm{c}}^k}^k &= \bigcap_{j=1}^m \{x^k \in \Omega_{\mathrm{ds}}^k \mid f_j^k(x^k) \leq f_{\mathrm{c},j}^k\} \\ &\subset \bigcap_{j=1}^m \{x^k \in \Omega_{\mathrm{ds}}^k \mid f_j^k(x^k) \leq g_{\mathrm{c},j}^k\} = \Omega_{g_{\mathrm{c}}^k}^k. \end{aligned}$$

Ferner ist mit Satz 2.12 dann

$$\prod_{k=1}^n \Omega_{f_{\mathrm{c}}^k}^k \subset \prod_{k=1}^n \Omega_{g_{\mathrm{c}}^k}^k \tag{5.13}$$

für $k = 1, \ldots, n$ und aufgrund Satz 3.23 auch

$$\lambda^d \left(\prod_{k=1}^n \Omega_{f_{\mathrm{c}}^k}^k \right) \leq \lambda^d \left(\prod_{k=1}^n \Omega_{g_{\mathrm{c}}^k}^k \right). \tag{5.14}$$

Die Ungleichung (5.14) besagt, dass die Funktion

$$\widetilde{J}_{\mathrm{an}} : \prod_{k=1}^{n} \mathbb{R}^m \to [0,\infty),\ (f_{\mathrm{c}}^1,\ldots,f_{\mathrm{c}}^n) \mapsto \lambda^d\left(\prod_{k=1}^{n} \Omega_{f_{\mathrm{c}}^k}^k\right)$$

monoton wachsend ist. Daher ist J_{an} wegen $J_{\mathrm{an}} = -\log\circ\widetilde{J}_{\mathrm{an}}$ monoton fallend.

(ii) Hierzu wird gezeigt, dass $\widetilde{J}_{\mathrm{an}}$ nach oben halbstetig ist. Mit [19] folgt dann, dass $\log\circ\widetilde{J}_{\mathrm{an}}$ als Komposition einer stetigen, monoton wachsenden Funktion und einer nach oben halbstetigen Funktion ebenfalls nach oben halbstetig ist. Damit ist $J_{\mathrm{an}} = -\log\circ\widetilde{J}_{\mathrm{an}}$ schließlich nach unten halbstetig.
Sei $\hat{f} \in \mathbb{R}^{mn}$ gegeben. So ist die Ungleichung

$$\lim_{\delta\searrow 0}\left(\sup_{\hat{g}\in B(\hat{f},\delta)} \widetilde{J}_{\mathrm{an}}(\hat{g})\right) \quad\leq\quad \widetilde{J}_{\mathrm{an}}(\hat{f})$$

nachzuweisen. Als Norm im $\mathbb{R}^{mn}$ kann hier die ∞-Norm gewählt werden, d.h. $B(\hat{f},\delta) = \{\hat{g} \in \mathbb{R}^{mn} \mid \|\hat{g}-\hat{f}\|_\infty < \delta\}$. Aufgrund der Monotonie von $\widetilde{J}_{\mathrm{an}}$ gilt für $\hat{\delta} := (\delta,\ldots,\delta) \in \mathbb{R}^{mn}$ mit $\delta \in (0,\infty)$

$$\sup_{\hat{g}\in B(\hat{f},\delta)} \widetilde{J}_{\mathrm{an}}(\hat{g}) \quad\leq\quad \widetilde{J}_{\mathrm{an}}(\hat{f}+\hat{\delta}).$$

Sei $(\delta_i)_{i\in\mathbb{N}}$ eine monoton fallende Folge mit $\delta_i \in (0,\infty)$, $i \in \mathbb{N}$ und $\lim\limits_{i\to\infty} \delta_i = 0$. So bleibt zu zeigen, dass die Ungleichung

$$\lim_{i\to\infty} \widetilde{J}_{\mathrm{an}}(\hat{f}+\hat{\delta}_i) \quad\leq\quad \widetilde{J}_{\mathrm{an}}(\hat{f})$$

erfüllt ist, wobei $\hat{\delta}_i := (\delta_i,\ldots,\delta_i) \in \mathbb{R}^{mn}$. Wegen Satz 2.12 und Bemerkung 5.2 gilt

$$\begin{aligned}\prod_{k=1}^{n} \Omega_{f_{\mathrm{c}}^k}^k &= \prod_{k=1}^{n}\bigcap_{j=1}^{m}\{x^k \in \Omega_{\mathrm{ds}}^k \mid f_j^k(x^k) \leq f_{\mathrm{c},j}^k\} \\ &= \Omega_{\mathrm{ds}} \cap \prod_{k=1}^{n}\bigcap_{j=1}^{m}(f_j^k)^{-1}\big((-\infty, f_{\mathrm{c},j}^k]\big). \qquad (5.15)\end{aligned}$$

Zur Übersichtlichkeit werden für den weiteren Beweis die beiden Schreibweisen $\Omega := \prod\limits_{k=1}^{n} \Omega_{f_{\mathrm{c}}^k}^k$ und $\Omega_i := \prod\limits_{k=1}^{n} \Omega_{f_{\mathrm{c}}^k+\delta_i^k}^k$ mit $\delta_i^k := (\delta_i,\ldots,\delta_i) \in \mathbb{R}^m$, $i \in \mathbb{N}$ eingeführt. Wie in (5.13) ist $\Omega_i \supset \Omega_{i+1}$ für $i \in \mathbb{N}$ und mit den Eigenschaften von

Bild und Urbild bzgl. dem Durchschnitt von Familien von Mengen auch

$$\begin{aligned}\bigcap_{i\in\mathbb{N}}\Omega_i &= \bigcap_{i\in\mathbb{N}}\left(\Omega_{\mathrm{ds}}\cap\prod_{k=1}^{n}\bigcap_{j=1}^{m}(f_j^k)^{-1}\big((-\infty,f_{\mathrm{c},j}^k+\delta_i]\big)\right)\\ &= \Omega_{\mathrm{ds}}\cap\prod_{k=1}^{n}\bigcap_{j=1}^{m}(f_j^k)^{-1}\left(\bigcap_{i\in\mathbb{N}}(-\infty,f_{\mathrm{c},j}^k+\delta_i]\right)\\ &= \Omega_{\mathrm{ds}}\cap\prod_{k=1}^{n}\bigcap_{j=1}^{m}(f_j^k)^{-1}\big((-\infty,f_{\mathrm{c},j}^k]\big) \;=\; \Omega,\end{aligned}$$

d.h. $\Omega_i \downarrow \Omega$. Ferner lässt sich mit der Darstellung aus Gleichung (5.15) auch erkennen, dass $\Omega_1 \subset \Omega_{\mathrm{ds}}$ ist und damit $\lambda^d(\Omega_1) \leq \lambda^d(\Omega_{\mathrm{ds}}) < \infty$. Somit sind die Voraussetzungen an die σ-Stetigkeit aus Satz 3.23 erfüllt, weshalb

$$\lim_{i\to\infty}\widetilde{J}_{\mathrm{an}}\big(\hat{f}+\hat{\delta}_i\big) \;=\; \lim_{i\to\infty}\lambda^d(\Omega_i) \;=\; \lambda^d(\Omega) \;=\; \widetilde{J}_{\mathrm{an}}\big(\hat{f}\big)$$

gilt und daher $\widetilde{J}_{\mathrm{an}}$ nach oben halbstetig ist.

□

Aus dem Beweis von Satz 5.8 wird ersichtlich, dass das analytische VER-Problem eigentlich auch mit den Ungleichungsbedingungen $\sum_{k=1}^{n} f_{\mathrm{c}}^k \leq f_{\mathrm{c}}$ anstatt mit den Gleichungsbedingungen $\sum_{k=1}^{n} f_{\mathrm{c}}^k = f_{\mathrm{c}}$ hätte formuliert werden können. Da die Zielfunktion J_{an} monoton fallend ist, wird eine optimale Lösung immer für $\sum_{k=1}^{n} f_{\mathrm{c}}^k = f_{\mathrm{c}}$ angenommen, falls sie existiert. Um die Existenz einer optimalen Lösung des VER-Problems nachzuweisen, ist der zweite Teil des Satzes 5.9 entscheidend. Zuvor müssen jedoch noch Eigenschaften der zulässigen Menge des Problems (5.11) untersucht werden.

Eigenschaften der zulässigen Menge

Satz 5.10. *Die zulässige Menge* $\mathcal{F}_{\mathrm{an}} = \{\hat{f}_{\mathrm{c}} \in \mathbb{R}^{mn} \mid \sum_{k=1}^{n} f_{\mathrm{c}}^k = f_{\mathrm{c}}\}$ *des analytischen VER-Problems (5.11) ist konvex.*

Beweis. Die Bedingung $\sum_{k=1}^{n} f_{\mathrm{c}}^k = f_{\mathrm{c}}$ kann geschrieben werden als $E\hat{f}_{\mathrm{c}} = f_{\mathrm{c}}$ mit $E := (I_m, \ldots, I_m) \in \mathbb{R}^{m\times mn}$, wobei I_m die $m \times m$-Einheitsmatrix ist. Somit gilt

$$\mathcal{F}_{\mathrm{an}} \;=\; \{\hat{f}_{\mathrm{c}} \in \mathbb{R}^{mn} \mid E\hat{f}_{\mathrm{c}} = f_{\mathrm{c}}\}$$

und nach Beispiel 4.2 ist $\mathcal{F}_{\text{an}}$ konvex. □

Mithilfe der Darstellung $\mathcal{F}_{\text{an}} = \{\hat{f}_{\text{c}} \in \mathbb{R}^{mn} \mid E\hat{f}_{\text{c}} = f_{\text{c}}\}$, $E := (I_m, \ldots, I_m) \in \mathbb{R}^{m \times mn}$ kann das analytische VER-Problem (5.11) auch in der kompakten Form

$$\begin{aligned} &\underset{\hat{f}_{\text{c}} \in \mathbb{R}^{mn}}{\text{minimiere}} && J_{\text{an}}(\hat{f}_{\text{c}}) \\ &\text{u.d.N.} && E\hat{f}_{\text{c}} = f_{\text{c}} \end{aligned} \tag{5.16}$$

geschrieben werden. Um eine optimale Lösung des analytischen VER-Problem zu finden, ist es vorteilhaft, diejenigen $\hat{f}_{\text{c}}$ aus $\mathcal{F}_{\text{an}}$ zu „entfernen“, für welche die Zielfunktion den Wert ∞ annimmt.

Satz und Definition 5.11. *Für das analytische VER-Problem (5.11) ist die Menge*

$$\mathcal{F}_{\text{an,c}} := \left\{ \hat{f}_{\text{c}} \in \mathcal{F}_{\text{an}} \;\middle|\; \underline{f}^k_{\text{c},j} \leq f^k_{\text{c},j} \leq f_{\text{c},j} - \sum_{\substack{\tilde{k}=1,\\ \tilde{k}\neq k}}^{n} \underline{f}^{\tilde{k}}_{\text{c},j},\ j = 1, \ldots, m,\ k = 1, \ldots, n \right\}$$

mit $\underline{f}^k_{\text{c},j} := \inf\limits_{x^k \in \Omega^k_{\text{ds}}} f^k_j(x^k)$ *für* $j = 1, \ldots, m$, $k = 1, \ldots, n$ *kompakt und die Zielfunktion nimmt für alle* $\hat{f}_{\text{c}} \in \mathcal{F}_{\text{an}} \backslash \mathcal{F}_{\text{an,c}}$ *den Wert* ∞ *an. Darum wird* $\mathcal{F}_{\text{an,c}}$ *auch* ***eingeschränkte zulässige Menge*** *des Problems (5.11) genannt.*

Beweis. Weil die Mengen Ω^k_{ds} beschränkt sind und die Funktionen f^k_j von $\mathbb{R}^{d^k}$ nach $\mathbb{R}$ abbilden, gilt $\inf\limits_{x^k \in \Omega^k_{\text{ds}}} f^k_j(x^k) \in \mathbb{R}$ für $j = 1, \ldots, m$, $k = 1, \ldots, n$. Die Menge $\mathcal{F}_{\text{an,c}}$ ist somit beschränkt und abgeschlossen und daher kompakt.
Sei $\hat{f}_{\text{c}} \in \mathcal{F}_{\text{an}}$ mit $f^k_{\text{c},j} < \underline{f}^k_{\text{c},j}$ für ein $j \in \{1, \ldots, m\}$ und ein $k \in \{1, \ldots, n\}$, also $\hat{f}_{\text{c}} \in \mathcal{F}_{\text{an}} \backslash \mathcal{F}_{\text{an,c}}$. Dann gilt $(f^k_j)^{-1}((-\infty, f^k_{\text{c},j}]) = \{x^k \in \Omega^k_{\text{ds}} \mid f^k_j(x^k) \leq f^k_{\text{c},j}\} = \emptyset$ für selbe $j \in \{1, \ldots, m\}$ und $k \in \{1, \ldots, n\}$ und ferner $\Omega^k_{f^k_{\text{c}}} = \bigcap\limits_{j=1}^{m} (f^k_j)^{-1}((-\infty, f^k_{\text{c},j}]) = \emptyset$ für eben dieses $k \in \{1, \ldots, n\}$. Somit ist

$$\lambda^d \left(\prod_{k=1}^{n} \Omega^k_{f^k_{\text{c}}} \right) = \prod_{k=1}^{n} \lambda^{d^k} \left(\Omega^k_{f^k_{\text{c}}} \right) = 0$$

bzw. $J_{\text{an}}(f^1_{\text{c}}, \ldots, f^n_{\text{c}}) = \infty$. Also muss $f^k_{\text{c},j} \geq \underline{f}^k_{\text{c},j}$ für alle $j \in \{1, \ldots, m\}$ und alle $k \in \{1, \ldots, n\}$, erfüllt sein, damit $J_{\text{an}}(f^1_{\text{c}}, \ldots, f^n_{\text{c}}) < \infty$ gilt.
Sei jetzt $\hat{f}_{\text{c}} \in \mathcal{F}_{\text{an}}$ mit $f^k_{\text{c},j} \geq \underline{f}^k_{\text{c},j}$ für $j = 1, \ldots, m$ und $k = 1, \ldots, m$. Ferner sei

$f^k_{\mathrm{c},j} > f_{\mathrm{c},j} - \sum\limits_{\substack{\tilde{k}=1,\\ \tilde{k}\neq k}}^{n} \underline{f}^{\tilde{k}}_{\mathrm{c},j}$ für ein $j \in \{1,\dots,m\}$ und ein $k \in \{1,\dots,n\}$, dann sind wegen $f^k_{\mathrm{c},j} + \sum\limits_{\substack{\tilde{k}=1,\\ \tilde{k}\neq k}}^{n} \underline{f}^{\tilde{k}}_{\mathrm{c},j} > f_{\mathrm{c},j}$ die Nebenbedingungen des analytischen VER-Problems verletzt und $\hat{f}_{\mathrm{c}}$ kann nicht in $\mathcal{F}_{\mathrm{an}}$ bzw. $\mathcal{F}_{\mathrm{an}}\backslash\mathcal{F}_{\mathrm{an,c}}$ liegen. □

Dieser Satz erlaubt es, den Bereich für die Suche einer optimalen Lösung des VER-Problems, also die zulässige Menge $\mathcal{F}_{\mathrm{c}}$, einzuschränken. Hierbei sei bemerkt, dass die Menge $\mathcal{F}_{\mathrm{an,c}}$ sowohl Elemente besitzen kann, für die J_{an} den Wert ∞ annimmt, als auch leer sein kann.
Schlussendlich kann damit die Existenz einer optimalen Lösung des VER-Problems gezeigt werden.

Existenz einer optimalen Lösung

Korollar 5.12. *Das analytische VER-Problem (5.11) besitzt mindestens eine optimale Lösung.*

Beweis. Sei $\mathcal{F}_{\mathrm{an,c}} = \emptyset$, so gilt für alle $\hat{f}_{\mathrm{c}} \in \mathcal{F}_{\mathrm{an}} = \mathcal{F}_{\mathrm{an}}\backslash\mathcal{F}_{\mathrm{an,c}}$ sofort $J_{\mathrm{an}}(\hat{f}_{\mathrm{c}}) = \infty$ und jedes $\hat{f}_{\mathrm{c}} \in \mathcal{F}_{\mathrm{an}}$ ist daher optimale Lösung.
Sei $\mathcal{F}_{\mathrm{an,c}}, \mathcal{F}_{\mathrm{an}}\backslash\mathcal{F}_{\mathrm{an,c}} \neq \emptyset$, dann gilt für alle $\hat{f}_{\mathrm{c}} \in \mathcal{F}_{\mathrm{an,c}}$ und alle $\hat{g}_{\mathrm{c}} \in \mathcal{F}_{\mathrm{an}}\backslash\mathcal{F}_{\mathrm{an,c}}$ gerade $J_{\mathrm{an}}(\hat{f}_{\mathrm{c}}) \leq J_{\mathrm{an}}(\hat{g}_{\mathrm{c}}) = \infty$. Des Weiteren ist die Zielfunktion J_{an} nach Satz 5.9 nach unten halbstetig und nimmt, wie in Kapitel 2.1 beschrieben, auf kompakten Mengen ihr Minimum an. Insgesamt existiert damit eine optimale Lösung auf $\mathcal{F}_{\mathrm{an,c}}$. □

An dieser Stelle wird nun ein Beispiel betrachtet, in welchem die Eigenschaften des analytischen VER-Problems zu erkennen sind. Für die Funktion f wird hierbei ein Polynom gewählt, siehe Beispiel 2.7. Polynome sind immer stetig und messbar und eigenen sich besonders für das analytische VER-Problem, da sie mehrere Strukturen zulassen, sofern die entsprechenden Koeffizienten des Polynoms null sind. Ist das Polynom von Grad eins, also eine lineare Funktion, so kann insbesondere jede Struktur s mit $|s| = d$ realisiert werden.

Beispiel 5.13. Für die Designmenge $\Omega_{\mathrm{ds}} = [-2,2] \times [-1{,}5,\ 1{,}5]$, die Outputfunktion $f : \mathbb{R}^2 \to \mathbb{R},\ x \mapsto (x_1)^3 + (x_2)^3$ mit Outputgrenze $f_{\mathrm{c}} = -1$ und Struktur $s = (1,1)$ ist eine optimale Lösung des analytischen VER-Problems (5.11) gesucht.

Die Anzahl der Designkomponenten beträgt in diesem Beispiel $d = 2$ und die Anzahl

der Outputkomponenten $m = 1$. Offensichtlich ist f mit der Struktur s additiv zerlegbar. In Abbildung 5.3 ist die zulässige Designmenge Ω_{f_c} veranschaulicht. Dabei ist zu erkennen, dass Ω_{f_c} nicht konvex ist, obwohl $f^1 : \mathbb{R} \to \mathbb{R},\ x_1 \mapsto (x_1)^3$ und $f^2 : \mathbb{R} \to \mathbb{R},\ x_2 \mapsto (x_2)^3$ quasikonvex sind, was bereits aus Bemerkung 4.11 bekannt ist.
Für die zulässige Menge $\mathcal{F}_{\text{an}}$ des Optimierungsproblems gilt

$$\mathcal{F}_{\text{an}} = \{(f_c^1, f_c^2) \in \mathbb{R}^2 \mid f_c^1 + f_c^2 = -1\}$$

und für die Mengen $\Omega^1_{f_c^1}$ und $\Omega^2_{f_c^2}$, welche in der Zielfunktion J_{an} enthalten sind,

$$\begin{aligned} \Omega^1_{f_c^1} &= \{x_1 \in [-2,2] \mid (x_1)^3 \leq f_c^1\}, \\ \Omega^2_{f_c^2} &= \{x_2 \in [-1{,}5,1{,}5] \mid (x_2)^3 \leq f_c^2\}. \end{aligned}$$

Der Funktionswert der Zielfunktion J_{an} ergibt sich damit zu

$$J_{\text{an}}(f_c^1, f_c^2) = \begin{cases} -\log\left((\sqrt[3]{f_c^1}+2)(\sqrt[3]{f_c^2}+1{,}5)\right) & \text{für } (f_c^1, f_c^2) \in \widetilde{D}_1, \\ -\log\left(4(\sqrt[3]{f_c^2}+1{,}5)\right) & \text{für } (f_c^1, f_c^2) \in \widetilde{D}_2, \\ -\log\left(3(\sqrt[3]{f_c^1}+2)\right) & \text{für } (f_c^1, f_c^2) \in \widetilde{D}_3, \\ -\log(12) & \text{für } (f_c^1, f_c^2) \in \widetilde{D}_4, \\ \infty & \text{sonst.} \end{cases}$$

wobei

$$\begin{aligned} \widetilde{D}_1 &:= (-8,8) \times (-3{,}375, 3{,}375), \\ \widetilde{D}_2 &:= [8,\infty) \times (-3{,}375, 3{,}375), \\ \widetilde{D}_3 &:= (-8,8) \times [3{,}375, \infty), \\ \widetilde{D}_4 &:= [8,\infty) \times [3{,}375, \infty). \end{aligned}$$

In Abbildung 5.4 ist der Graph von J_{an} visualisiert und in Abbildung 5.5 der Graph der Funktion

$$J_{\text{an,p}} : \mathbb{R} \to (-\infty, \infty],\ f_c^2 \mapsto J_{\text{an}}(-1 - f_c^2, f_c^2).$$

Nun wird versucht mithilfe der Optimalitätskriterien aus Kapitel 4.2 mögliche lokale Lösungen des Beispielproblems zu finden. Eine mögliche lokale Lösung mit minimalem Funktionswert der Zielfunktion ist dann eine optimale Lösung. Dazu ist

zu beachten, dass die Zielfunktion nur auf der Menge $D_\mathrm{d} = ((-8,\infty)\backslash\{0,8\}) \times ((-3{,}375,\infty)\backslash\{0,3{,}375\}) \subset \mathbb{R}^2$ differenzierbar ist. Deshalb können mit den Optimalitätskriterien nur auf dieser Menge Kandidaten für lokale Lösungen gefunden werden. Anschließend können die Funktionswerte in Umgebungen von Punkten $(f_\mathrm{c}^1, f_\mathrm{c}^2) \in \mathcal{F}_\mathrm{an}$, in denen J_an nicht differenzierbar ist, mit diesen Kandidaten verglichen werden.
Die KKT-Bedingungen für das Beispielproblem werden nur in $D_\mathrm{d} \cap \widetilde{D}_1$ erfüllt. Es gilt

$$\begin{aligned} -\frac{1}{3\sqrt[3]{f_\mathrm{c}^1}^2\left(\sqrt[3]{f_\mathrm{c}^1}+2\right)} + \lambda &= 0, \\ -\frac{1}{3\sqrt[3]{f_\mathrm{c}^2}^2\left(\sqrt[3]{f_\mathrm{c}^2}+1{,}5\right)} + \lambda &= 0, \\ f_\mathrm{c}^1 + f_\mathrm{c}^2 &= -1 \end{aligned}$$

für $(f_\mathrm{c}^1, f_\mathrm{c}^2, \lambda) \in \{(-1{,}35, 0{,}35, 0{,}31), (-0{,}21, -0{,}79, 0{,}68), (-0{,}09, -1{,}09, 0{,}67)\}$, wobei alle Komponenten der Einträge dieser Menge auf zwei Nachkommastellen gerundet sind. Die sich daraus ergebenden, möglichen Lösungen sind zusammen mit ihren Zielfunktionswerten sowohl in Abbildung 5.4 als auch in Abbildung 5.5 dargestellt. Zusätzlich werden die zugehörigen Mengen $\Omega^1_{f_\mathrm{c}^1} \times \Omega^2_{f_\mathrm{c}^2}$ in Abbildung 5.3 veranschaulicht. Wie oben beschrieben kann insgesamt gezeigt werden, dass der Vektor $(f_\mathrm{c}^{1,*}, f_\mathrm{c}^{2,*}) = (-1{,}35, 0{,}35)$ die eindeutige, optimale Lösung des Beispielproblems ist mit $J_\mathrm{an}(f_\mathrm{c}^{1,*}, f_\mathrm{c}^{2,*}) = -0{,}07$, ebenfalls auf zwei Nachkommastellen gerundet.

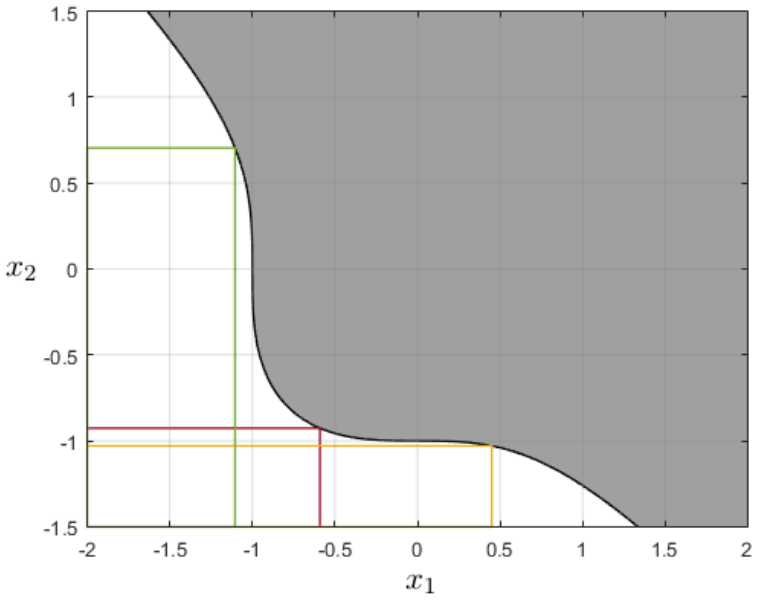


Abbildung 5.3: Zulässige Designmenge Ω_{f_c} (weiß) und zu den möglichen Lösungen gehörigen Mengen $\Omega^1_{f_c^1} \times \Omega^2_{f_c^2}$

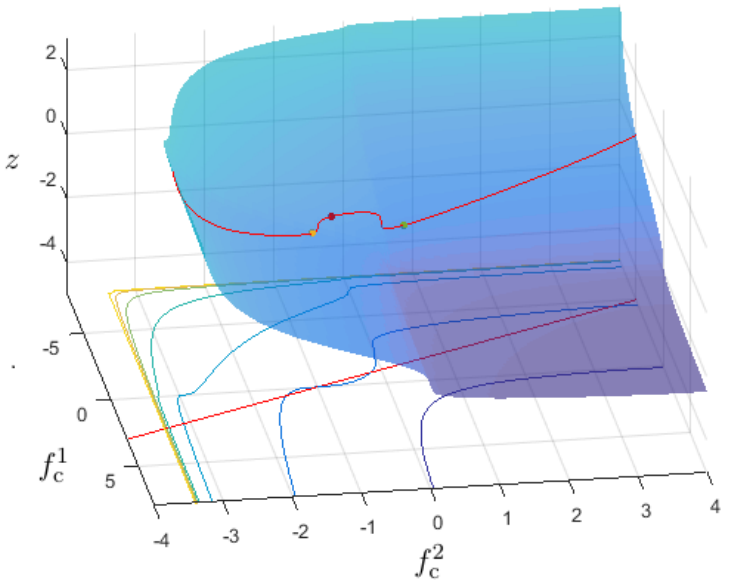


Abbildung 5.4: Graph von J_{an} (Blautöne), Graph von J_{an} auf die zulässige Menge eingeschränkt (rot) und mögliche Lösungen zusammen mit ihren Funktionswerten als Punkte

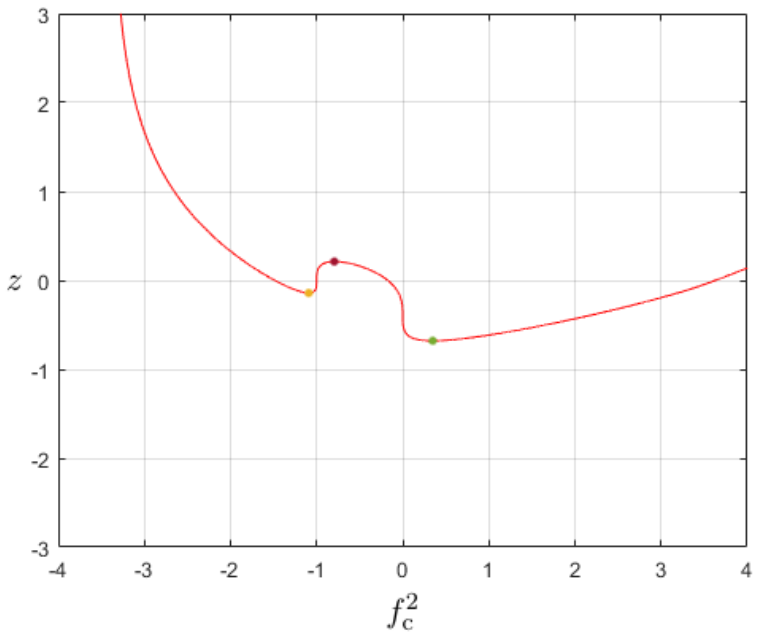


Abbildung 5.5: Graph von $J_{\text{an,p}}$ und parametrisierte mögliche Lösungen mit ihren Funktionswerten als Punkte

Bemerkung 5.14. In Beispiel 5.13 ist erkennbar, dass die Zielfunktion J_{an} des analytischen VER-Problem in der Regel nicht differenzierbar ist. Darüber hinaus ist in diesem Beispiel zu sehen, dass die Aufstellung der Zielfunktion selbst bei handlichen Outputfunktionen recht komplex werden kann und damit auch die Bestimmung des Gradienten der Zielfunktion auf D_{d}.

6. Das konvexe VER-Problem

Nun wird das analytische VER-Problem für konvexe Outputfunktionen betrachtet und gezeigt, dass dies ein konvexes Optimierungsproblem darstellt. Falls die Struktur der Outputfunktion eindimensional ist, gibt es eine alternative Formulierung für das konvexe Problem.

6.1. Allgemeine Strukturen

Satz 6.1. *Sei das analytische VER-Problems (5.11) mit konvexer Outputfunktion f gegeben. Dann ist auch die Zielfunktion J_{an} konvex.*

Beweis. Der Beweis erfolgt in drei Schritten. Zunächst sei jedoch an Satz 4.10 erinnert, welcher begründet, dass für eine konvexe Outputfunktion f auch die Funktionen f^k, $k = 1, \ldots, n$, mit $f(x) = \sum_{k=1}^{n} f^k(x^k)$ für alle $x \in \mathbb{R}^d$ konvex sind.

Schritt 1: Es wird gezeigt, dass für $k \in \{1, \ldots, n\}$ die Funktion χ_{f^k}, definiert über

$$\chi_{f^k} : \mathbb{R}^{d^k} \times \mathbb{R}^m \to \{0,1\},\ (x^k, f_{\mathrm{c}}^k) \mapsto \chi_{f^k}(x^k, f_{\mathrm{c}}^k) := \begin{cases} 1 & \text{falls } f^k(x) \leq f_{\mathrm{c}}^k, \\ 0 & \text{sonst,} \end{cases}$$

logarithmisch-konkav ist, wobei das Ungleichungssystem $f^k(x) \leq f_{\mathrm{c}}^k$ wie gehabt komponentenweise zu verstehen ist. Hierzu reicht es nachzuweisen, dass für alle $(x^k, f_{\mathrm{c}}^k), (y^k, g_{\mathrm{c}}^k) \in \mathbb{R}^{d^k} \times \mathbb{R}^m$ und alle $\theta \in (0,1)$ die Ungleichung

$$\begin{aligned} \chi_{f^k}(\theta x^k + (1-\theta)y^k, \theta f_{\mathrm{c}}^k + (1-\theta) g_{\mathrm{c}}^k) &\geq \chi_{f^k}(x^k, f_{\mathrm{c}}^k)^{\theta}\, \chi_{f^k}(y^k, g_{\mathrm{c}}^k)^{1-\theta} \\ &= \chi_{f^k}(x^k, f_{\mathrm{c}}^k)\, \chi_f(y^k, g_{\mathrm{c}}^k) \end{aligned}$$

erfüllt ist. Für $\chi_{f^k}(x^k, f_{\mathrm{c}}^k) = 0$ oder $\chi_{f^k}(y^k, g_{\mathrm{c}}^k) = 0$ ist dies stets der Fall. Seien daher $\chi_{f^k}(x^k, f_{\mathrm{c}}^k) = \chi_{f^k}(y^k, g_{\mathrm{c}}^k) = 1$, d.h. für $(x^k, f_{\mathrm{c}}^k), (y^k, g_{\mathrm{c}}^k) \in \mathbb{R}^{d^k} \times \mathbb{R}^m$ gilt $f^k(x^k) \leq f_{\mathrm{c}}^k$ und $f^k(y^k) \leq g_{\mathrm{c}}^k$. Mit der Konvexität von f^k folgt daraus

$$\begin{aligned} f^k(\theta x^k + (1-\theta)y^k) &\leq \theta \underbrace{f^k(x^k)}_{\leq f_{\mathrm{c}}^k} + (1-\theta) \underbrace{f^k(y^k)}_{\leq g_{\mathrm{c}}^k} \\ &\leq \theta f_{\mathrm{c}}^k + (1-\theta) g_{\mathrm{c}}^k, \end{aligned}$$

weshalb auch $\chi_{f^k}(\theta x^k + (1-\theta)y^k, \theta f_{\mathrm{c}}^k + (1-\theta) g_{\mathrm{c}}^k) = 1$ gilt. Somit wurde nachgewiesen, dass χ_{f^k} logarithmisch-konkav ist.

Schritt 2: Nun wird für $k \in \{1, \ldots, n\}$ gezeigt, dass die Abbildung

$$\widetilde{J}^k : \mathbb{R}^m \mapsto [0, \infty),\ f_{\mathrm{c}}^k \mapsto \lambda^{d^k}\big(\Omega_{f_{\mathrm{c}}^k}^k\big)$$

logarithmisch-konkav ist. Wegen $\Omega_{f_{\mathrm{c}}^k}^k = \{x^k \in \Omega_{\mathrm{ds}}^k \mid f^k(x^k) \leq f_{\mathrm{c}}^k\}$ gilt

$$\widetilde{J}^k(f_{\mathrm{c}}^k) \quad = \quad \int_{\Omega_{\mathrm{ds}}^k} \chi_{f^k}\big(x^k, f_{\mathrm{c}}^k\big)\, \mathrm{d}x^k$$

für $f_{\mathrm{c}}^k \in \mathbb{R}^m$ mit den Darstellungen aus Kapitel 3.3. Die Menge Ω_{ds}^k ist aufgrund ihrer Gestalt, vgl. (5.6) und Satz 4.3, konvex und liegt in $\mathcal{B}(\mathbb{R}^{d^k})$. Somit sind insgesamt alle Voraussetzungen an Korollar 4.17 erfüllt und $\widetilde{J}^k$ ist folglich logarithmisch-konkav.

Schritt 3: Im letzten Schritt wird die Konvexität der Zielfunktion $\widetilde{J}_{\mathrm{an}}$ nachgewiesen. Es sei

$$\widetilde{J} : \prod_{k=1}^{n} \mathbb{R}^m \to [0, \infty),\ (f_{\mathrm{c}}^1, \ldots, f_{\mathrm{c}}^n) \mapsto \prod_{k=1}^{n} \widetilde{J}^k(f_{\mathrm{c}}^k),$$

wobei im Folgenden wieder $\hat{f}_{\mathrm{c}} = (f_{\mathrm{c}}^1, \ldots, f_{\mathrm{c}}^n)^{\mathrm{T}} \in \mathbb{R}^{mn}$ geschrieben wird. Da die Funktionen $\widetilde{J}^k$, $k = 1, \ldots, n$, logarithmisch konkav sind, gilt für alle $\hat{f}_{\mathrm{c}}, \hat{g}_{\mathrm{c}} \in \mathbb{R}^{mn}$ und alle $\theta \in (0, 1)$

$$\begin{aligned}
\widetilde{J}\big(\theta \hat{f}_{\mathrm{c}} + (1-\theta)\hat{g}_{\mathrm{c}}\big) &= \prod_{k=1}^{n} \widetilde{J}^k(\theta f_{\mathrm{c}}^k + (1-\theta) g_{\mathrm{c}}^k) \\
&\geq \prod_{k=1}^{n} \widetilde{J}^k(f_{\mathrm{c}}^k)^{\theta}\, \widetilde{J}^k(g_{\mathrm{c}}^k)^{1-\theta} \\
&= \left(\prod_{k=1}^{n} \widetilde{J}^k(f_{\mathrm{c}}^k)\right)^{\theta} \left(\prod_{k=1}^{n} \widetilde{J}^k(g_{\mathrm{c}}^k)\right)^{1-\theta} \\
&= \widetilde{J}\big(\hat{f}_{\mathrm{c}}\big)^{\theta} f\big(\hat{g}_{\mathrm{c}}\big)^{1-\theta},
\end{aligned}$$

d.h. $\widetilde{J}$ ist ebenfalls logarithmisch-konkav. Daraus folgt, dass $\log \circ \widetilde{J}$ konkav und

$-\log \circ \widetilde{J}$ konvex ist. Unter Verwendung von Gleichung (3.4) kann

$$\begin{aligned} J_{\mathrm{an}}(\hat{f}_{\mathrm{c}}) &= -\log\left(\lambda^d\left(\prod_{k=1}^{n}\Omega^k_{f^k_{\mathrm{c}}}\right)\right) \\ &= -\log\left(\prod_{k=1}^{n}\lambda^{d^k}\left(\Omega^k_{f^k_{\mathrm{c}}}\right)\right) \\ &= -\log\left(\prod_{k=1}^{n}\widetilde{J}^k(f^k_{\mathrm{c}})\right) \\ &= -\log\left(\widetilde{J}(\hat{f}_{\mathrm{c}})\right) \end{aligned}$$

für $\hat{f}_{\mathrm{c}} \in \mathbb{R}^{mn}$ gezeigt werden, womit die Konvexität von J_{an} nachgewiesen ist. □

Da im Falle einer konvexen Outputfunktion neben der zulässigen Menge auch die Zielfunktion des analytischen VER-Problems konvex ist, ergibt sich ein konvexes Optimierungsproblem.

Definition 6.2. *Seien die Designmenge $\Omega_{\mathrm{ds}} \subset \mathbb{R}^d$ sowie die konvexe Outputfunktion $f : \mathbb{R}^d \to \mathbb{R}^m$ mit Outputgrenze $f_{\mathrm{c}} \in \mathbb{R}^m$ gegeben. Ferner sei f additiv zerlegbar mit Struktur $s \in \mathbb{N}^n$. Das* ***konvexe VER-Problem*** *lautet dann*

$$\begin{aligned} &\underset{f^k_{\mathrm{c}}\in\mathbb{R}^m,\ k=1,\ldots,n}{\text{minimiere}} && -\log\left(\lambda^d\left(\prod_{k=1}^{n}\Omega^k_{f^k_{\mathrm{c}}}\right)\right) \\ &\text{u.d.N.} && \sum_{k=1}^{n} f^k_{\mathrm{c}} = f_{\mathrm{c}}. \end{aligned} \tag{6.1}$$

Da das konvexe VER-Problem (6.1) ein Spezialfall des analytischen VER-Problems (5.11) ist, übertragen sich die weiteren Eigenschaften aus Kapitel 5.2 direkt. Dass die Zielfunktion im konvexen Fall ebenso nicht stetig sein muss, zeigt das Beispiel 6.4.

Beispiele

Beispiel 6.3. Für die Designmenge $\Omega_{\mathrm{ds}} = [-1{,}5, 1{,}5]^3$, die konvexe Outputfunktion $f : \mathbb{R}^3 \to \mathbb{R},\ x \mapsto \|x\|_2^2 = (x_1)^2 + (x_2)^2 + (x_3)^2$ mit Outputgrenze $f_{\mathrm{c}} = 1$ und Struktur $s = (2, 1)$ ist eine optimale Lösung des konvexen VER-Problems (6.1) gesucht.

Die Anzahl der Designkomponenten beträgt in diesem Beispiel $d = 3$ und die Anzahl der Outputkomponenten $m = 1$. Offensichtlich ist f mit der Struktur s additiv zerlegbar, wobei $f^1 : \mathbb{R}^2 \to \mathbb{R},\ (x_1, x_2) \mapsto (x_1)^2 + (x_2)^2$ und $f^2 : \mathbb{R}^2 \to \mathbb{R},\ x_3 \mapsto (x_3)^2$

ist. In Abbildung 6.1 ist die zulässige Designmenge Ω_{f_c} veranschaulicht.
Für die zulässige Menge $\mathcal{F}_{\text{an}}$ des Optimierungsproblems gilt

$$\mathcal{F}_{\text{an}} = \{(f_c^1, f_c^2) \in \mathbb{R}^2 \mid f_c^1 + f_c^2 = 1\}$$

und für die Mengen $\Omega^1_{f_c^1}$ und $\Omega^2_{f_c^2}$, welche in der Zielfunktion J_{an} enthalten sind,

$$\begin{aligned} \Omega^1_{f_c^1} &= \{(x_1, x_2) \in [-1{,}5, 1{,}5]^2 \mid (x_1)^2 + (x_2)^2 \leq f_c^1\}, \\ \Omega^2_{f_c^2} &= \{x_2 \in [-1{,}5, 1{,}5] \mid (x_3)^2 \leq f_c^2\}. \end{aligned}$$

Bei diesem Beispielproblem genügt es die Zielfunktion J_{an} auf $(0,1]^2 \subset \mathbb{R}$ zu betrachten, da es keine Vektoren $(f_c^1, f_c^2) \in \mathcal{F}_{\text{an}} \backslash (0,1]^2$ gibt mit $J_{\text{an}}(f_c^1, f_c^2) < \infty$. Für $(f_c^1, f_c^2) \in (0,1]^2$ gilt $J_{\text{an}}(f_c^1, f_c^2) = -\log(2\pi f_c^1 \sqrt{f_c^2})$. In Abbildung 6.2 ist der Graph von J_{an} visualisiert und in Abbildung 6.3 der Graph der Funktion

$$J_{\text{an,p}} : (0,1] \to (-\infty, \infty], \ f_c^2 \mapsto J_{\text{an}}(1 - f_c^2, f_c^2).$$

Ferner ist J_{an} in $(f_c^1, f_c^2) \in (0,1]^2$ differenzierbar. Deshalb kann mithilfe der Optimalitätskriterien aus Kapitel 4.2 versucht werden eine Lösung zu finden. Erfüllt ein Vektor $(f_c^{1,*}, f_c^{2,*}, \lambda^*) \in (0,1]^2 \times \mathbb{R}$ die KKT-Bedingungen

$$\begin{aligned} -\frac{1}{f_c^{1,*}} + \lambda^* &= 0, \\ -\frac{1}{2f_c^{2,*}} + \lambda^* &= 0, \\ f_c^{1,*} + f_c^{2,*} &= 1 \end{aligned}$$

so ist $(f_c^{1,*}, f_c^{2,*}) \in (0,1]^2$ bereits optimale Lösung des konvexen Beispielproblems. Dies ist gerade für $(f_c^{1,*}, f_c^{2,*}, \lambda^*) = (\frac{2}{3}, \frac{1}{3}, 2) \in (0,1]^2 \times \mathbb{R}$ mit $J_{\text{an}}(f_c^{1,*}, f_c^{2,*}) = -0{,}88$, auf zwei Nachkommastellen gerundet, der Fall. Die eindeutige, optimale Lösung ist zusammen mit ihrem Zielfunktionswert sowohl in Abbildung 5.4 als auch in Abbildung 5.5 dargestellt. Zusätzlich wird die zugehörige Menge $\Omega^1_{f_c^{1,*}} \times \Omega^2_{f_c^{2,*}}$ in Abbildung 5.3 veranschaulicht.

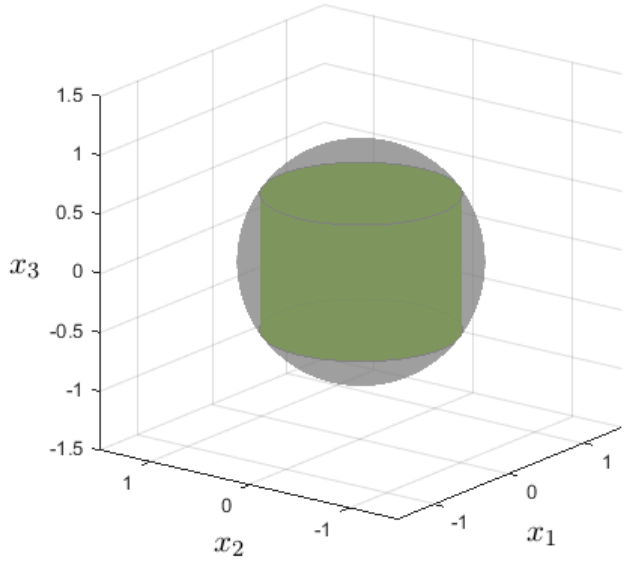


Abbildung 6.1: Zulässige Designmenge Ω_{f_c} (transparent schwarz umrandet) und zur optimalen Lösung gehörige Menge $\Omega^1_{f_c^{1,*}} \times \Omega^2_{f_c^{2,*}}$ (transparent grün umrandet)

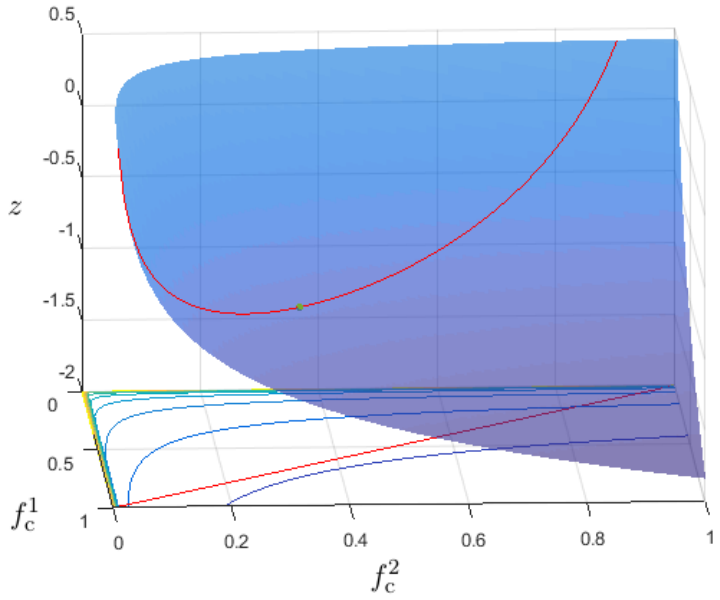


Abbildung 6.2: Graph von J_{an} (Blautöne), Graph von J_{an} auf die zulässige Menge eingeschränkt (rot) und optimale Lösung zusammen mit ihrem Funktionswert als Punkt (grün)

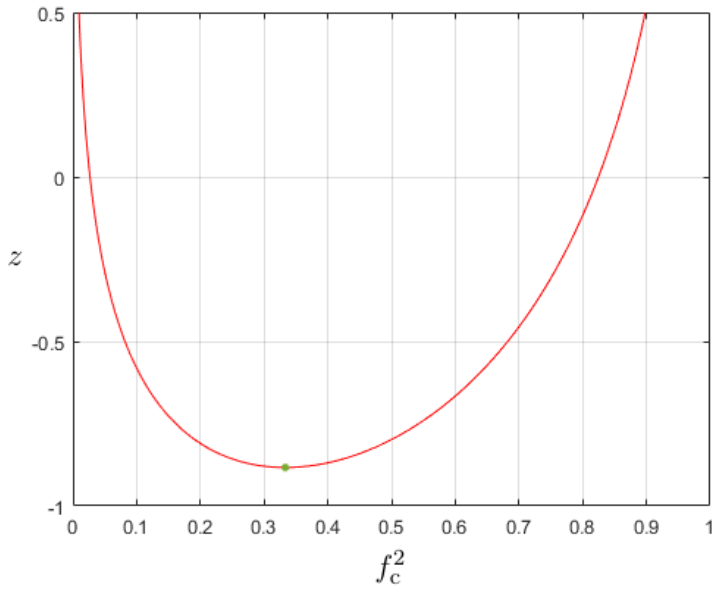


Abbildung 6.3: Graph von $J_{\text{an,p}}$ und parametrisierte optimale Lösung mit ihrem Funktionswert als Punkt

Beispiel 6.4. Für die Designmenge $\Omega_{\text{ds}} = [-1{,}5, 2{,}5] \times [-1{,}5,\ 1{,}5]$, die konvexe Outputfunktion $f = (f_1, f_2)^{\text{T}} : \mathbb{R}^2 \to \mathbb{R}^2,\ x \mapsto ((x_1)^2 + (x_2)^2, (x_1 - 1)^2 + (x_2)^2$ mit Outputgrenze $f_{\text{c}} = (1, 1)^{\text{T}}$ und Struktur $s = (1, 1)$ ist eine optimale Lösung des konvexen VER-Problems (6.1) gesucht.

Die Anzahl der Designkomponenten beträgt in diesem Beispiel $d = 2$ und die Anzahl der Outputkomponenten $m = 2$. Offensichtlich ist f mit der Struktur s additiv zerlegbar und es ist $f^1 = (f_1^1, f_2^1)^{\text{T}} : \mathbb{R}^2 \to \mathbb{R},\ x_1 \mapsto ((x_1)^2, (x_1 - 1)^2)^{\text{T}}$ und $f^2 = (f_1^2, f_2^2)^{\text{T}} : \mathbb{R}^2 \to \mathbb{R},\ x_2 \mapsto ((x_2)^2, (x_2)^2)^{\text{T}}$. In Abbildung 6.4 ist die zulässige Designmenge $\Omega_{f_{\text{c}}}$ veranschaulicht.
Für die zulässige Menge $\mathcal{F}_{\text{an}}$ des Optimierungsproblems gilt

$$\mathcal{F}_{\text{an}} \quad = \quad \{(f_{\text{c}}^1, f_{\text{c}}^2) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid f_{\text{c}}^1 + f_{\text{c}}^2 = (1, 1)^{\text{T}}\}$$

und für die Mengen $\Omega_{f_{\text{c}}^1}^1$ und $\Omega_{f_{\text{c}}^2}^2$, welche in der Zielfunktion J_{an} enthalten sind,

$$\begin{aligned} \Omega_{f_{\text{c}}^1}^1 &= \{x_1 \in [-1{,}5, 2{,}5] \mid ((x_1)^2, (x_1 - 1)^2)^{\text{T}} \leq f_{\text{c}}^1\}, \\ \Omega_{f_{\text{c}}^2}^2 &= \{x_2 \in [-1{,}5, 1{,}5] \mid ((x_2)^2, (x_2)^2)^{\text{T}} \leq f_{\text{c}}^2\}. \end{aligned}$$

In diesem Beispiel genügt es die Zielfunktion J_{an} auf $(0, 1]^2 \times (0, 1]^2 \subset \mathbb{R}^2 \times \mathbb{R}^2$

zu betrachten, da es keine Vektoren $(f_c^1, f_c^2) \in \mathcal{F}_{an} \backslash ((0,1]^2 \times (0,1]^2)$ gibt mit der Eigenschaft $J_{an}(f_c^1, f_c^2) < \infty$. Für $(f_c^1, f_c^2) \in (0,1]^2 \times (0,1]^2$ gilt

$$J_{an}(f_c^1, f_c^2) = \begin{cases} -\log\left(2\left(\sqrt{f_{c,1}^1} + \sqrt{f_{c,1}^2} - 1\right)\sqrt{f_{c,2}^1}\right) & \text{für } (f_c^1, f_c^2) \in \widetilde{D}_1, \\ -\log\left(2\left(\sqrt{f_{c,1}^1} + \sqrt{f_{c,1}^2} - 1\right)\sqrt{f_{c,2}^2}\right) & \text{für } (f_c^1, f_c^2) \in \widetilde{D}_2, \end{cases}$$

wobei

$$\begin{aligned} \widetilde{D}_1 &:= \{(f_c^1, f_c^2) \in (0,1]^2 \times (0,1]^2 \mid f_{c,2}^1 \leq f_{c,2}^2\}, \\ \widetilde{D}_2 &:= \{(f_c^1, f_c^2) \in (0,1]^2 \times (0,1]^2 \mid f_{c,2}^1 > f_{c,2}^2\}. \end{aligned}$$

In Abbildung 6.5 ist der Graph der Funktion

$$J_{an,p} : (0,1] \to (-\infty, \infty],\ f_c^2 \mapsto J_{an}((1,1)^T - f_c^2, f_c^2).$$

visualisiert. Die Zielfunktion J_{an} ist insbesondere für $f_{c,2}^1 = f_{c,2}^2$ nicht differenzierbar. Bei diesem Beispiel kann gezeigt werden, dass die KKT-Bedingungen im differenzierbaren Bereich nicht erfüllt werden können. Die eindeutige, optimale Lösung des Problems beträgt auf zwei Nachkommastellen gerundet $f_c^{1,*} = (0{,}71, 0{,}71)^T$, $f_c^{2,*} = (0{,}21, 0{,}21)^T$. Diese kann später direkt aus Beispiel 6.9 abgeleitet werden. Die zugehörige Menge $\Omega^1_{f_c^{1,*}} \times \Omega^2_{f_c^{2,*}}$ ist in Abbildung 6.4 veranschaulicht.

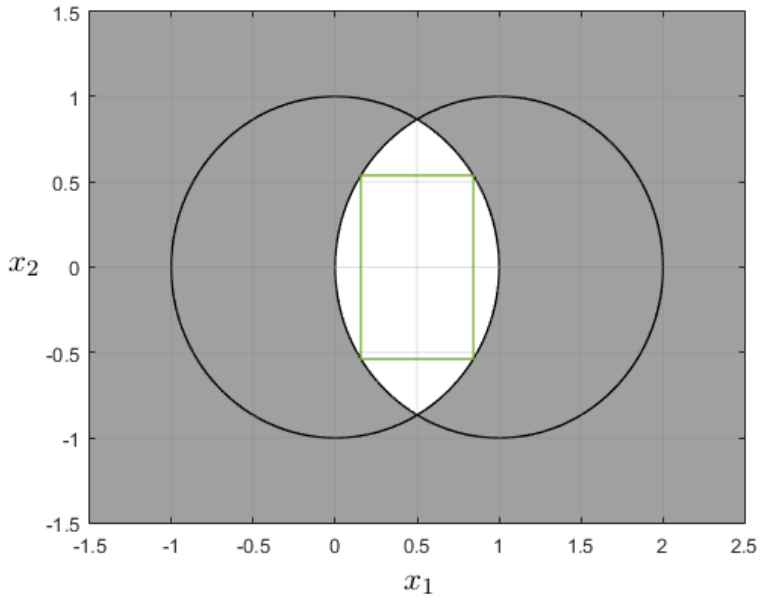


Abbildung 6.4: Zulässige Designmenge Ω_{f_c} (weiß) und zur optimalen Lösung gehörige Menge $\Omega^1_{f_c^{1,*}} \times \Omega^2_{f_c^{2,*}}$ (grün umrandet)

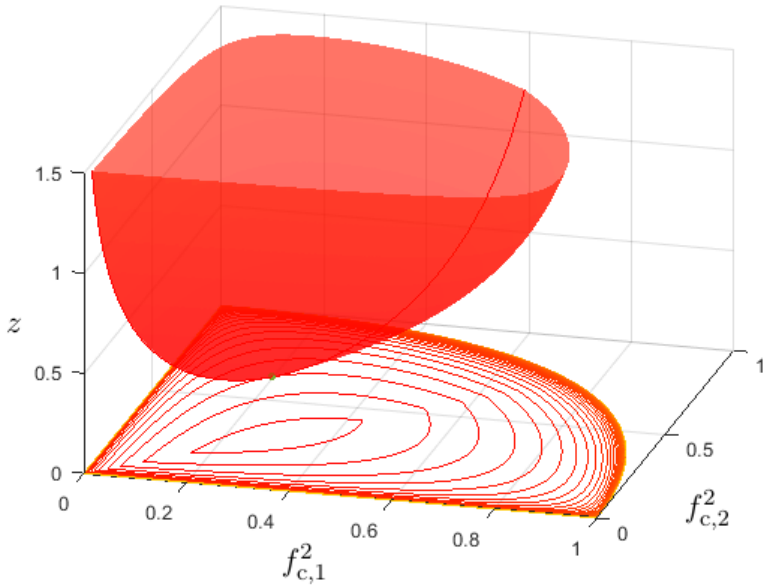


Abbildung 6.5: Graph von $J_{\text{an,p}}$ und parametrisierte optimale Lösung mit ihrem Funktionswert als Punkt

Auch in Beispiel 6.4 ist zu erkennen, dass die Zielfunktion auch bei konvexen Outputfunktionen nicht zwingend differenzierbar ist. Bei eindimensionalen Strukturen ergibt sich eine zur optimalen Lösung gehörige Menge stets in Form eines d-dimensionalen Hyperrechtecks, vgl. Beispiel 2.19. Deshalb wird an dieser Stelle noch kurz auf eine äquivalente Formulierung des konvexen VER-Problems (6.1) eingegangen, bei welcher die Zielfunktion differenzierbar ist.

6.2. Eindimensionale Strukturen

In den folgenden Betrachtungen sei die Outputfunktion f stets konvex und additiv zerlegbar mit eindimensionaler Struktur, d.h. $s = (1, \ldots, 1) \in \mathbb{N}^d$ und $n = d$. Für die Vektoren $f_{\mathrm{c}}^i \in \mathbb{R}^m$, $i = 1, \ldots, d$, kann das Lebesgue-Maß der zugehörigen Mengen $\Omega^i_{f_{\mathrm{c}}^i} \subset \mathbb{R}$ über $l_i := \lambda^1(\Omega^i_{f_{\mathrm{c}}^i}) \in [0, \infty)$ definiert werden. Aufgrund der Konvexität und somit Stetigkeit der Outputfunktion sind alle $\Omega^i_{f_{\mathrm{c}}^i}$ als Schnitt abgeschlossener und konvexer Mengen, vgl. Gleichung (5.5), abgeschlossen und konvex. Daher können sie im Fall $\lambda^1(\Omega^i_{f_{\mathrm{c}}^i}) = 0$ als leere Menge $\emptyset$ oder einelementige Menge $\{x_i^{\mathrm{M}}\}$ und im Fall $\lambda^1(\Omega^i_{f_{\mathrm{c}}^i}) > 0$ als Intervall $[x_i^{\mathrm{M}} - \frac{l_i}{2}, x_i^{\mathrm{M}} + \frac{l_i}{2}]$ mit $x_i^{\mathrm{M}} \in \mathbb{R}$, $i = 1, \ldots, d$, dargestellt werden. Für einen positiven Wert des Maßes muss hierbei

$$\max_{x_i \in [x_i^{\mathrm{M}} - \frac{l_i}{2}, x_i^{\mathrm{M}} + \frac{l_i}{2}]} f_j^i(x_i) \;=\; \max\left(f_j^i\left(x_i^{\mathrm{M}} - \frac{l_i}{2}\right), f_j^i\left(x_i^{\mathrm{M}} + \frac{l_i}{2}\right)\right) \;\leq\; f_{\mathrm{c},j}^i \qquad (6.2)$$

für $j = 1, \ldots, m$ gewährleistet sein, wobei $\max(x, y) := \max\{x, y\}$ für $x, y \in \mathbb{R}$.

Insgesamt gibt es deshalb auch für eine optimale Lösung des konvexen VER-Problems $f_{\mathrm{c}}^{i,*} \in \mathbb{R}^m$, $i = 1, \ldots, d$, welche $\lambda^d\left(\prod_{i=1}^d \Omega^i_{f_{\mathrm{c}}^{i,*}}\right) > 0$ erfüllt, zugehörige $l_i^* \in (0, \infty)$ und $x_i^{\mathrm{M},*} \in \mathbb{R}$ mit $[x_i^{\mathrm{M},*} - \frac{l_i^*}{2}, x_i^{\mathrm{M},*} + \frac{l_i^*}{2}] = \Omega^i_{f^{i,*}} \subset \Omega^i_{\mathrm{ds}}$ für $i = 1, \ldots, d$.
Somit können die entsprechenden Mengen der optimalen Lösung des konvexen VER-Problems auch über das nachfolgende Problem gefunden werden.

Definition 6.5. *Seien die Designmenge $\Omega_{\mathrm{ds}} \subset \mathbb{R}^d$ sowie die konvexe Outputfunktion $f : \mathbb{R}^d \to \mathbb{R}^m$ mit Outputgrenze $f_{\mathrm{c}} \in \mathbb{R}^m$ gegeben. Ferner sei f additiv zerlegbar mit Struktur $s = (1, \ldots, 1) \in \mathbb{N}^d$. Das* ***eindimensionale konvexe VER-Problem***

lautet dann

$$\begin{aligned} &\underset{l_i\in(0,\infty),\ x_i^{\mathrm{M}}\in\mathbb{R},\ i=1,\dots,d}{\text{minimiere}} && -\log\left(\prod_{i=1}^{d} l_i\right) \\ &\text{u.d.N.} && \sum_{i=1}^{d} \max(f^i(x_i^{\mathrm{M}}-\tfrac{l_i}{2}), f^i(x_i^{\mathrm{M}}+\tfrac{l_i}{2})) \le f_{\mathrm{c}}, \\ & && -x_i^{\mathrm{M}}+\tfrac{l_i}{2} \le -x_i^{\mathrm{lb}},\ i=1,\dots,d, \\ & && x_i^{\mathrm{M}}+\tfrac{l_i}{2} \le x_i^{\mathrm{ub}},\ i=1,\dots,d. \end{aligned} \tag{6.3}$$

Dabei stellt die Funktion $\max(\cdot,\cdot):\mathbb{R}^m\times\mathbb{R}^m\to\mathbb{R}^m$ *das komponentenweise Maximum zweier Vektoren in* $\mathbb{R}^m$ *dar, vgl. Gleichung (6.2).*

Wie bei den anderen Problemen kann die Zielfunktion hierbei wieder als Funktion

$$J_1:(0,\infty)^d\times\mathbb{R}^d\to\mathbb{R},\ (l_1,\dots,l_d,x_1^{\mathrm{M}},\dots,x_d^{\mathrm{M}})\mapsto -\log\left(\prod_{i=1}^{d} l_i\right)$$

aufgefasst werden.

Eigenschaften des Problems

Nun wird noch kurz auf die Eigenschaften des eindimensionalen konvexen VER-Problems eingegangen. Dazu werden Elemente in $(0,\infty)^d\times\mathbb{R}^d$ als $2d$-Tupel zusammengefasst, vgl. die Darstellung von $\hat{f}_{\mathrm{c}}$ in Kapitel 5.2. An dieser Stelle wird die Schreibweise $\hat{l}:=(l_1,\dots,l_d,x_1^{\mathrm{M}},\dots,x_d^{\mathrm{M}})\in(0,\infty)^d\times\mathbb{R}^d\subset\mathbb{R}^{2d}$ benutzt.

Satz 6.6. *Die Zielfunktion* J_1 *des eindimensionalen konvexen VER-Problems (6.3) ist*

(i) konvex und

(ii) differenzierbar.

Beweis. (i) Zu zeigen ist, dass für alle $\hat{l},\hat{r}\in(0,\infty)^d\times\mathbb{R}^d$ und alle $\theta\in(0,1)$ die Ungleichung

$$J_1\big(\theta\hat{l}+(1-\theta)\hat{r}\big) \le \theta J_1\big(\hat{l}\big)+(1-\theta)J_{\mathrm{d}}(\hat{r}\big) \tag{6.4}$$

erfüllt ist, wobei $\hat{l}:=(l_1,\dots,l_d,x_1^{\mathrm{M}},\dots,x_d^{\mathrm{M}})$ und $\hat{r}:=(r_1,\dots,r_d,y_1^{\mathrm{M}},\dots,y_d^{\mathrm{M}})$. Die Logarithmusfunktion log ist wegen $\frac{\mathrm{d}^2}{\mathrm{d}x^2}\log(x)=-\frac{1}{x^2}$ für alle $x\in(0,\infty)$

konkav. Daher gilt für alle $l_i, r_i \in (0,\infty)$, $i = 1,\dots,d$ und alle $\theta \in (0,1)$ gerade

$$\log(\theta l_i + (1-\theta) r_i)) \quad \geq \quad \theta \log(l_i) + (1-\theta)\log(r_i)$$

und somit auch

$$\begin{aligned} \sum_{i=1}^{d} \log(\theta l_i + (1-\theta) r_i)) &\geq \sum_{i=1}^{d} \left(\theta \log(l_i) + (1-\theta)\log r_i\right), \\ \log\left(\prod_{i=1}^{d} (\theta l_i + (1-\theta) r_i)\right) &\geq \theta \log\left(\prod_{i=1}^{d} l_i\right) + (1-\theta)\log\left(\prod_{i=1}^{d} r_i\right). \end{aligned}$$

Daraus folgt direkt die Ungleichung (6.4) für alle $\hat{l}, \hat{r} \in (0,\infty)^d \times \mathbb{R}^d$ und für alle $\theta \in (0,1)$.

(ii) Für alle $\hat{l} \in (0,\infty)^d \times \mathbb{R}^d$ gilt

$$\begin{aligned} \nabla J_1(\hat{l}) &= \nabla J_{\mathrm{d}}(l_1,\dots,l_d,x_1^{\mathrm{M}},\dots,x_d^{\mathrm{M}}) \\ &= \left(-\frac{1}{l_1},\dots,-\frac{1}{l_d},0,\dots,0\right)^{\mathrm{T}} \in \mathbb{R}^{2d}. \end{aligned}$$

Damit ist J_1 differenzierbar.

□

Satz 6.7. *Die zulässige Menge*

$$\begin{aligned} \mathcal{F}_1 = \Bigg\{ \hat{l} \in (0,\infty)^d \times \mathbb{R}^d \;\Big|\; \sum_{i=1}^{d} \max\left(f^i\left(x_i^{\mathrm{M}} - \frac{l_i}{2}\right), f^i\left(x_i^{\mathrm{M}} + \frac{l_i}{2}\right)\right) \leq f_{\mathrm{c}}, \\ - x_i^{\mathrm{M}} + \frac{l_i}{2} \leq -x_i^{\mathrm{lb}},\ x_i^{\mathrm{M}} + \frac{l_i}{2} \leq x_i^{\mathrm{ub}},\ i = 1,\dots,d \Bigg\} \end{aligned}$$

des eindimensionalen konvexen VER-Problems (6.3) ist konvex.

Beweis. Es reicht zu zeigen, dass die zugrundeliegenden Funktionen der Nebenbedingungen konvex sind, dann ist $\mathcal{F}_1$ nach Satz 4.8 konvex.
Die Funktionen, welche die Bedingungen $-x_i^{\mathrm{M}} + \frac{l_i}{2} \leq -x_i^{\mathrm{lb}}$,$x_i^{\mathrm{M}} + \frac{l_i}{2} \leq x_i^{\mathrm{ub}}$, $i = 1,\dots,d$ beschreiben, sind linear und daher konvex.
Für $\hat{l} = (l_1,\dots,l_d,x_1^{\mathrm{M}},\dots,x_d^{\mathrm{M}}) \in (0,\infty)^d \times \mathbb{R}^d$ sei $x_i^{\mathrm{l}} := x_i^{\mathrm{M}} - \frac{l_i}{2}$, $x_i^{\mathrm{u}} := x_i^{\mathrm{M}} + \frac{l_i}{2}$ und für $\hat{r} = (r_1,\dots,r_d,y_1^{\mathrm{M}},\dots,y_d^{\mathrm{M}}) \in (0,\infty)^d \times \mathbb{R}^d$ entsprechend $y_i^{\mathrm{l}} := y_i^{\mathrm{M}} - \frac{r_i}{2}$,

$y_i^{\mathrm{u}} := y_i^{\mathrm{M}} + \frac{r_i}{2}$, $i = 1, \ldots, d$. Dann gilt für alle $\hat{l}, \hat{r} \in (0, \infty)^d \times \mathbb{R}^d$ und alle $\theta \in (0, 1)$ weiterhin

$$\begin{aligned}
&\sum_{i=1}^{d} \max\Big(f_j^i\big(\theta x_i^{\mathrm{l}} + (1-\theta) y_i^{\mathrm{l}}\big), f_j^i\big(\theta(x_i^{\mathrm{u}} + (1-\theta) y_i^{\mathrm{u}}\big)\Big) \\
&\qquad \leq \sum_{i=1}^{d} \max\Big(\theta f_j^i\big(x_i^{\mathrm{l}}\big) + (1-\theta) f_j^i\big(y_i^{\mathrm{l}}\big), \theta f_j^i\big(x_i^{\mathrm{u}}\big) + (1-\theta) f_j^i\big(y_i^{\mathrm{u}}\big)\Big) \\
&\qquad \leq \sum_{i=1}^{d} \Big(\theta \max\Big(f_j^i\big(x_i^{\mathrm{l}}\big), f_j^i\big(x_i^{\mathrm{u}}\big)\Big) + (1-\theta) \max\Big(f_j^i\big(y_i^{\mathrm{l}}\big), f_j^i\big(y_i^{\mathrm{u}}\big)\Big)\Big) \\
&\qquad = \theta \sum_{i=1}^{d} \max\Big(f_j^i\big(x_i^{\mathrm{l}}\big), f_j^i\big(x_i^{\mathrm{u}}\big)\Big) + (1-\theta) \sum_{i=1}^{d} \max\Big(f_j^i\big(y_i^{\mathrm{l}}\big), f_j^i\big(y_i^{\mathrm{u}}\big)\Big),
\end{aligned}$$

$j = 1, \ldots, m$. Somit ist gezeigt, dass die zulässige Menge $\mathcal{F}_1$ konvex ist. □

Aufgrund dieser Eigenschaften ist das eindimensionale konvexe VER-Problem (6.3) ebenfalls ein konvexes Optimierungsproblem. Im Unterschied zum konvexen VER-Problem (6.1) ist hierbei die Zielfunktion stets differenzierbar. Allerdings ist die Funktion $\max(\cdot, \cdot)$, welche in den Nebenbedingungen des eindimensionalen konvexen VER-Problems auftritt, nicht differenzierbar.

Beim konvexen VER-Problem (6.1) mit eindimensionaler Struktur liegen die Variablen der Zielfunktion im $\mathbb{R}^{md}$ und die Anzahl der Nebenbedingungen beträgt m. Vergleichsweise dazu sind die Variablen der Zielfunktion des eindimensionalen konvexen VER-Problems (6.3) im $\mathbb{R}^{2d}$ und die Anzahl der Nebenbedingungen beträgt $m + 3d$. Unter Verwendung eines identischen Optimierungsalgorithmus kann somit ein geringerer Rechenaufwand bei der Berechnung einer optimalen Lösung mithilfe des eindimensionalen konvexen VER-Problems erst ab $m > 2$ erzielt werden.

Bemerkung 6.8. Differenzierbarkeit in den Nebenbedingungen beim eindimensionalen konvexen VER-Problem (6.3) könnte erreicht werden, indem anstatt des Terms $\max(f^i(x_i^{\mathrm{M}} - \frac{l_i}{2}), f^i(x_i^{\mathrm{M}} + \frac{l_i}{2}))$ alle möglichen Kombinationen der $f^i(x_i^{\mathrm{M}} \pm \frac{l_i}{2})$ mit $i = 1, \ldots, d$ untersucht werden würden. D.h. anstatt m Ungleichungen müssten dann $m \cdot 2^d$ Ungleichungen erfüllt sein. Da in diesem Fall die Anzahl der Ungleichungen exponentiell mit d wächst, wird dieses Vorgehen im Folgenden nicht weiter verfolgt.

Nachfolgend wird das Beispiel 6.4 erneut aufgegriffen und als eindimensionales konvexes VER-Problem dargestellt. Hierbei können die Nebenbedingungen so vereinfacht werden, dass die zugrundeliegende Funktionen allesamt differenzierbar sind.

Beispiel 6.9. Für die Designmenge $\Omega_{\mathrm{ds}} = [-1{,}5, 2{,}5] \times [-1{,}5,\ 1{,}5]$ und die konvexe Outputfunktion $f = (f_1, f_2)^{\mathrm{T}} : \mathbb{R}^2 \to \mathbb{R}^2,\ x \mapsto ((x_1)^2 + (x_2)^2, (x_1 - 1)^2 + (x_2)^2$ mit Outputgrenze $f_{\mathrm{c}} = (1, 1)^{\mathrm{T}}$ ist eine optimale Lösung des eindimensionalen konvexen VER-Problems (6.3) gesucht.

Die Nebenbedingungen zu diesem Beispielproblem lauten

$$\begin{aligned}
\max\left(\left(x_1^{\mathrm{M}} - \tfrac{l_1}{2}\right)^2, \left(x_1^{\mathrm{M}} + \tfrac{l_1}{2}\right)^2\right) + \max\left(\left(x_2^{\mathrm{M}} - \tfrac{l_2}{2}\right)^2, \left(x_2^{\mathrm{M}} + \tfrac{l_2}{2}\right)^2\right) &\leq 1, \\
\max\left(\left(x_1^{\mathrm{M}} - \tfrac{l_1}{2} - 1\right)^2, \left(x_1^{\mathrm{M}} + \tfrac{l_1}{2} - 1\right)^2\right) + \max\left(\left(x_2^{\mathrm{M}} - \tfrac{l_2}{2}\right)^2, \left(x_2^{\mathrm{M}} + \tfrac{l_2}{2}\right)^2\right) &\leq 1, \\
-x_1^{\mathrm{M}} + \tfrac{l_1}{2} &\leq 1{,}5, \\
-x_2^{\mathrm{M}} + \tfrac{l_2}{2} &\leq 1{,}5, \\
x_1^{\mathrm{M}} + \tfrac{l_1}{2} &\leq 2{,}5, \\
x_2^{\mathrm{M}} + \tfrac{l_2}{2} &\leq 1{,}5
\end{aligned}$$

für $(l_1, l_2, x_1^{\mathrm{M}}, x_2^{\mathrm{M}}) \in (0, \infty)^2 \times \mathbb{R}^2 \subset \mathbb{R}^{2d}$. Aus Symmetriegründen muss für eine optimale Lösung $(x_1^{\mathrm{M},*}, x_2^{\mathrm{M},*}) = (0{,}5, 0)$ gelten. Zusammen mit der Eigenschaft $\{x \in \mathbb{R}^2 \mid f(x) \leq f_{\mathrm{c}}\} = \Omega_{f_{\mathrm{c}}}$ reduzieren sich die Nebenbedingungen des Optimierungsproblems zu

$$\left(\frac{l_1}{2} + 0{,}5\right)^2 + \left(\frac{l_2}{2}\right)^2 = 1.$$

Die KKT-Bedingungen

$$\begin{aligned}
l_2^* + \lambda\left(\frac{l_1^*}{2} + 0{,}5\right) &= 0, \\
l_1^* + \lambda\frac{l_2^*}{2} &= 0, \\
\left(\frac{l_1^*}{2} + 0{,}5\right)^2 + \left(\frac{l_2^*}{2}\right)^2 &= 1
\end{aligned}$$

sind dann für $(l_1^*, l_2^*, \lambda^*) = (0{,}69, 1{,}08, -1{,}28)$ erfüllt, wobei die Einträge auf zwei Nachkommastellen gerundet wurden. Deshalb ist $(0{,}69, 1{,}08, 0{,}5, 0) \in (0, \infty)^2 \times \mathbb{R}^2$ eine optimale Lösung des Beispielproblems. Damit kann die optimale Lösung des zugehörigen analytischen Problems aus Beispiel 6.4 zurückgerechnet werden.

7. Numerische Lösung des konvexen VER-Problems

In diesem Kapitel wird vorgestellt, wie das konvexe VER-Problem mit ein- oder zweidimensionaler Struktur numerisch gelöst werden kann. Dafür wird in dieser Arbeit die Software Matlab® verwendet. Im Vordergrund steht hierbei das Gesamtvorgehen um eine numerische Lösung zu erhalten. Daher werden einige Aspekte nur grob behandelt und mathematische Aussagen nicht detailliert begründet.

7.1. Grundlegendes

Um Optimierungsprobleme der Form (4.10) zu lösen, gibt es in Matlab® implementierte Algorithmen. Für den Fall, dass die zugrundeliegenden Funktionen des Problems (4.10) stetig differenzierbar sind, eignet sich beispielsweise ein sogenanntes Innere-Punkte-Verfahren, welches über den Befehl „`fmincon`“ aufgerufen werden kann und sich durch eine relativ schnelle Berechnung der Lösung auszeichnet. Falls die entsprechenden Ableitungen nicht existieren, ist ein direktes Suchverfahren hilfreich, was beispielsweise in der Matlab®-Routine „`patternsearch`“ hinterlegt ist und mehr Zeit für die Lösungsfindung benötigt. Im Folgenden werden die beiden erwähnten Verfahren kurz vorgestellt.

Innere-Punkte-Verfahren mit `fmincon`

Der Innere-Punkte-Algorithmus der Routine `fmincon` versucht das Minimierungsproblem (4.10) mit $D = \mathbb{R}^d$ zu lösen, indem es eine Folge approximierter Minimierungsprobleme löst, die jeweils durch den Parameter σ beschrieben werden. Für $\sigma \in (0, \infty)$ lautet das **approximierte Minimierungsproblem**

$$\begin{aligned} &\underset{(x,s)\in\mathbb{R}^d\times(0,\infty)^m}{\text{minimiere}} && J(x) - \sigma \sum_{j=1}^{m} \log(s_j) \\ &\text{u.d.N.} && e(x) = e_{\mathrm{c}},\ f(x) + s = f_{\mathrm{c}}. \end{aligned} \tag{7.1}$$

Die Variable $s \in (0, \infty)^m$ heißt „Schlupfvariable“ und der hinzugefügte Term der Zielfunktion „logarithmische Barriere“, vgl. [11]. Hiermit wird dafür gesorgt, dass die Ungleichungsbedingung $f(x) \leq f_{\mathrm{c}}$ für $x \in \mathbb{R}^d$ strikt eingehalten wird, was die Bezeichnung „innere Punkte“ motiviert. Kann $s \in (0, \infty)^m$ garantiert werden, so enthält das approximierte Problem keine Ungleichungsbedingung mehr. Dadurch vereinfachen sich die KKT-Bedingungen (4.13)-(4.17). Für $\sigma \to 0$ wird erwartet, dass die Lösung des approximierten Problems (7.1) gegen die Lösung von (4.10) konvergiert.

Ausgehend von einem Startwert $(x_0, s_0) \in \mathbb{R}^d \times (0, \infty)^m$ versucht der Matlab®-Algorithmus zu jeder Iteration ein Newton-Schritt $(\Delta x, \Delta s)$ durchzuführen, um das approximative Problem zu lösen. Dabei werden die entsprechenden Gleichungen der KKT-Bedingungen des approximierten Minimierungsproblems (7.1) mithilfe linearer Approximation gelöst. Für konvexe Probleme ist der Newton-Schritt in der Regel durchführbar. Falls dies nicht der Fall ist, wird ein sogenannter CG-Schritt getestet, siehe [17] für weitere Informationen. Der Algorithmus bricht insbesondere dann ab, wenn die Optimalitätskriterien erster Ordnung bis auf eine bestimmte Toleranz erfüllt sind. Die Grundlagen des von Matlab® implementierten Innere-Punkte-Verfahrens finden sich in [5].

Direktes Suchverfahren mit `patternsearch`

Der allgemeine Patternsearch-Algorithmus in der Matlab®-Routine `patternsearch` generiert eine Folge von Punkten, welche sich einer Lösung des Minimierungsproblems (4.10) annähert. In jedem Schritt wird auf einer Menge von Punkten, welche in einer Umgebung eines aktuellen Punktes liegen und Netz genannt werden, nach einem Punkt gesucht, dessen Zielfunktionswert geringer ist als der des aktuellen Punktes. Im Erfolgsfall wird ein solcher Punkt dann zum nächsten aktuellen Punkt des Algorithmus. Bei diesem Vorgehen werden keine Ableitungsinformationen der Zielfunktion benötigt, weshalb auch von einer direkten Suchmethode gesprochen wird.
Sofern keine Nebenbedingungen verletzt werden, besteht das Netz aus $2d$ Punkten, welche entlang der Koordinatenachsen ausgerichtet sind und vom aktuellen Punkt einen gewissen Abstand s, auch Netzgröße genannt, haben. Treten lineare Nebenbedingungen auf, so wird das Netz so modifiziert, dass nur zulässige Punkte für das Netz erlaubt werden. Falls auch nichtlineare Nebenbedingungen vorliegen, wird ein sogenannter erweiterter Lagrangescher Patternsearch-Algorithmus verwendet, der, wie die `fmincon`-Routine, Barrieren in der Zielfunktion benutzt, siehe [15].
Sobald ein Punkt des Netzes gefunden wurde, welcher einen geringeren Wert der Zielfunktion aufweist als der aktuelle Punkt, wird dieser zum neuen aktuellen Punkt und ein neues Netz mit doppelter Netzgröße wird erzeugt. Wenn keiner der Zielfunktionsfunktionswerte der Netzpunkte kleiner ist als der des aktuellen Punktes, wird ein neues Netz um den selben Punkt mit halber Netzgröße erzeugt. Der Algorithmus bricht insbesondere dann ab, wenn s einen kritischen Wert unterschreitet.
Der Patternsearch-Algorithmus von Matlab® funktioniert auch, wenn die Zielfunktion auf Netzpunkten den Wert ∞ annimmt, indem diese Punkte ignoriert werden.

Ist dies für alle Netzpunkte der Fall, so wird ein neues Netz mit halbem Abstand erzeugt. Daher ist es wichtig mit einem Punkt zu starten, dessen Zielfunktionswert kleiner als ∞ ist, worauf im Folgenden näher eingegangen wird.
In [17] sind Details zum Matlab® implementierten Algorithmus zu finden und in [3] theoretische Untersuchungen zu diesem.

An dieser Stelle sei außerdem bemerkt, dass es neben den beiden vorgestellten Algorithmen noch viele weitere gibt, um Optimierungsprobleme zu lösen, siehe [11]. Auch in [2] werden Optimierungsalgorithmen vorgestellt, die insbesondere für konvexe, nichtdifferenzierbare Probleme geeignet sind.

Startwertbestimmung

Für das konvexe VER-Problem (6.1) ist ein Startwert $(f_{\mathrm{c},0}^1, \ldots, f_{\mathrm{c},0}^n) \in \prod_{k=1}^{n} \mathbb{R}^m$ gesucht mit $(f_{\mathrm{c},0}^1, \ldots, f_{\mathrm{c},0}^n) \in \mathcal{F}_{\mathrm{an}}$ und $J_{\mathrm{an}}(f_{\mathrm{c},0}^1, \ldots, f_{\mathrm{c},0}^n) < \infty$. Im Folgenden wird beschrieben, wie ein solcher Startwert gefunden werden kann.

Zunächst wird dazu das Optimierungsproblem

$$\begin{aligned} &\underset{x \in \mathbb{R}^d}{\text{minimiere}} && \|f(x) - f_{\mathrm{c}}\|_\infty \\ &\text{u.d.N.} && -x_i \leq -x_i^{\mathrm{lb}},\ x_i \leq x_i^{\mathrm{ub}},\ i = 1, \ldots, d \end{aligned} \tag{7.2}$$

gelöst. Dieses ist wie in Beispiel 4.20 beschrieben ein konvexes Problem, wobei die Zielfunktion generell nicht differenzierbar ist. Zur Lösung des Problems kann deshalb die Matlab®-Routine `patternsearch` verwendet werden. Analog zu Problem (4.12) könnte dieses Problem auch auf ein Problem transformiert werden, bei welchem die zugrundeliegende Funktionen differenzierbar sind, sofern f differenzierbar ist.
Falls für eine optimale Lösung $x^* \in \mathcal{F}$ von (7.2) nun $\|f(x^*) - f_{\mathrm{c}}\|_\infty > 0$ gilt, so existiert kein $x \in \Omega_{\mathrm{ds}}$ mit $f(x) \leq f_{\mathrm{c}}$. Deshalb hat die zulässige Designmenge $\Omega_{f_{\mathrm{c}}}$ ein Lebesgue-Maß vom Wert 0 und somit auch jedes kartesische Produkt $\prod_{k=1}^{n} \Omega_{f_{\mathrm{c}}^k}^k \subset \Omega_{f_{\mathrm{c}}}$ mit $(f_{\mathrm{c}}^1, \ldots, f_{\mathrm{c}}^n) \in \mathcal{F}_{\mathrm{an}}$.
Im Falle $\|f(x^*) - f_{\mathrm{c}}\|_\infty < 0$ gibt es ein $s = (s_1, \ldots, s_m) \in (0, \infty)^m$ mit $f(x^*) + s = f_{\mathrm{c}}$. Als Startwert eignet sich dann $(f_{\mathrm{c},0}^1, \ldots, f_{\mathrm{c},0}^n) \in \prod_{k=1}^{n} \mathbb{R}^m$ mit $f_{\mathrm{c},0,j}^k = f^k(x^*) + \frac{1}{n} s_j$ für $j = 1, \ldots, m$, $k = 1, \ldots, n$, d.h. $\sum_{k=1}^{n} f_{\mathrm{c},0}^k = f_{\mathrm{c}}$, also $(f_{\mathrm{c},0}^1, \ldots, f_{\mathrm{c},0}^n) \in \mathcal{F}$. Aufgrund der Stetigkeit von f und damit der Funktionen f^k, existieren $r^k \in (0, \infty)$, sodass für alle

$x^k \in B(x^{*,k}, r^k) \cap \Omega^k_{\mathrm{ds}}$ gilt $f^k(x^k) < f^k_{\mathrm{c},0}$, $k = 1, \ldots, n$. Daraus folgt, dass $\prod_{k=1}^{n} \Omega^k_{f^k_{\mathrm{c},0}}$ ein positives Lebesque-Maß hat und damit $J_{\mathrm{an}}(f^1_{\mathrm{c},0}, \ldots, f^n_{\mathrm{c},0}) < \infty$.
Falls $\|f(x^*) - f_{\mathrm{c}}\|_\infty = 0$ gilt, so muss zunächst geprüft werden, ob $\Omega_{f_{\mathrm{c}}}$ eine Nullmenge ist. Wenn dies nicht zutrifft, kann der Startwert für die Optimierung wie im Fall $\|f(x^*) - f_{\mathrm{c}}\|_\infty < 0$ gewählt werden mit $s \in [0, \infty)^m$.

Analog kann über eine Lösung $x^* \in \mathcal{F}$ des Minimierungsproblems (7.2) ein Startwert $(l_{0,1}, \ldots, l_{0,d}, x^{\mathrm{M}}_{0,1}, \ldots, x^{\mathrm{M}}_{0,d}) \in (0, \infty)^d \times \mathbb{R}^d$ für das eindimensionale konvexe VER-Problem (6.3) gefunden werden. Gilt für $x^* \in \mathcal{F}$ gerade $\|f(x^*) - f_{\mathrm{c}}\|_\infty < 0$, mit $-x^*_i < -x^{\mathrm{lb}}_i$ und $x^*_i < x^{\mathrm{ub}}_i$ für $i = 1, \ldots, d$, so existiert aufgrund der Stetigkeit von f ein $l_0 \in (0, \infty)^d$ mit $\sum_{i=1}^{d} \max(f^i(x^*_i - \frac{l_{0,i}}{2}), f^i(x^*_i + \frac{l_{0,i}}{2})) < f_{\mathrm{c}}$, $-x^*_i + \frac{l_{0,i}}{2} < -x^{\mathrm{lb}}_i$ und $x^*_i + \frac{l_{0,i}}{2} < x^{\mathrm{ub}}_i$ für $i = 1, \ldots, d$. Darum kann hier $x^{\mathrm{M}}_0 = x^*$ und l_0 komponentenweise hinreichend klein gewählt werden. Dann ist für das so gewählte $(l_{0,1}, \ldots, l_{0,d}, x^{\mathrm{M}}_{0,1}, \ldots, x^{\mathrm{M}}_{0,d}) \in (0, \infty)^d \times \mathbb{R}^d$ außerdem die Regularitätsbedingung von Slater erfüllt.
Für den Fall $\|f(x^*) - f_{\mathrm{c}}\|_\infty \geq 0$ gelten die Aussagen des vorherigen Absatzes.

Numerische Approximation des Lebesgue-Maßes

Da eine analytische Beschreibung der Zielfunktion des analytischen VER-Problems im Allgemeinen sehr komplex ist, vgl. Bemerkung 5.14, wird in diesem Kapitel vorgestellt, wie sie numerisch für zwei- bzw. eindimensionale Strukturen s berechnet werden kann. Insbesondere wird dazu beschrieben, wie das Lebesgue-Maß einer Menge $\Omega^k_{f^k_{\mathrm{c}}} = \{x^k \in \Omega^k_{\mathrm{ds}} \mid f^k(x^k) \leq f^k_{\mathrm{c}}\} \subset \mathbb{R}^2$ mit $k \in \{1, \ldots, n\}$, das gerade dem Flächeninhalt dieser Menge entspricht, berechnen werden kann. Im Folgenden wird der Index k der Einfachheit halber weggelassen, d.h. es wird ein beliebiger Index $k \in \{1, \ldots, n\}$ mit $d^k = 2$ betrachtet.

In einem ersten Schritt werden zwei äquidistante Gitter, G und $\widetilde{G}$, über $\Omega_{\mathrm{ds}} \subset \mathbb{R}^2$ gelegt. Seien $n_{G,1} \in \mathbb{N}$ die Anzahl der Gitterpunkte des Gitters G in x_1-Richtung und $n_{G,2} \in \mathbb{N}$ die Anzahl der Gitterpunkte in x_2-Richtung. Die beiden Gitter sind definiert über

$$\begin{aligned} G &= \{(x^{\mathrm{lb}}_1 + i_1 h_1, x^{\mathrm{lb}}_2 + i_2 h_2) \mid i_1 \in I_{G,1},\ i_2 \in I_{G,2}\}, \\ \widetilde{G} &= \{(x^{\mathrm{lb}}_1 + i_1 h_1, x^{\mathrm{lb}}_2 + i_2 h_2) \mid i_1 \in I_{\widetilde{G},1},\ i_2 \in I_{\widetilde{G},2}\} \end{aligned}$$

mit Gitterschrittweiten $h_1 := \frac{x_1^{\mathrm{ub}} - x_1^{\mathrm{lb}}}{n_{G,1}}$ und $h_2 := \frac{x_2^{\mathrm{ub}} - x_2^{\mathrm{lb}}}{n_{G,2}}$ sowie Gitterindexmengen $I_{G,1} := \{\frac{1}{2}, \frac{3}{2}, \ldots, n_{\mathrm{G_d},1} - \frac{1}{2}\}$, $I_{G,2} := \{\frac{1}{2}, \frac{3}{2}, \ldots, n_{G,2} - \frac{1}{2}\}$, $I_{\widetilde{G},1} := \{0, 1, \ldots, n_{G,1}\}$ und $I_{\widetilde{G},2} := \{0, 1, \ldots, n_{G,2}\}$. In Abbildung 7.1 sind die beiden Gitter dargestellt.

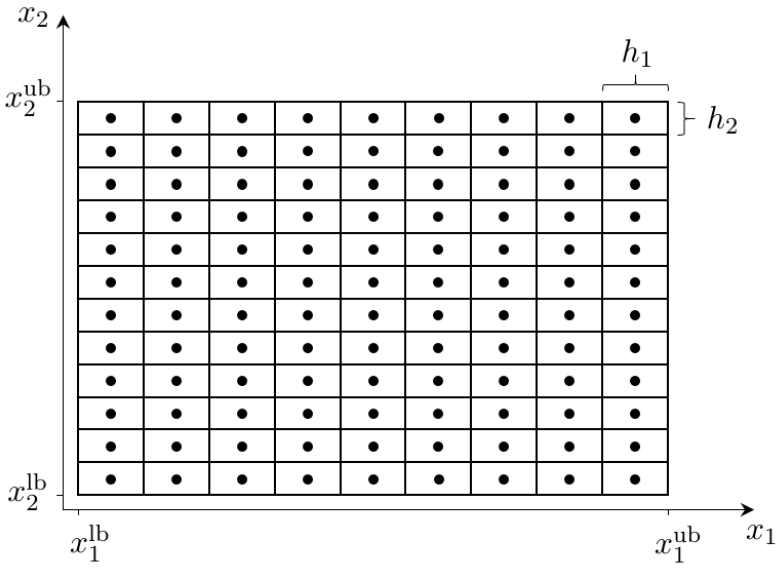


Abbildung 7.1: Darstellung des Gitters G anhand von Punkten und des Gitter $\widetilde{G}$ anhand von Schnittpunkten der durchgezogenen Linien

Im nächsten Schritt wird die Funktion f an jedem Gitterpunkt des Gitters G ausgewertet. Sofern ein Punkt $x_G \in G$ mit $(x_{G,1}, x_{G,2}) := (x_1^{\mathrm{lb}} + i_1 h_1, x_2^{\mathrm{lb}} + i_2 h_2)$, welcher eindeutig über $i_1 \in I_{G,1}$ und $i_2 \in I_{G,2}$ bestimmt ist, das Ungleichungssystem $f(x_G) \leq f_{\mathrm{c}}$ erfüllt, heißt dieser „gültig“. In diesem Fall wird das zugehörige Rechteck R_{x_G}, definiert über

$$R_{x_G} := \{x \in \mathbb{R}^2 \mid x_{G,1} - \frac{h_1}{2} < x_1 < x_{G,1} + \frac{h_1}{2},\ x_{G,2} - \frac{h_2}{2} < x_2 < x_{G,2} + \frac{h_2}{2}\}$$

mit Flächeninhalt $\lambda^2(R_{x_G}) = h_1 h_2$ ebenfalls als gültig erachtet. Das Rechteck R_{x_G} ist offen und wird durch die Gitterpunkte $(x_{G,1} - \frac{h_1}{2}, x_{G,2} - \frac{h_2}{2})$, $(x_{G,1} + \frac{h_1}{2}, x_{G,2} - \frac{h_2}{2})$, $(x_{G,1} + \frac{h_1}{2}, x_{G,2} + \frac{h_2}{2})$ und $(x_{G,1} - \frac{h_1}{2}, x_{G,2} + \frac{h_2}{2})$ des Gitters $\widetilde{G}$ festgelegt. Erfüllt der Gitterpunkt $x_G \in G$ das Ungleichungssystem $f(x_G) \leq f_{\mathrm{c}}$ nicht, dann wird sowohl x_G, als auch das zugehöriges Rechteck R_{x_G} als „ungültig“ erachtet.
Im letzten Schritt lässt sich nun der Flächeninhalt der Menge $\Omega_{f\mathrm{c}}$ über die Anzahl der gültigen Gitterpunkte $n_{G,\mathrm{v}}$, die Gesamtanzahl aller Gitterpunkte und dem

Flächeninhalt von Ω_{ds} bestimmen, d.h. genauer

$$\lambda^2(\Omega_{f_\mathrm{c}}) \quad = \quad \frac{n_{G,\mathrm{v}}}{n_{G,1}n_{G,2}}(x_1^{\mathrm{ub}} - x_1^{\mathrm{lb}})(x_2^{\mathrm{ub}} - x_2^{\mathrm{lb}}) + \varepsilon_G, \tag{7.3}$$

wobei ε_G der Approximationsfehler ist, der bei dieser diskreten Art der Berechnung entsteht und sogleich genauer untersucht wird.
In Abbildung 7.2 wird dieses Verfahren der numerischen Berechnung des Flächeninhalts anhand einer Ellipse, welche über eine Ungleichung der Form $f(x) \leq f_\mathrm{c}$ für $x \in \Omega_{\mathrm{ds}}$ dargestellt werden kann, illustriert.

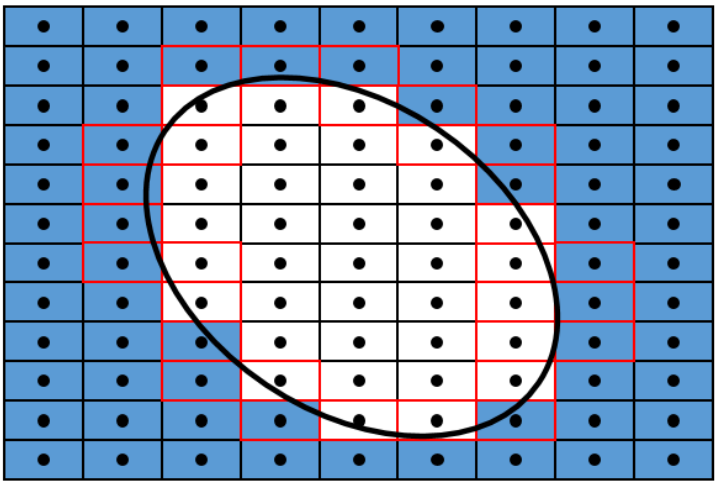


Abbildung 7.2: Zur numerischen Berechnung des Flächeninhalts mit gültigen Rechtecken (weiß), ungültigen Rechtecken (blau), Rechtecken, die zu einem Fehler in der Flächenberechnung führen (rot umrandet)

Der Rand der Menge Ω_{f_c}, welcher mit $\partial\Omega_{f_\mathrm{c}}$ bezeichnet wird, trennt die gültigen von den ungültigen Punkten, wobei alle Punkte auf dem Rand gültig sind. Ein Rechteck R_{x_G} kann also nur dann einen Beitrag zum Approximationsfehler ε_G leisten, wenn $\partial\Omega_{f_\mathrm{c}} \cap R_{x_G} \neq \emptyset$ gilt, d.h. R_{x_G} sowohl gültige als auch ungültige Punkte enthält. Der absolute Anteil von R_{x_G} am Gesamtfehler wird von nun an mit $\Delta\varepsilon_G$ bezeichnet.
Sei $R_{x_G} \cap \partial\Omega_{f_\mathrm{c}} = \partial\Omega_{f_\mathrm{c}}$ für ein $x_G \in G$. Der Approximationsfehler wird maximal, wenn Ω_{f_c} eine Nullmenge mit $\{x_G\} \in \Omega_{f_\mathrm{c}}$ ist, d.h. $\varepsilon_G = \Delta\varepsilon_G = h_1 h_2$. Sei im Folgenden also $R_{x_G} \cap \partial\Omega_{f_\mathrm{c}} \subsetneq \partial\Omega_{f_\mathrm{c}}$ für $x_G \in G$. Der Rand ist daher stets ein eindimensionales Objekt mit Länge $L(\partial\Omega_{f_\mathrm{c}})$, wobei L einer Kurve ihre Länge zuordnet. Der absolute Anteil der Länge des Randes, der auf R_{x_G} fällt, sei nun $\Delta L(\partial\Omega_{f_\mathrm{c}})$. Gesucht ist

dann das maximal mögliche Verhältnis $\frac{\Delta\varepsilon_G}{\Delta L(\partial\Omega_{f_c})}$. Durch Ausloten der verschiedenen Möglichkeiten, wie ein Teil des Randes einer konvexen Menge in R_{x_G} liegen kann, ergibt sich, dass

$$\frac{\Delta\varepsilon_G}{\Delta L(\partial\Omega_{f_c})} \leq \frac{h_1 h_2}{\frac{1}{2}\min\{h_1, h_2\}} = 2\max\{h_1, h_2\}$$

gilt. Dies ist gerade der Fall wenn Ω_{f_c} selbst ein eindimensionales Objekt ist und wie Abbildung 7.3 zu sehen in R_{x_G} liegt, sodass x_G und damit R_{x_G} gültig sind.

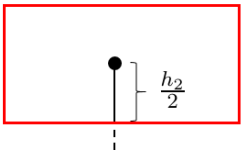


Abbildung 7.3: Beispiel für das Auftreten des maximalen Verhältnisses $\frac{\Delta\varepsilon_G}{\Delta L(\partial\Omega_{f_c})} = 2h_1$ für $h_2 < h_1$

Mithilfe $\frac{\Delta\varepsilon_G}{\Delta L(\partial\Omega_{f_c})} L(\partial\Omega_{f_c})$ kann ε_G nach oben abgeschätzt werden. Es gilt, dass die Länge des Randes einer konvexen Menge stets kleiner als die Länge des Randes einer Menge ist, welche diese einschließt. Dies kann damit begründet werden, dass eine „nächste-Punkte-Projektion“ von einer einschließenden auf eine konvexe Menge kontrahierend ist. Mit $\Omega_{f_c} \subset \Omega_{\mathrm{ds}}$ folgt daher

$$\varepsilon_G \leq 4\max\{h_1, h_2\}\big((x_1^{\mathrm{ub}} - x_1^{\mathrm{lb}}) + (x_2^{\mathrm{ub}} - x_2^{\mathrm{lb}})\big).$$

Mit $h := \max\{h_1, h_2\}$ ergibt sich dann insgesamt $\varepsilon_G = \mathcal{O}(h)$, d.h. es handelt sich um einen Fehler erster Ordnung in h.

Das approximierte Lebesgue-Maß einer Menge $\Omega_{f_c^k}^k \subset \mathbb{R}$ mit $k \in \{1, \ldots, n\}$ lässt sich im Eindimensionalen auf eine ähnliche, aber unkompliziertere Weise bestimmen. Hierbei ist der Fehler ebenfalls erster Ordnung in der Gitterschrittweite h.

Insgesamt können damit nun Rückschlüsse auf den Gesamtfehler bei der Berechnung des Volumens eines n-fachen kartesischen Produkts $\prod_{k=1}^{n} \Omega_{f_c}^k$ mit $\Omega_{f_c}^k \subset \mathbb{R}^{d^k}$ und $d^k \leq 2$ für $k = 1, \ldots, n$, gezogen werden. Wird das Volumen über $\prod_{k=1}^{n} \lambda^{d^k}(\Omega_{f_c}^k)$ bestimmt mit h als größte auftretende Gitterschrittweite für die numerische Berechnung von

$\lambda^{d^k}(\Omega^k_{f_c})$, $k = 1, \ldots, n$, so ist auch der Gesamtfehler von erster Ordnung in h.

Analoge Betrachtungen können auch für ein n-faches kartesisches Produkt $\prod_{k=1}^{n} \Omega^k_{f_c}$ mit $\Omega^k_{f_c} \subset \mathbb{R}^{d^k}$ konvex und $d^1 > 2$ gemacht werden. Es sei bemerkt, dass auch im nicht konvexen Fall die Approximation des Lebesgue-Maßes auf diese Weise erfolgen kann. Allerdings gilt dann die Abschätzung für den Fehler ε_G nicht mehr.

7.2. Beispielprobleme

Nun stehen alle Hilfsmittel bereit, um das konvexe analytische VER-Problem (6.1) für eine ein- bzw. zweidimensionale Struktur numerisch lösen zu können. Als Beispiel wird zunächst das Problem aus Beispiel 6.3 aufgegriffen und die numerische Lösung mit der tatsächlichen, optimalen Lösung verglichen. Anschließend wird ein Praxisbeispiel mit einer realistischen Outputfunktion und realistischen Outputgrenzen betrachtet. Zur numerischen Lösung wird ein Rechner verwendet, dessen Eigenschaften in Anhang A aufgeführt sind. Für die Lösungsfindung wird hierbei außerdem die Gitterschrittweite in allen Dimensionen konstant gewählt.

Numerische Lösung zu Beispiel 6.3

Da für die zulässige Designmenge in Beispiel 6.3 gerade $\Omega_{f_c} \subset [-1,1]^3$ gilt, wird hier $\Omega_{ds} = [-1,1]^3$ als Designmenge gewählt.

Beispiel 7.1. Für die Designmenge $\Omega_{ds} = [-1,1]^3$, die konvexe Outputfunktion $f : \mathbb{R}^3 \to \mathbb{R},\ x \mapsto \|x\|_2^2 = (x_1)^2 + (x_2)^2 + (x_3)^2$ mit Outputgrenze $f_c = 1$ und Struktur $s = (2,1)$ ist eine optimale Lösung des konvexen VER-Problems (6.1) gesucht.

Mit den Methoden aus Kapitel 7.1 kann die Lösung des Problems numerisch bestimmt werden. Dazu wird die Matlab®-Routine `patternsearch` unter Standardeinstellungen verwendet und die Approximation des Lebesgue-Maßes je einmal mit der Gitterschrittweite $h = 0{,}1$, $h = 0{,}01$ und $h = 0{,}001$ durchgeführt. Als Startwert für die Optimierung ergibt sich mit der in Kapitel 7.1 beschriebenen Vorgehensweise $f_{c,0} = (0{,}5, 0{,}5)$. Die zur numerisch berechneten Lösung gehörigen Mengen sind für die drei Varianten der Gitterschrittweite in Abbildung 7.4 dargestellt.
Abbildung 7.5 stellt die Differenz zwischen der zweiten Komponente der numerisch berechneten Lösung und der zweiten Komponente der tatsächlichen Lösung, d.h. $f_c^{*,2}$ aus Beispiel 6.3, dar. Die Differenz in der ersten Komponente ist identisch mit dieser, da für jedes Element $\hat{f}_c \in \mathbb{R}^2$ der zulässigen Menge $f_c^1 + f_c^2 = 1$ gelten muss.

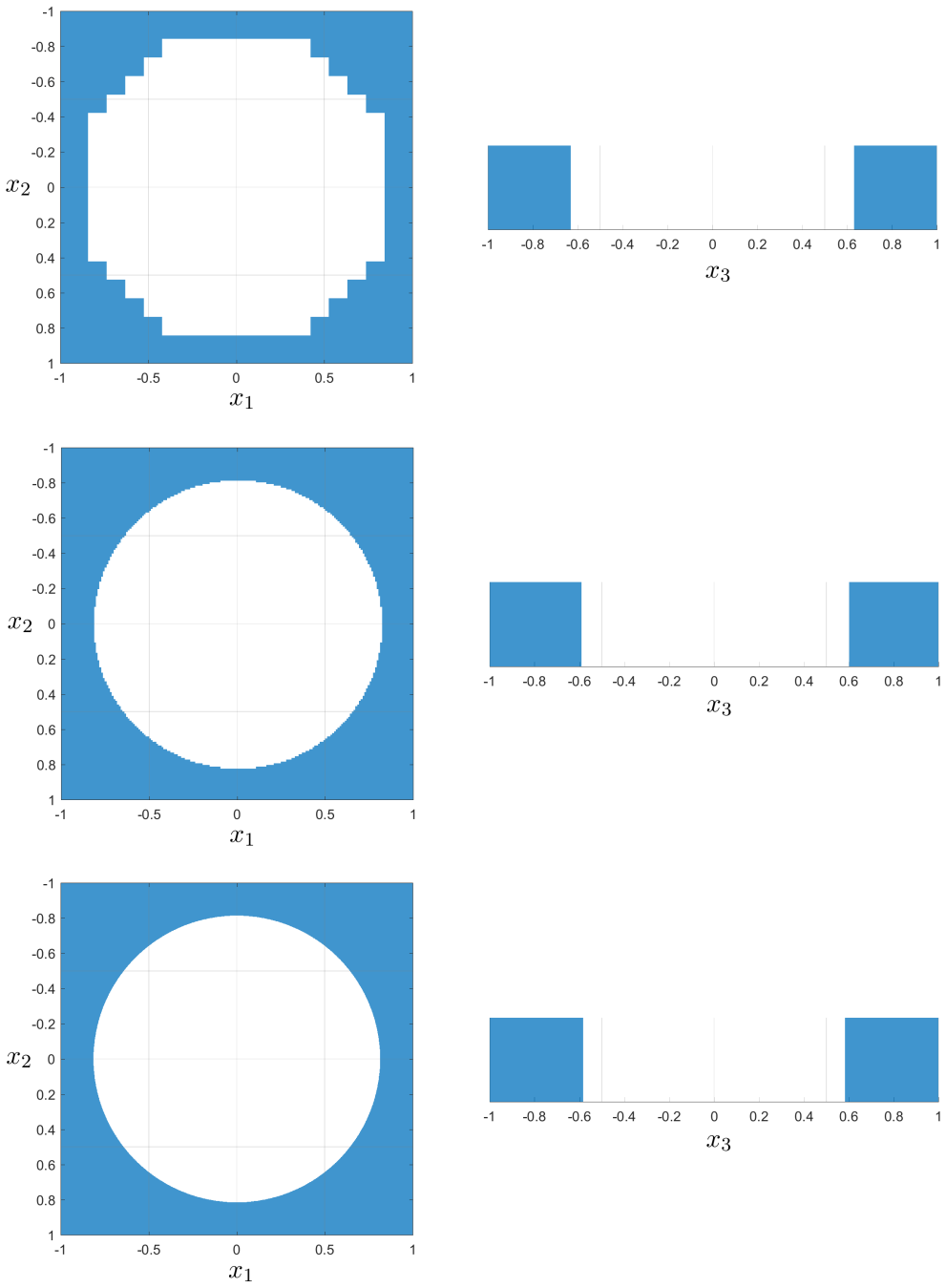


Abbildung 7.4: Zur numerisch berechneten, optimalen Lösung des Beispielproblems gehörige Mengen (weiß) für $h = 0{,}1$ (oben), $h = 0{,}01$ (Mitte) und $h = 0{,}001$ (unten)

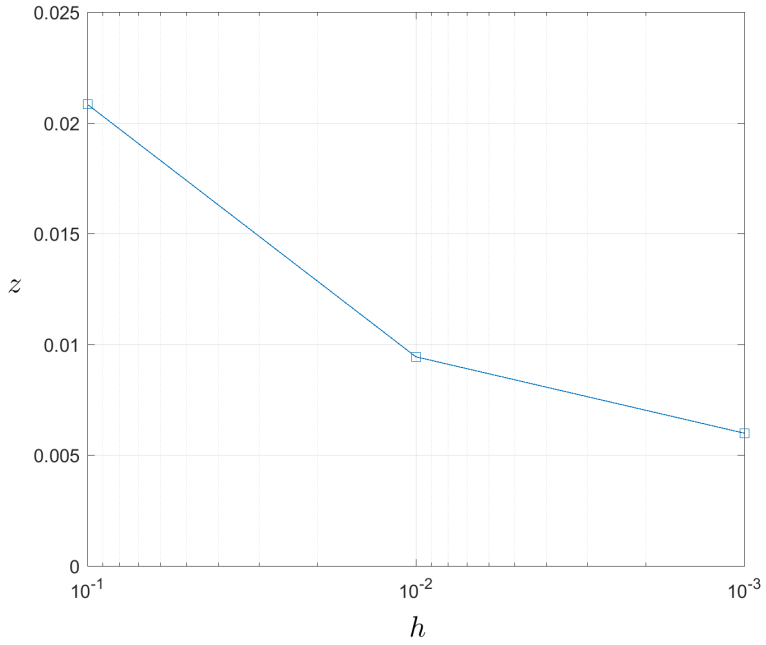


Abbildung 7.5: Differenz zwischen der zweiten Komponente der numerisch berechneten Lösung und der zweiten Komponente der tatsächlichen Lösung auf der z-Achse für die verschiedenen Gitterschrittweiten auf einer logarithmischen Skala der h-Achse

In Abbildung 7.5 ist zu erkennen, dass sich die numerisch berechnete Lösung für eine kleinere Gitterschrittweite der tatsächlichen Lösung annähert. Um auch die Anwendbarkeit der verschiedenen Gitterschrittweiten vergleichen zu können, werden die zur Berechnung benötigten CPU-Zeiten in Tabelle 7.1 angegeben. Außerdem sind dort die zugehörigen Funktionswerte dargestellt. Die Werte sind dabei auf zwei Nachkommastellen gerundet.

h	$J_{\text{an}}(\hat{f})$	CPU-Zeit in s
0,1	-0,95	0,03
0,01	-0,89	0,11
0,001	-0,89	12,50

Tabelle 7.1: Funktionswerte der Zielfunktion und die CPU-Zeiten der Simulation für verschiedene Gitterschrittweiten

Genau wie die numerische Lösung sich der exakten optimalen Lösung für kleiner werdendes h annähert, so nähert sich auch der Zielfunktionswert dem der exakten optimalen Lösung, d.h. $-0{,}88$, siehe Tabelle 7.1. Die Genauigkeit in der numerischen Approximation korreliert jedoch mit der Simulationszeit. Je kleiner die Gitterschrittweite für die Approximation des Lebesgue-Maßes ist, desto größer ist die CPU-Zeit der Simulation.

Praxisbeispiel

Jetzt wird ein Praxisbeispiel mit zwölf Designkomponenten, neun Anforderungen und einer linearen Outputfunktion betrachtet. Dabei ist die Designmenge Ω_{ds} auf $[-1,1]^{12}$ transformiert, d.h. es gilt $x_i^{\text{lb}} = -1$ und $x_i^{\text{ub}} = 1$ für $i = 1, \ldots, d$.

Beispiel 7.2. Für die Designmenge $\Omega_{\text{ds}} = [-1,1]^{12}$, die lineare Outputfunktion $f : \mathbb{R}^{12} \to \mathbb{R}^9,\ x \mapsto Fx$ mit

$$F = \begin{pmatrix}
0{,}00 & 0{,}00 & -0{,}03 & -0{,}08 & 0{,}00 & 0{,}00 & \ldots & \ldots & -0{,}02 & -0{,}05 & -0{,}04 & -0{,}16 & -0{,}24 & -0{,}16 \\
0{,}01 & 0{,}01 & 0{,}72 & 1{,}58 & -0{,}01 & 0{,}03 & \ldots & \ldots & -0{,}29 & -0{,}46 & -0{,}65 & -2{,}22 & -3{,}03 & 2{,}83 \\
-0{,}01 & -0{,}01 & -0{,}72 & -1{,}58 & 0{,}01 & -0{,}03 & \ldots & \ldots & 0{,}29 & 0{,}46 & 0{,}65 & 2{,}22 & 3{,}03 & -2{,}83 \\
0{,}00 & 0{,}00 & -0{,}03 & -0{,}05 & 0{,}00 & 0{,}00 & \ldots & \ldots & 0{,}00 & -0{,}01 & 0{,}02 & 0{,}07 & 0{,}07 & -0{,}12 \\
0{,}62 & 0{,}91 & 0{,}00 & -0{,}04 & -0{,}04 & -0{,}10 & \ldots & \ldots & -0{,}03 & -0{,}09 & -0{,}11 & -0{,}24 & 0{,}02 & 0{,}01 \\
0{,}00 & -0{,}02 & -0{,}02 & -0{,}08 & 0{,}28 & 0{,}40 & \ldots & \ldots & 0{,}03 & 0{,}09 & 0{,}11 & 0{,}24 & -0{,}02 & -0{,}01 \\
-0{,}62 & -0{,}91 & 0{,}00 & 0{,}04 & 0{,}04 & 0{,}10 & \ldots & \ldots & 0{,}00 & -0{,}02 & -0{,}02 & -0{,}06 & 0{,}00 & 0{,}00 \\
0{,}00 & 0{,}02 & 0{,}02 & 0{,}08 & -0{,}28 & -0{,}40 & \ldots & \ldots & 0{,}00 & 0{,}02 & 0{,}02 & 0{,}06 & 0{,}00 & 0{,}00 \\
0{,}09 & -0{,}03 & -1{,}47 & -0{,}91 & -2{,}48 & -2{,}11 & \ldots & \ldots & 0{,}04 & 0{,}15 & 0{,}26 & 0{,}45 & 0{,}02 & 0{,}01
\end{pmatrix} \in \mathbb{R}^{9\times 12}$$

und Outputgrenze $f_{\text{c}} = (-0{,}03, 1{,}52, 3{,}48, 0{,}15, -0{,}24, 0{,}74, -0{,}17, 0{,}67, 2{,}56)^{\text{T}}$ ist eine optimale Lösung des konvexen VER-Problems (6.1) sowohl für eine eindimensionale Struktur als auch für eine zweidimensionale Struktur, deren Einträge alle den

Wert zwei haben, gesucht.

Wie im vorherigen Beispiel wird die Lösung des VER-Problems mit zweidimensionaler Struktur numerisch mit der Matlab®-Routine `patternsearch` unter Standardeinstellungen berechnet. Der Startwert wird dazu wie in 7.1 beschrieben bestimmt und als Gitterschrittweite für die Approximation des Lebesgue-Maßes wird $h = 0.01$ gewählt. Die Lösung, die sich daraus ergibt, ist in Abbildung 7.6 in Form ihrer zugehörigen Mengen dargestellt.
Um die Lösung des Problems mit eindimensionaler Struktur zu erhalten, wird das eindimensionale konvexe VER-Problem (6.3) gelöst. Da die Outpufunktion linear ist, sind die Funktionen, welche den Nebenbedingungen zugrunde liegen, differenzierbar. Numerisch kann das Problem deshalb mit der Matlab®-Routine `fmincon` gelöst werden, wobei der Startwert wieder wie in Kapitel 7.1 beschrieben gewählt wird und für die Matlab®-Routine `fmincon` Standardeinstellungen verwendet werden. Die zur optimalen Lösung gehörigen Mengen mit eindimensionaler Struktur sind in Abbildung 7.6 zusammen mit den zur optimalen Lösung gehörigen Mengen des Problems mit zweidimensionaler Struktur dargestellt.

In Tabelle 7.2 sind zusätzlich die Funktionswerte der Zielfunktionen sowie die CPU-Zeiten der Simulation für die hier verwendeten Berechnungsmethoden angegeben. Darüber hinaus werden entsprechende Eigenschaften weiterer Berechnungsmethoden aufgeführt. Alle ermittelte Werte sind dabei auf zwei Nachkommastellen gerundet.

Struktur	Matlab®-Routine	h	$J_{\text{an}}(\hat{f})$ bzw. $J_1(\hat{l})$	CPU-Zeit in s
eindimensional	`fmincon`	-	10.47	0,56
eindimensional	`patternsearch`	0,1	13,11	5,75
eindimensional	`patternsearch`	0,01	12,98	27,89
zweidimensional	`patternsearch`	0,1	1,85	8,41
zweidimensional	`patternsearch`	0,01	1.69	1340,39

Tabelle 7.2: Funktionswerte der Zielfunktionen und die CPU-Zeiten der Simulation für verschiedene Berechnungsmethoden

Nun können die Berechnungsmethoden untereinander verglichen werden. Zunächst ist dabei der Fokus auf den Vergleich der beiden verwendeten Strukturen untereinander gerichtet, später auf den Vergleich der verschiedenen Berechnungsmethoden innerhalb einer Struktur.

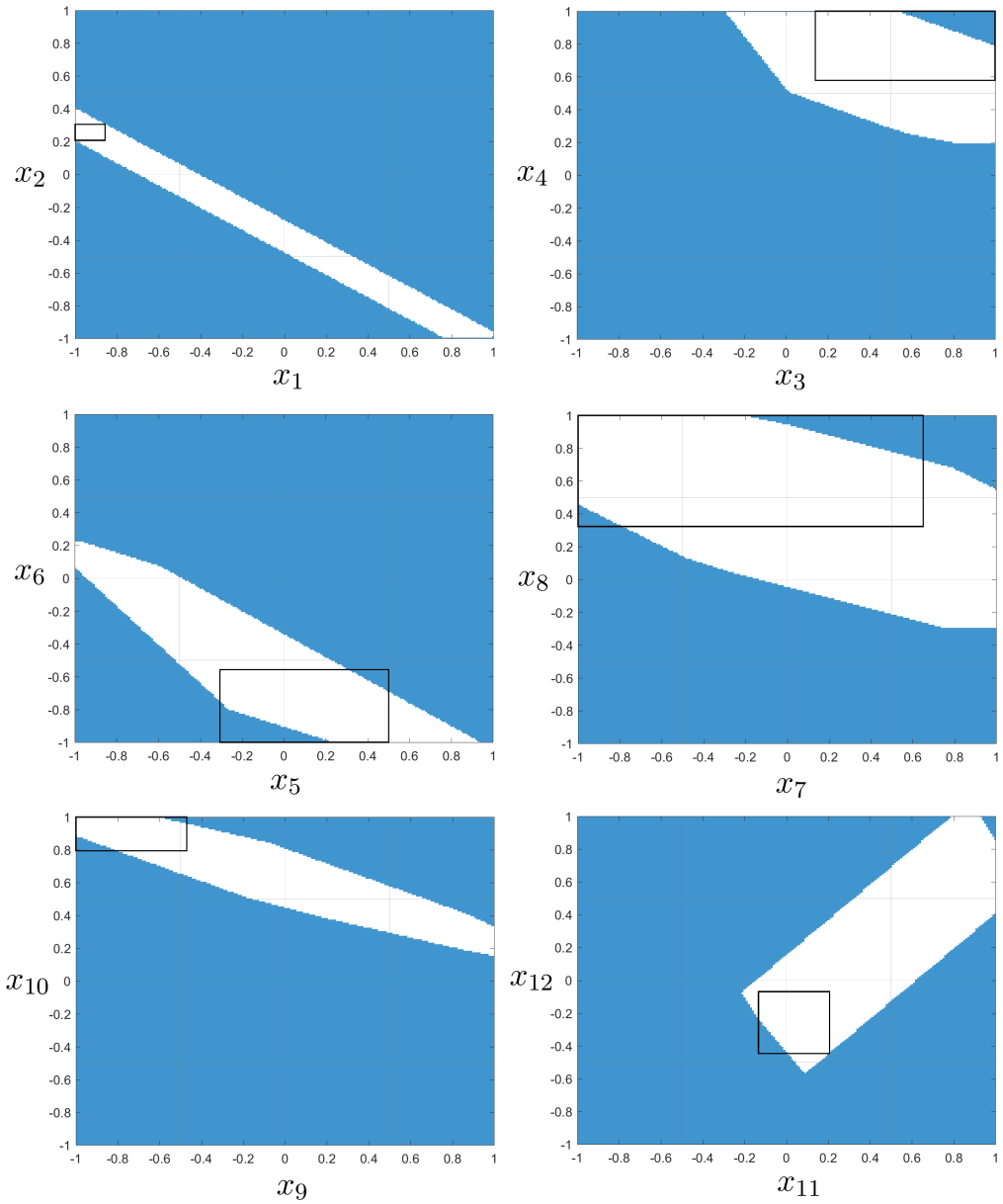


Abbildung 7.6: Zur numerisch berechneten, optimalen Lösung des Beispielproblems gehörige Mengen für die zweidimensionale Struktur (weiß) und für die eindimensionale Struktur (schwarz umrandet)

In Tabelle 7.2 ist zu erkennen, dass der Zielfunktionswert der numerisch berechneten, optimalen Lösung für eine zweidimensionialen Struktur deutlich kleiner ist als der für eine eindimensionale Struktur. Dies bedeutet, dass die Verwendung einer höherdimensionalen Struktur ein größeres Volumen ermöglicht. Dies ist auch in Abbildung

7.6 gut zu sehen. Außerdem stellt sich hier heraus, dass das kartesische Produkt der zur optimalen Lösung gehörigen Mengen für eine eindimensionale Struktur nicht unbedingt eine Teilmenge des kartesischen Produkts der zur optimalen Lösung gehörigen Mengen für eine zweidimensionale Struktur sein muss.

Wie bereits in Beispiel 7.1 steigt auch hier die CPU-Zeit bei einer Verfeinerung der Gitterschrittweite unter Verwendung der Matlab®-Routine `patternsearch` stark an. Dieser Effekt ist bei einer zweidimensionalen Struktur gegenüber einer eindimensionalen Struktur nochmals verstärkt. Dies beruht darauf, dass die Anzahl der Funktionsauswertungen, welche bei der numerischen Approximation des Lebesgue-Maßes im $\mathbb{R}^2$ gegenüber der numerischen Approximation im $\mathbb{R}$ anfallen, quadriert werden müssen.
Auch im Vergleich zur Matlab®-Routine `fmincon` ist die CPU-Zeit bei der Verwendung der Routine `patternsearch` generell größer. Einerseits ist dies darin begründet, dass eine Suche mit Gradienteninformationen gegenüber einer direkten Suche schneller ist. Andererseits sind die Optimierungsvariablen des zugrundeliegenden, konvexen VER-Problems (6.1) für eine eindimensionale Struktur im $\mathbb{R}^{md}$, wobei $md = 108$. Dies bedeutet, dass die Dimension der Optimierungsvariablen des konvexen VER-Problems (6.1) ebenfalls groß gegenüber der Dimension der Optimierungsvariablen des eindimensionalen konvexen VER-Problems (6.3) mit $2d = 24$ ist.

Durch die Approximation des Lebesgue-Maßes ist außerdem die Zielfunktion des konvexen analytischen VER-Problems (6.1) eine Stufenfunktion, weshalb es nicht immer möglich ist, die tatsächliche Lösung des Problems bestmöglich zu approximieren. Damit kann erklärt werden, warum der Zielfunktionswert bei der Verwendung der Matlab®-Routine `patternsearch` größer ist als der bei der Verwendung von `fmincon`.

8. Fazit

Das Ziel der Arbeit war es zunächst das Problem „eine volumenmaximale Rechtecksmenge zu finden, welche in eine über mehrere Anforderungen definierte Menge einbeschrieben ist“ in einem mathematischen Kontext zu beschreiben. Hierzu wurde in Kapitel 5 das sogenannte volumenmaximale-einbeschriebene-Rechtecksmengen-Problem, oder kurz VER-Problem, als abstraktes Optimierungsproblem formuliert. Diese Formulierung basiert unter anderem auf den Produktraumeigenschaften des euklidischen Raums $\mathbb{R}^d$ und den darin enthaltenen Borel-Mengen, denen ein Volumen zugeordnet werden kann. Die entsprechenden Grundlagen dafür wurden in den Kapiteln 2 und 3 untersucht.
Um das VER-Problem mit gängigen Optimierungsalgorithmen lösen zu können, wurde eine analytische Version des VER-Problems eingeführt. Diese gilt für Anforderungen, welche durch additiv zerlegbare und messbare Funktionen festgelegt sind. Es konnte gezeigt werden, dass die Zielfunktion des analytischen VER-Problems unterhalb-stetig ist und somit das Problem eine optimale Lösung besitzt. Allerdings traten in einem Beispiel auch lokale, nicht globale Lösungen auf. Deshalb ist nicht garantiert, dass ein Optimierungsalgorithmus die zu einem VER-Problem gehörige, größtmögliche Rechtecksmenge findet.

Unter der stärkeren Voraussetzung, dass die Funktionen, welche die Anforderung beschrieben, sowohl additiv zerlegbar als auch konvex sind, konnte in Kapitel 6 nachgewiesen werden, dass das VER-Problem ein konvexes Optimierungsproblem ist. Somit ist jede lokale Lösung des konvexen VER-Problems auch eine globale. Die Nachweise dazu beruhen auf den in Kapitel 4 gemachten, theoretischen Betrachtungen über konvexe Mengen und Funktionen und insbesondere auch logarithmisch-konkave Funktionen.
Des Weiteren konnte für das konvexe VER-Problem eine alternative Formulierung der Problemstellung gefunden werden, sofern die zugehörige Rechtecksmenge sich aus eindimensionalen Mengen zusammensetzt. Für eine große Anzahl von Anforderungen kann die Dimension der Optimierungsvariablen so deutlich verringert werden, woraus ein geringerer Rechenaufwand bei der Problemlösung resultiert.

An die mathematische Formulierung anknüpfend wurde in Kapitel 7 eine Möglichkeit vorgestellt, wie das konvexe VER-Problem numerisch mit der Software Matlab® gelöst werden kann. Dazu wurden in Matlab® implementierte Algorithmen zur Lösung von Optimierungsproblemen betrachtet und Methoden eingeführt, mithilfe derer ein Startwert für die Optimierung bestimmt und das Volumen einer konvexen Menge

numerisch berechnet werden kann.
Durch mehrere Beispielprobleme wurden Anwendungsbereiche der VER-Probleme, wie sie in dieser Arbeit betrachtet werden, gezeigt. Die numerische Berechnung des konvexen VER-Problems ist insbesondere für wenige Anforderungen und eine niedrige Dimension des euklidischen Raums $\mathbb{R}^d$ relativ schnell und genau. Für eine größere Anzahl an Anforderungen bzw. höhere Dimensionen des euklidischen Raums $\mathbb{R}^d$ dauern die numerischen Berechnungen länger und eine gute Approximation einer optimalen Lösung ist nicht immer garantiert.

Eine Untersuchung weiterer Optimierungsalgorithmen könnte mögliche Verbesserungen in der Güte der Approximation einer optimalen Lösung liefern. Falls damit nicht zusätzlich eine Verkürzung der Berechnungszeit erreicht wird, könnte eine solche über die Reduzierung der Anforderungen erzielt werden. Hierfür müsste es gelingen eine additiv zerlegbare Funktion zu finden, welche alle Anforderungen gemeinsam und möglichst genau beschreibt. Ebenso könnten Anforderungen, welche nicht durch additiv zerlegbare Funktionen beschrieben werden, durch additiv zerlegbare Funktionen angenähert werden. Eine intelligente Bestimmung derartiger Funktionen könnte daher die Praxisanwendung dieser Arbeit erweitern und soll Gegenstand zukünftiger Forschung sein.

A. Anhang

Beweise

Beweis zu Satz 2.12

Satz A.1. *Seien A, B, C und D Mengen. Für das kartesische Produkt gelten die Eigenschaften*

(i) $A \times (B \cap C) = (A \times B) \cap (A \times C)$ *bzw.* $A \times (B \cup C) = (A \times B) \cup (A \times C)$,

(ii) $A \times B = \emptyset \Leftrightarrow A = \emptyset$ *oder* $B = \emptyset$,

(iii) $A \times B \subset C \times D \Leftrightarrow A \subset C$ *und* $B \subset D$,

(iv) $(A \times B)^{\mathrm{c}} = (A^{\mathrm{c}} \times B^{c}) \cup (A^{\mathrm{c}} \times B) \cup (A \times B^{\mathrm{c}})$,

(v) $(A \cap B) \times (C \cap D) = (A \times C) \cap (B \times D)$.

Beweis. (i) Siehe [16] für einen Beweis.

(ii) Siehe [16] für einen Beweis.

(iii) „$\Leftarrow$“: 1. Fall: Für $A = \emptyset$ oder $B = \emptyset$ gilt stets $A \times B \subseteq C \times D$, vgl. (ii). 2. Fall: Seien A und B nichtleer und $(a, b) \in A \times B$, d.h. $a \in A$ und $b \in B$. Wegen $A \subseteq C$ und $B \subseteq D$ ist auch $a \in C$ und $b \in D$, also gilt $(a, b) \in C \times D$.
„$\Rightarrow$“: Siehe [16] für einen Beweis.

(iv) „$\subset$“: Sei $(a, b) \in (A \times B)^{\mathrm{c}}$, also $(a, b) \notin (A \times B)$, d.h. $a \notin A$ oder $b \notin B$. Somit gilt ($a \notin A$ und $b \notin B$) oder ($a \notin A$ und $b \in B$) oder ($a \in A$ und $b \notin B$), woraus $(a, b) \in A^{\mathrm{c}} \times B^{\mathrm{c}}$ oder $(a, b) \in A^{\mathrm{c}} \times B$ oder $(a, b) \in A \times B^{\mathrm{c}}$ folgt. Deshalb gilt insgesamt $(a, b) \in (A^{\mathrm{c}} \times B^{\mathrm{c}}) \cup (A^{\mathrm{c}} \times B) \cup (A \times B^{\mathrm{c}})$. „$\supset$“: Analog.

(v) Falls A, B, C oder D die leere Menge ist, so ist die Gleichung stets erfüllt. Seien diese also im Folgenden nichtleer.
„$\subset$“: Sei $(a, c) \in (A \cap B) \times (C \cap D)$, d.h. $a \in A$, $a \in B$, $c \in C$ und $c \in D$. Somit ist $(a, c) \in (A \times C)$ und $(a, c) \in (B \times D)$, woraus $(a, c) \in (A \times C) \cap (B \times D)$ folgt. „$\supset$“: Analog.

□

Beweis zu Satz 3.3

Satz A.2. *Für beliebiges $\mathcal{E} \subset \mathcal{P}(X)$ ist die von $\mathcal{E}$ erzeugte σ-Algebra gegeben als der Durchschnitt aller $\mathcal{E}$ umfassenden σ-Algebren über X, genauer*

$$\sigma(\mathcal{E}) = \bigcap \{\mathcal{A} \mid \mathcal{A} \text{ ist } \sigma\text{-Algebra über } X \text{ und } \mathcal{E} \subset \mathcal{A}\}.$$

Beweis. Da $\mathcal{P}(X)$ eine σ-Algebra ist mit $\mathcal{E} \subset \mathcal{P}(X)$, ist das System aller $\mathcal{E}$ umfassenden σ-Algebren über X nichtleer.
Sei $\mathcal{S} := \bigcap \{\mathcal{A} \mid \mathcal{A}$ ist σ-Algebra über X und $\mathcal{E} \subset \mathcal{A}\}$. Es wird nun gezeigt, dass auch $\mathcal{S}$ eine σ-Algebra darstellt. Die leere Menge ist Element jeder σ-Algebra und somit auch Element des Durchschnitts von σ-Algebren, also $\emptyset \in \mathcal{S}$. Für $A \in \mathcal{S}$ gilt $A \in \mathcal{A}$ für alle $\mathcal{A} \in \{\mathcal{A} \mid \mathcal{A}$ ist σ-Algebra über X und $\mathcal{E} \subset \mathcal{A}\}$ und aufgrund der Tatsache, dass $\mathcal{A}$ eine σ-Algebra ist, auch $A^c \in \mathcal{A}$ und deshalb $A^c \in \mathcal{S}$. Analog gilt für $A_j \in \mathcal{S}$, $n \in \mathbb{N}$ auch $A_j \in \mathcal{A}$ für alle $\mathcal{A} \in \{\mathcal{A} \mid \mathcal{A}$ ist σ-Algebra über X und $\mathcal{E} \subset \mathcal{A}\}$ und da $\mathcal{A}$ wieder eine σ-Algebra ist, $\bigcup_{j \in \mathbb{N}} A_j \in \mathcal{A}$ und darum $\bigcup_{j \in \mathbb{N}} A_j \in \mathcal{S}$.
Außerdem folgt für jede beliebige σ-Algebra $\mathcal{A}$ mit $\mathcal{E} \subset \mathcal{A}$ bereits $\mathcal{S} \subset \mathcal{A}$ und $\mathcal{E} \subset \mathcal{S}$, was die Behauptung ergibt, d.h. $\sigma(\mathcal{E}) = \mathcal{S}$. □

Beweis zu Satz 3.4

Satz A.3. *Sei $\mathcal{E}_1$ ein Erzeuger der σ-Algebra $\mathcal{A}_1$ und $\mathcal{E}_2$ ein Erzeuger der σ-Algebra $\mathcal{A}_2$. Falls $\mathcal{E}_1 \subset \mathcal{A}_2$ und $\mathcal{E}_2 \subset \mathcal{A}_1$, dann gilt $\mathcal{A}_1 = \mathcal{A}_2$.*

Beweis. „$\mathcal{A}_1 \subset \mathcal{A}_2$“: Sei $\mathcal{E}_1 \subset \mathcal{A}_2$, dann ist $\mathcal{A}_1 \subset \mathcal{A}_2$.
„$\mathcal{A}_1 \supset \mathcal{A}_2$“: Analog. □

Beweis zu Satz 3.6

Satz A.4. *Die Mengensysteme*

(i) $\mathcal{T}^d = \{U \subset \mathbb{R}^d \mid U \text{ ist offen in } \mathbb{R}^d\}$,

(ii) $\mathcal{C}^d = \{U \subset \mathbb{R}^d \mid U \text{ ist abgeschlossen in } \mathbb{R}^d\}$,

(iii) $\mathcal{I}^d = \{(a, b] \mid a, b \in \mathbb{R}^d, a < b\}$,

(iv) $\mathcal{I}^d_\infty = \{(a, \infty) \mid a \in \mathbb{R}^d\}$

sind Erzeuger der Borel-σ-Algebra $\mathcal{B}(\mathbb{R}^d)$, wobei (a, ∞) für $a \in \mathbb{R}^d$ analog zu (3.1) erklärt ist.

Beweis. (i) Nach Definition der Borel-σ-Algebra gilt bereits $\sigma(\mathcal{T}^d) = \mathcal{B}(\mathbb{R}^d)$.

(ii) Da jede abgeschlossene Teilmenge das Komplement einer offenen Teilmenge des $\mathbb{R}^d$ ist, folgt $\sigma(\mathcal{C}^d) = \sigma(\mathcal{T}^d) = \mathcal{B}(\mathbb{R}^d)$.

(iii) Wegen Satz 3.4 reicht es $\mathcal{I}^d \subset \sigma(\mathcal{T}^d)$ und $\mathcal{T}^d \subset \sigma(\mathcal{I}^d)$ zu zeigen.
„$\mathcal{I}^d \subset \sigma(\mathcal{T}^d)$“: Aus $(a,b] = \bigcap_{k\in\mathbb{N}} \prod_{i=1}^{d} (a_i, b_i + \frac{1}{k})$ resultiert $(a,b] \in \sigma(\mathcal{T}^d)$ für alle $a, b \in \mathbb{R}^d$ mit $a < b$ und deshalb $\mathcal{I}^d \subset \sigma(\mathcal{T}^d)$.
„$\mathcal{T}^d \subset \sigma(\mathcal{I}^d)$“: Sei $U \subset \mathbb{R}^d$ offen. Wie in Kapitel 2.1 bemerkt, lässt sich U als abzählbare Vereinigung offener Mengen der Form $(a,b) \subset U$, vgl. (3.1), mit $a, b \in \mathbb{Q}^d$ und $a < b$ darstellen. Ferner ist $(a,b) = \bigcup_{\substack{c\in\mathbb{Q}^d \\ c<b}} (a,c]$, woraus $(a,b) \in \sigma(\mathcal{I}^d)$ und somit auch $U \in \sigma(\mathcal{I}^d)$ und $\mathcal{T}^d \subset \sigma(\mathcal{I}^d)$ folgt.

(iv) Wegen Satz 3.4 reicht es auch hier $\mathcal{I}^d_\infty \subset \sigma(C^d)$ und $\mathcal{I}^d \subset \sigma(\mathcal{I}^d_\infty)$ zu zeigen.
„$\mathcal{I}^d_\infty \subset \sigma(\mathcal{I}^d)$“: Es gilt $(a,\infty) = \bigcup_{k\in\mathbb{N}} \prod_{i=1}^{d} (a_i, k]$ für alle $a \in \mathbb{R}^d$ und deshalb auch $\mathcal{I}^d_\infty \subset \sigma(\mathcal{I}^d)$.
„$\mathcal{I}^d \subset \sigma(\mathcal{I}^d_\infty)$“: Aus $(a,b] = (a,\infty)\backslash(b,\infty)$ resultiert $(a,b] \in \sigma(\mathcal{I}^d)$ für alle $a, b \in \mathbb{R}^d$ mit $a < b$ und darum $\mathcal{I}^d \subset \sigma(\mathcal{I}^d_\infty)$.

□

Beweis zu Satz 3.11

Satz A.5. *Seien $n \in \mathbb{N}$ und $d^k \in \mathbb{N}$ fest für $k = 1, \ldots, n$ mit $d := \sum_{k=1}^{n} d^k$. Dann gilt*

$$\bigotimes_{k=1}^{n} \mathcal{B}(\mathbb{R}^{d^k}) = \mathcal{B}(\mathbb{R}^d).$$

Beweis.

„$\supset$“: Jedes d-dimensionale halboffene Rechteck $(a,b]$ in $\mathbb{R}^d$ mit $a, b \in \mathbb{R}^d$ im Sinne von (3.1) kann als kartesisches Produkt d^k-dimensionaler halboffenener Rechtecke $(a^k, b^k]$, $k = 1, \ldots, n$, dargestellt werden, d.h. $(a,b] = \prod_{k=1}^{n} (a^k, b^k]$, wobei $a^k, b^k \in \mathbb{R}^{d^k}$. Daher gilt $(a,b] \in \bigotimes_{k=1}^{n} \mathcal{B}(\mathbb{R}^{d^k})$ und da $\mathcal{B}(\mathbb{R}^d)$ die kleinste σ-Algebra ist, die alle halboffenen Rechtecke enthält, auch $\mathcal{B}(\mathbb{R}^d) \subset \bigotimes_{k=1}^{n} \mathcal{B}(\mathbb{R}^{d^k})$.

„$\subset$“: Sei $\mathcal{A}^k$ definiert als

$$\mathcal{A}^k = \left\{ A^k \in \mathbb{R}^{d^k} \mid \mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^{k-1}} \times A^k \times \mathbb{R}^{d^{k+1}} \times \ldots \times \mathbb{R}^{d^n} \in \mathcal{B}(\mathbb{R}^d) \right\}$$

für $k = 1, \ldots, n$. Mit Satz 2.12 verallgemeinert auf mehrfache kartesische Produkte gilt

$$\begin{aligned} &\mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^{k-1}} \times \emptyset \times \mathbb{R}^{d^{k+1}} \times \ldots \times \mathbb{R}^{d^n} = \emptyset, \\ &\mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^{k-1}} \times (A^k)^{\mathrm{c}} \times \mathbb{R}^{d^{k+1}} \times \ldots \times \mathbb{R}^{d^n} \\ &\qquad = \left(\mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^{k-1}} \times A^k \times \mathbb{R}^{d^{k+1}} \times \ldots \times \mathbb{R}^{d^n} \right)^{\mathrm{c}} \end{aligned}$$

für alle $A^k \in \mathbb{R}^{d^k}$ und

$$\begin{aligned} &\mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^{k-1}} \times \left(\bigcup_{j \in \mathbb{N}} A_j^k \right) \times \mathbb{R}^{d^{k+1}} \times \ldots \times \mathbb{R}^{d^n} \\ &\qquad = \bigcup_{j \in \mathbb{N}} \left(\mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^{k-1}} \times A_j^k \times \mathbb{R}^{d^{k+1}} \times \ldots \times \mathbb{R}^{d^n} \right) \end{aligned}$$

für alle $A_j^k \in \mathbb{R}^{d^k}$, $j \in \mathbb{N}$. Weil $\mathcal{B}(\mathbb{R}^d)$ eine σ-Algebra ist, kann gefolgert werden, dass auch $\mathcal{A}^k$ eine σ-Algebra ist. Aufgrund Bemerkung 2.16 enthält die σ-Algebra $\mathcal{A}^k$ alle offenen Mengen, weshalb $\mathcal{B}(\mathbb{R}^{d^k}) \subset \mathcal{A}^k$ und somit

$$\{ \mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^{k-1}} \times A^k \times \mathbb{R}^{d^{k+1}} \times \ldots \times \mathbb{R}^{d^n} \mid A^k \in \mathcal{B}(\mathbb{R}^{d^k}) \} \subset \mathcal{B}(\mathbb{R}^d)$$

für $k = 1, \ldots, n$ gilt. Weiterhin ist

$$\prod_{k=1}^{n} A^k = \bigcap_{k=1}^{n} \mathbb{R}^{d^1} \times \ldots \times \mathbb{R}^{d^{k-1}} \times A^k \times \mathbb{R}^{d^{k+1}} \times \ldots \times \mathbb{R}^{d^n} \in \mathcal{B}(\mathbb{R}^d)$$

für $A^k \in \mathcal{B}(\mathbb{R}^{d^k})$, $k = 1, \ldots, n$, woraus direkt $\bigotimes_{k=1}^{n} \mathcal{B}(\mathbb{R}^{d^k}) \subset \mathcal{B}(\mathbb{R}^d)$ folgt. □

Beweis zu Satz 3.23

Satz A.6. *Seien $(X, \mathcal{A}, \mu)$ ein Maßraum und $A_i \in \mathcal{A}$, $i \in \mathbb{N}$. Dann gelten*

(i) Subtraktivität, d.h. für $A_1 \subset A_2$ und $\mu(A_1) < \infty$ ist $\mu(A_2 \backslash A_1) = \mu(A_2) - \mu(A_1)$,

(ii) Monotonie, d.h. für $A_1 \subset A_2$ ist $\mu(A_1) \leq \mu(A_2)$,

(iii) *σ-Stetigkeit, d.h. für $A_i \uparrow A$ gilt $\lim_{i\to\infty} \mu(A_i) = \mu(A)$ und für $A_i \downarrow A$ mit $\mu(A_1) < \infty$ ebenfalls $\lim_{i\to\infty} \mu(A_i) = \mu(A)$.*

Beweis. Sei $A_1 \subset A_2$, dann gilt $A_1 \dot\cup (A_2 \backslash A_1) = A_2$ und aus der Additivität und Nichtnegativität des Maßes folgt

$$\mu(A_2) = \mu(A_1) + \mu(A_2 \backslash A_1) \geq \mu(A_1),$$

womit die ersten beiden Teile des Satzes bereits gezeigt sind.
Sei $A_i \uparrow A$. Dann sind die Mengen $\tilde{A}_1 := A_1$, $\tilde{A}_j = A_j \backslash A_{j-1}$, $j \geq 2$ paarweise disjunkt und es gilt $A_i = \bigcup_{j=1}^{i} \tilde{A}_j$ und $A = \bigcup_{j\in\mathbb{N}} \tilde{A}_j$. Ferner ist

$$\mu(A) \quad = \quad \sum_{i\in\mathbb{N}} \mu(\tilde{A}_i) = \lim_{i\to\infty} \sum_{j=1}^{i} \mu(\tilde{A}_j) = \lim_{i\to\infty} \mu(\tilde{A}_i).$$

Sei nun $A_i \downarrow A$ mit $\mu(A_1) < \infty$. Dann gilt für die Folge $(\tilde{A}_i)_{i\in\mathbb{N}}$ mit $\tilde{A}_i = A_1 \backslash A_i$, $i \in \mathbb{N}$ gerade $\tilde{A}_i \uparrow A_1 \backslash A$. Mit den obigen Erkenntnissen und der Subtraktivität folgt

$$\mu(A_1) - \mu(A) = \mu(A_1 \backslash A) = \lim_{i\to\infty} \mu(A_1 \backslash A_i) = \mu(A_1) - \lim_{i\to\infty} \mu(A_i)$$

und somit die Behauptung. □

Beweis zu Satz 4.3

Satz A.7. *(i) Seien $C_j \subset \mathbb{R}^d$, $j = 1, \ldots, m$, $m \in \mathbb{N}$ konvexe Mengen, dann ist auch $C := \bigcap_{j=1}^{m} C_j$ eine konvexe Menge.*

(ii) Seien $n \in \mathbb{N}$ und $d^k \in \mathbb{N}$ fest für $k = 1, \ldots, n$ mit $d := \sum_{k=1}^{n} d^k$. Ferner seien $C^k \subset \mathbb{R}^{d^k}$ nichtleer. Die Menge $C := \prod_{k=1}^{n} C^k \subset \mathbb{R}^d$ ist genau dann konvex, wenn die Mengen C^k, $k = 1, \ldots, n$, konvex sind.

Beweis. (i) Seien $x, y \in C_j$ für $j = 1, \ldots, m$ und $\theta \in [0, 1]$. Für alle $j \in \{1, \ldots, m\}$ gilt dann $\theta x + (1-\theta)y \in C_j$ und somit auch sofort $\theta x + (1-\theta)y \in C$.

(ii) „⇒“: Seien $x, y \in C$ mit $x = (x^1, \ldots, x^n)$ und $y = (y^1, \ldots, y^n)$, d.h. $x^k, y^k \in C^k$ für $k = 1, \ldots, n$ und $\theta \in [0, 1]$. Dann ist auch $z := \theta x + (1-\theta)y \in C$ mit $z = (z^1, \ldots, z^n)$, da C konvex ist, was bedeutet, dass $z^k = \theta x^k + (1-\theta)y^k \in C^k$

gilt.
„⇐“: Seien $x^k, y^k \in C^k$ für $k = 1, \ldots, n$ und deshalb auch $x, y \in C$. Dann gilt für $\theta \in [0,1]$ gerade $z^k := \theta x^k + (1-\theta) y^k \in C^k$, $k = 1, \ldots, n$ und darum $z \in C$.
□

Beweis zu Satz 4.5

Satz A.8. *Seien $C \subset \mathbb{R}^d$ offen und konvex und $f : C \to \mathbb{R}$ differenzierbar. So gilt, dass f genau dann konvex ist, wenn für alle $x, y \in C$ mit $x \neq y$ die Ungleichung*

$$f(x) + \nabla f(x)^{\mathrm{T}}(y - x) \leq f(y)$$

erfüllt ist.

Beweis. „⇒“: Wegen $f(\theta y + (1-\theta)x) \leq \theta f(y) + (1-\theta) f(x)$ für alle $x, y \in C \subset \mathbb{R}^d$ mit $x \neq y$ und alle $\theta \in (0,1)$ gilt auch

$$\frac{f(\theta y + (1-\theta)x) - f(x)}{\theta} \leq f(y) - f(x).$$

Grenzwertbildung $\theta \searrow 0$ liefert $\lim\limits_{\theta \searrow 0} \frac{f(\theta y + (1-\theta)x) - f(x)}{\theta} = \nabla f(x)^{\mathrm{T}}(y - x)$, woraus

$$\nabla f(x)^{\mathrm{T}}(y - x) \leq f(y) - f(x)$$

resultiert.
„⇐“: Für alle $x, y \in C \subset \mathbb{R}^d$ und alle $\theta \in (0,1)$ gilt

$$\begin{aligned} f(\theta y + (1-\theta)x) + \theta \nabla f(\theta y + (1-\theta)x)^{\mathrm{T}}(x - y) &\leq f(x), \qquad \text{(A.1)} \\ f(\theta y + (1-\theta)x) + (1-\theta) \nabla f(\theta y + (1-\theta)x)^{\mathrm{T}}(y - x) &\leq f(y). \qquad \text{(A.2)} \end{aligned}$$

Multiplikation von (A.1) mit $(1-\theta)$ und (A.2) mit θ und anschließende Addition ergibt

$$f(\theta y + (1-\theta)x) \leq \theta f(y) + (1-\theta) f(x).$$

□

Beweis zu Satz 4.6

Satz A.9. *Seien $C \subset \mathbb{R}^d$ offen und konvex und $f : \mathbb{R}^d \to \mathbb{R}$ zweimal differenzierbar. So ist f genau dann konvex, wenn $\nabla^2 f(x)$ für alle $x \in C$ positiv semidefinit ist.*

Beweis. „⇒“: Sei f konvex. Ferner seien $x \in C \subset \mathbb{R}^d$ und $h \in \mathbb{R}^d$ gegeben. Weil C offen ist, gibt es ein $\theta_0 \in \mathbb{R}$ mit $x + \theta h \in C$ für alle $\theta \in (0, \theta_0]$. Aufgrund Satz 4.5 gilt dann für alle $\theta \in (0, \theta_0]$ gerade

$$f(x + \theta h) - f(x) - \theta \nabla f(x)^{\mathrm{T}} h \quad \geq \quad 0.$$

Mit der qualitativen Taylor-Formel

$$f(x + h) = f(x) + \nabla f(x)^{\mathrm{T}} h + \frac{1}{2} h^{\mathrm{T}} \nabla^2 f(x) h + o(\|h\|^2),$$

welche aus dem Satz von Taylor folgt, vgl. (4.3), gilt

$$\frac{\theta^2}{2} h^{\mathrm{T}} \nabla^2 f(x) h + o(\|\theta h\|^2) \quad \geq \quad 0$$

für $\theta \searrow 0$. Dabei steht o für die Landau-Notation. Divison durch θ^2 und Grenzwertbildung liefert dann, dass $\nabla^2 f(x)$ positiv semidefinit für alle $x \in C$ sein muss.
„⇐“: Sei die Hessematrix $\nabla^2 f(x)$ für alle $x \in C \subset \mathbb{R}^d$ positiv semidefinit. Aus Gleichung (4.3) und Satz 4.5 resultiert nun sofort die Konvexität von f. □

Sonstiges

Betriebssystem	Windows 10 Home 64-Bit, Version 1607
Prozessor	Intel(R) Core(TM) i7-6500 CPU @ 2.50 GHz
RAM	8 GB
Matlab	9.0.0.307022 (R2016a)

Tabelle A.1: Computer- und Softwareinformationen zu den numerischen Simulationen

Literatur

[1] C. Aliprantis und K. Border. *Infinite Dimensional Analysis.* Springer-Verlag Berlin Heidelberg, 1994.

[2] W. Alt. *Numerische Verfahren der konvexen, nichtglatten Optimierung.* B. G. Teubner Verlag, Wiesbaden, 2004.

[3] C. Audet und J. Dennis Jr. *Analysis of Generalized Pattern Searches.* B. SIAM Journal on Optimization, 13(3),889–903, 2003.

[4] S. Boyd und L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[5] R. Byrd, M. Hribar und J. Nocedal. *An Interior Point Algorithm for Large Scale Nonlinear Programming.* SIAM Journal on Optimization, 9(4), 877-900, 1999.

[6] M. Brokate und G. Kersting. *Maß und Integral.* Springer Basel AG, 2011.

[7] R. Denk und R. Racke. *Kompendium der Analysis*, Band 1. Vieweg+Teubner Verlag, Wiesbaden, 2011.

[8] R. Denk und R. Racke. *Kompendium der Analysis*, Band 2. Springer Spektrum, Wiesbaden, 2012.

[9] S. Erschen, F. Duddeck und M. Zimmermann. *Robust Design using classical optimization.* PAMM Proc. Appl. Math. Mech., 15, 565 – 566, 2015.

[10] C. Geiger und C. Kanzow. *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben.* Springer-Verlag Berlin Heidelberg, 1999.

[11] C. Geiger und C. Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben.* Springer-Verlag Berlin Heidelberg, 2002.

[12] M. Gerdts. *Nichtdifferenzierbare Optimierung.* Vorlesungsskript, `https://www.unibw.de/lrt1/gerdts/lehre/nichtdiff_optimierung.pdf`, 2007.

[13] G. Hörmann. *Grundbegriffe der Topologie.* Vorlesungsskript, `http://www.mat.univie.ac.at/~gue/lehre/13gbtop/GBTopSS13.pdf`, 2013.

[14] W. Kaballo. *Einführung in die Analysis II.* Spektrum Akademischer Verlag, Heidelberg-Berlin, 1997.

[15] T. Kolda, r. Lewis und V. Torczon. *A generating set direct search augmented Lagrangian algorithm for optimization with a combination of general and linear constraints.* Technical Report SAND 2006-5315, Sandia National Laboratories, 2006.

[16] I. Lehmann und W. Schulz. *Mengen - Relationen - Funktionen.* Springer Spektrum, Wiesbaden, 2016.

[17] MathWorks. *Optimization Toolbox^{TM} User's Guide.* `https://de.mathworks.com/help/pdf_doc/optim/optim_tb.pdf`, 2016.

[18] D. Meintrup und S. Schäffler. *Stochastik - Theorie und Anwendungen.* Springer-Verlag Berlin Heidelberg, 2005.

[19] J. Moore. *Mathematical Methods for Economic Theory 1.* Springer-Verlag Berlin Heidelberg, 1999.

[20] A. Prékopa. *On logarithmic concave measures and functions.* Acta Scientiarum Mathematicarum, 34, 335-343, 1973.

[21] A. Prékopa. *Logarithmic Concave Measures and Related Topics.* Stochastic Programming, Academic Press, 63-82, 1980.

[22] B. von Querenburg. *Mengentheoretische Topologie.* Springer-Verlag Berlin Heidelberg, 1973.

[23] S. Volkwein. *Numerische Verfahren der restringierten Optimierung.* Vorlesungsskript, `http://www.math.uni-konstanz.de/numerik/personen/volkwein/teaching/2014WS-Opt2-Script.pdf`, 2014.

[24] M. Zimmermann und J. Hoessle. *Computing solution spaces for robust design.* Int. J. Numer. Meth. Engng, 94, 290–307, 2013.