Notwendige Optimalitätsbedingungen für nichtkonvexe Standortoptimierungsprobleme

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät II

der Martin-Luther-Universität Halle-Wittenberg

vorgelegt von

Herrn Marcus Hillmann

geboren am 15. April 1988 in Halle (Saale)

Betreuerin und Erstgutachterin:

Frau Prof. Dr. Christiane Tammer
Martin-Luther-Universität
Halle-Wittenberg

Zweitgutachter:

Herr Prof. Akhtar A. Khan, Ph.D.
Rochester Institute of Technology
Rochester, NY

Datum der Einreichung: 29. Juni 2023
Datum der Verteidigung: 30. November 2023

Für Julia

„Frodo wäre ohne Sam nicht weit gekommen."

Frodo Beutlin

und für meine Familie

„Wir sind die krasseste Herde, die ich je gesehen habe!"

Sid

Danksagung

Das Anfertigen dieser Dissertation wäre ohne die Unterstützung einiger Personen nicht möglich gewesen, denen ich daher an dieser Stelle meinen tiefsten Dank aussprechen möchte.

Liebe Christiane, du hast mich bereits im Bachelorstudium für die Optimierung begeistert, seitdem meine beiden Masterarbeiten und nun auch die Promotion betreut und mich stets unterstützt und gefördert. Deine unglaublich herzliche Art hat die Zeit mit dir immer angenehm gemacht und dein immenses Wissen hat mir oft neue Impulse vermittelt. Dafür und für deine Freundschaft weit über die Arbeit hinaus danke ich dir von ganzem Herzen.

Lieber Akhtar, meine Zeit als Doktorand hat direkt bei dir in Rochester begonnen und dank deiner Freundlichkeit und Wärme habe ich mich sofort gut aufgehoben gefühlt. Ich danke dir für deine Freundschaft über die Jahre hinweg und freue mich sehr, dass ich meine Arbeit auch mit dir als Gutachter abschließen durfte.

Liebste Julia, ich könnte nicht glücklicher und dankbarer sein, dich in meinem Leben zu haben. Besonders deine bedingungslose Unterstüzung, deine Hingabe und der Halt, den du mir auch in den schwierigsten Phasen dieser Arbeit gegeben hast, erfüllen mich mit tiefster Liebe und Dankbarkeit.

Meiner ganzen Familie möchte ich nicht nur für die überwältigende Unterstützung und das fortwährende Anschubsen während der Promotion danken, sondern für einfach alles, auch davor und danach. Ich bin sehr froh, dass ich euch habe!

Ich danke all meinen Kolleginnen und Kollegen am Institut für Mathematik und speziell innerhalb der Arbeitsgruppe Optimierung und Stochastik, von denen mir viele in den letzten Jahren zu guten Freundinnen und Freunden geworden sind.

Ganz besonders danke ich Ernest, Niklas und Christian für die wunderbare Arbeitsatmosphäre, viele interessante und anregende Gespräche und natürlich unzählige Mittagsrunden. Darüber hinaus danke ich dir, Christian, für die besten Konferenzreisen, die man sich wünschen kann, sowie für deine Freundschaft, obwohl du es so lange mit mir in einem Büro aushalten musstest.

Der letzte Dank gehört meinen Freundinnen und Freunden, für ihre Unterstützung und für einfach viele tolle Momente in den letzten Jahren.

Inhaltsverzeichnis

Symbolverzeichnis

Grundlegende Notation

K	Körper
$\emptyset$	leere Menge
$A \cup B$	Vereinigung von A und B, $x \in A \cup B \Rightarrow x \in A$ oder $x \in B$
$A \cap B$	(Durch-)Schnitt von A und B, $x \in A \cap B \Rightarrow x \in A$ und $x \in B$
$A \setminus B$	relatives Komplement von B in A, $x \in A \setminus B \Rightarrow x \in A$ und $x \notin B$
$A \subseteq B$	A ist Teilmenge von B, $x \in A \Rightarrow x \in B$
$A \subset B$	A ist echte Teilmenge von B, $A \subseteq B$ und $A \neq B$
$A + B$	Addition von A und B, $A + B = \{a + b \mid a \in A, b \in B\}$
$A - B$	Subtraktion von A und B, $A - B = \{a - b \mid a \in A, b \in B\}$
$t \cdot A$	Skalarmultiplikation mit $t \in \mathbb{R}$, $t \cdot A = \{t \cdot a \mid a \in A\}$
$(V, +, \cdot)$	Vektorraum über K oder K-Vektorraum
$\mathbb{N}$	Menge der natürlichen Zahlen
$\mathbb{Q}$	Körper der rationalen Zahlen
$\mathbb{R}$	Körper der reellen Zahlen
$\mathbb{R}_+$	Menge der nichtnegativen reellen Zahlen, $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$
$\mathbb{R}^n$	n-dimensionaler Vektorraum über $\mathbb{R}$
$\mathbb{R}^n_+$	gewöhnlicher Ordnungskegel in $\mathbb{R}^n$
$\overline{\mathbb{R}}$	Menge der erweiterten reellen Zahlen, $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$
$\mathbb{C}$	Körper der komplexen Zahlen
(M, d_M)	metrischer Raum mit der Metrik $d_M \colon M \times M \to \mathbb{R}$
$(X, \\|\cdot\\|_X)$	normierter Raum mit der Norm $\\|\cdot\\|_X \colon X \to \mathbb{R}_+$
sgn	Signum- oder Vorzeichenfunktion
$\lvert \cdot \rvert$	Betragsfunktion auf $\mathbb{R}$, Absolutbetrag
$\\|\cdot\\|_p$	ℓ_p-Norm mit $p \in \overline{\mathbb{R}}$ und $p \geq 1$, $\\|\cdot\\|_p \colon \mathbb{R}^n \to \mathbb{R}$
$\\|\cdot\\|_1$	ℓ_1-Norm bzw. Manhattan-Norm, $\\|\cdot\\|_1 \colon \mathbb{R}^n \to \mathbb{R}$
$\\|\cdot\\|_2$	ℓ_2-Norm bzw. euklidische Norm, $\\|\cdot\\|_2 \colon \mathbb{R}^n \to \mathbb{R}$
$\\|\cdot\\|_\infty$	ℓ_∞-Norm bzw. Maximums-Norm oder Tschebyscheff-Norm, $\\|\cdot\\|_\infty \colon \mathbb{R}^n \to \mathbb{R}$
$\operatorname{dom} f$	effektiver Definitionsbereich oder Domain von f
$\operatorname{epi} f$	Epigraph von f
τ	Topologie (auf der Menge T)
(T, τ), (S, σ)	topologische Räume
O	offene Menge, $O \in \tau$
A	abgeschlossene Menge, $T \setminus A \in \tau$
$B_r(x)$	offene Kugel mit Radius r um $x \in M$
$\overline{B_r(x)}$	abgeschlossene Kugel mit Radius r um $x \in M$

τ_M	metrische Topologie auf M
U	Umgebung von $x \in T$, es existiert eine offene Menge $O \in \tau$ mit $x \in O \subseteq U$
$(V, +, \cdot, \nu)$	topologischer K-Vektorraum
L	Lipschitz-Konstante, $L \geq 0$
f'	Fréchet-Ableitung der Funktion f
X^*	stetiger Dualraum des normierten Raumes $(X, \|\cdot\|_X)$
$\|\cdot\|_*$	Operatornorm auf X^* mit $\|\cdot\|_* \colon X^* \to \mathbb{R}_+$
$\operatorname{int} A$	(algebraisch) Inneres der Menge A
$\operatorname{bd} A$	Rand der Menge A
$\operatorname{cl} A$	Abschluss der Menge A
$\operatorname{Min}(F, D)$	Menge aller minimalen Elemente von F bezüglich D
$\operatorname{WMin}(F, D)$	Menge aller schwach minimalen Elemente von F bezüglich D
$\varphi_{A,k}$	nichtlineares Skalarisierungsfunktional bezüglich der Menge A und der Richtung k
C	Kegel, $c \in C, \lambda \in \mathbb{R}_+ \Rightarrow \lambda c \in C$
min	Minimum
max	Maximum
inf	Infimum, größte untere Schranke
sup	Supremum, kleinste obere Schranke
lim sup	Limes superior, größter Häufungspunkt
lim inf	Limes inferior, kleinster Häufungspunkt
Lim sup	sequentieller oberer Painlevé-Kuratowski-Grenzwert
$\nabla f(x_0)$	Gradient von f an der Stelle x_0
$d^- f(x, h)$	(untere) Dini-Richtungsableitung von f an der Stelle x in Richtung h
$\partial f(x_0)$	konvexes Subdifferential von f an der Stelle x_0
$\partial^- f(x_0)$	Dini-Hadamard-Subdifferential von f an der Stelle x_0
$\partial_a f(x_0)$	approximierendes Subdifferential von f an der Stelle x_0
$\partial_L f(x_0)$	Limiting-Subdifferential von f an der Stelle x_0
$N_a(x_0, A)$	approximierender Normalenkegel an A im Punkt x_0
$\widehat{N}_\varepsilon(x, \Omega)$	Menge aller ε-Normalen an Ω im Punkt x
$\widehat{N}(x, \Omega)$	Prä-Normalenkegel an Ω im Punkt x
$N_L(x_0, \Omega)$	Limiting-Normalenkegel an Ω im Punkt x_0
$\Pi(x, \Omega)$	euklidischer Projektor von x auf Ω
$\mathcal{O}$	Landau-Symbol
$\max\limits_{i \in \{1,\dots,k\}} f_i$	Maximums-Funktion der Teilfunktionen $f_1 \dots, f_k$
Λ_x, Φ_x	Mengen von zulässigen Gewichtsvektoren im Punkt x

Klassifikationsschema

Position 1:

1	1-Standortproblem
N	N-Standortproblem

Position 2:

P	planares Standortproblem (Standortproblem in $\mathbb{R}^2$)
$\mathbb{R}^n$	Standortproblem in $\mathbb{R}^n$

Position 3:

$w_k \gtrless 0$	Standortproblem mit beliebigen positiven und negativen Gewichten
$w_k = \pm 1$	Standortproblem mit positiven und negativen Gewichten vom Betrag 1
$\bullet$	keine Besonderheiten

Position 4:

ℓ_p^m	Standortproblem mit Abstandsfunktionen, die alle von (möglicherweise) unterschiedlichen ℓ_p-Normen mit $p \in \overline{\mathbb{R}}$ und $p \geq 1$ induziert werden
ℓ_p	Standortproblem mit Abstandsfunktionen, die alle von derselben ℓ_p-Norm mit $p \in \overline{\mathbb{R}}$ und $p \geq 1$ induziert werden
$\ell_{p,1<p<+\infty}^m$	Standortproblem mit Abstandsfunktionen, die alle von (möglicherweise) unterschiedlichen ℓ_p-Normen mit $p \in \mathbb{R}$ und $p > 1$ induziert werden
$\ell_{p,1<p<+\infty}$	Standortproblem mit Abstandsfunktionen, die alle von derselben ℓ_p-Norm mit $p \in \mathbb{R}$ und $p > 1$ induziert werden
ℓ_1	Standortproblem mit Abstandsfunktionen, die alle von der ℓ_1-Norm induziert werden
ℓ_2	Standortproblem mit Abstandsfunktionen, die alle von der ℓ_2-Norm induziert werden
ℓ_∞	Standortproblem mit Abstandsfunktionen, die alle von der ℓ_∞-Norm induziert werden
$\bullet$	noch unbestimmt

Position 5:

Σ	Standortproblem mit Summenzielfunktion
max	Standortproblem mit Centerzielfunktion
CD	Standortproblem mit Centdian-Zielfunktion
$m - \sum_{wpar}$	Standortproblem mit vektorieller Zielfunktion aus m Summenzielfunktionen und mit schwacher Pareto-Minimalität als Lösungskonzept
$\bullet$	noch unbestimmt

Standortoptimierungsprobleme

x	neu zu platzierender Standort, $x \in \mathbb{R}^n$
m	Anzahl der existierenden Standorte, $m \in \mathbb{N}$
$a^1, \ldots, a^m$	existierende Standorte, $a^1, \ldots, a^m \in \mathbb{R}^n$
l	Laufindex der vektoriellen Komponenten, $l \in \{1, \ldots, n\}$
k	Laufindex der existierenden Standorte, $k \in \{1, \ldots, m\}$
w	Gewichtsvektor zu den existierenden Standorten, $w = (w_1, \ldots, w_m) \in \mathbb{R}^m$
$d_k(x, a^k)$	Abstandsfunktion zum existierenden Standort a^k, $d_k \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$
r	Anzahl der existierenden anziehenden Standorte, $r \in \mathbb{N} \cup \{0\}$ mit $r \leq m$
$a^1, \ldots, a^r$	existierende anziehende Standorte
i	Laufindex der existierenden anziehenden Standorte, $i \in \{1, \ldots, r\}$
w_i	Gewicht zum existierenden anziehenden Standort a^i, $w_i > 0$
$m - r$	Anzahl der existierenden abstoßenden Standorte
$a^{r+1}, \ldots, a^m$	existierende abstoßende Standorte
j	Laufindex der existierenden abstoßenden Standorte, $j \in \{r+1, \ldots, m\}$
w_j	Gewicht zum existierenden abstoßenden Standort a^j, $w_j < 0$
α	Steuerparameter im Centdianproblem, $\alpha \in [0, 1]$
J_x	Menge der Indizes der betragsmäßig größten Komponenten in x
K_s	Menge von Indizes gleicher Standorte, $s \in \{1, \ldots, u\}$
u	Anzahl der verschiedenen Indexmengen
$f, f_w, f_{w,\alpha}$	Zielfunktion des Standortoptimierungsproblems, ggf. unter Nutzung des Gewichtsvektors w und des Parameters α
$\to \min\limits_{x \in \mathbb{R}^n}$	skalare Minimierungsaufgabe über $\mathbb{R}^n$
$\to \operatorname*{v-wmin}\limits_{x \in \mathbb{R}^n}$	vektorwertige Minimierungsaufgabe über $\mathbb{R}^n$

Abbildungsverzeichnis

1 Einleitung

Ziel der vorliegenden Arbeit ist es, mittels eines neuartigen Ansatzes für verschiedene Problemstellungen der nichtkonvexen Standortoptimierung notwendige Optimalitätsbedingungen herzuleiten und zu vereinheitlichen. Unser Ansatz basiert dabei auf der Nutzung zweier verallgemeinerter Subdifferentiale, dem approximierenden Subdifferential nach Alexander D. Ioffe und dem Limiting-Subdifferential nach Alexander Y. Kruger und Boris S. Mordukhovich, sowie der dazugehörigen Kalküle. Diese stellen jeweils Erweiterungen des klassischen Subdifferentialbegriffs auf nicht notwendigerweise konvexe Funktionen dar.

Obgleich eine Vielzahl theoretischer Resultate zu diesen Subdifferentialen existiert, war es bisher dennoch kaum möglich, sie für praktische Zwecke zu etablieren. Wir werden im Folgenden aber einen nach unserem besten Wissen und Gewissen bislang ununtersuchten Ansatz verfolgen und nutzen dabei neben der Struktur der zugrunde liegenden Problemstellungen sowie Teilen der Subdifferentialkalküle auch unsere (ebenfalls zum Großteil neuen) expliziten Berechnungen nichtkonvexer Subdifferentiale für mehrere wichtige Funktionen. Die erzielten Resultate generalisieren und erweitern dabei nicht nur bekannte Ergebnisse aus konvexen Aufgabenstellungen, sondern vereinfachen durch den einheitlichen Zugang auch weitere Forschungen.

Basierend auf diesen Resultaten lassen sich außerdem Algorithmen entwickeln, welche etablierte Algorithmen der Standortoptimierung verallgemeinern und zur Lösung weitaus praxisrelevanterer Problemstellungen mit entsprechenden nichtkonvexen Zielsetzungen befähigen.

1.1 Historische Entwicklung und Motivation

Die Standortoptimierung ist ein in besonders hohem Maße von praktischen Problemstellungen geprägter Teilbereich der Mathematik. Ziel der Aufgabenstellungen ist es dabei, einen oder mehrere neue Standorte so zu lokalisieren, dass diese unter Berücksichtigung gewisser Abstandsbegriffe und Nebenbedingungen optimal zu bereits existierenden Standorten platziert sind.

Historisch gesehen wurde diese Problemklasse dabei das erste Mal von Pierre de Fermat in einem Brief an Evangelista Torricelli um das Jahr 1640 aufgeworfen. Dabei stellte Fermat die Aufgabe, zu einem gegebenen Dreieck ABC einen Punkt D zu konstruieren, so dass die Summe der Abstände des Punktes zu den Dreieckseckpunkten minimal wird. Torricelli konnte diese Aufgabe geometrisch lösen und wurde somit zum ersten Standortoptimierer der Geschichte. Abbildung 1.1 zeigt diese geometrische Lösung über die Konstruktion gleichseitiger Dreiecke auf den Seiten des gegebenen Dreiecks, der Schnittpunkt der Umkreise dieser Dreiecke ist der gesuchte Fermat-Torricelli-Punkt D.

Ein gutes Jahrhundert später verallgemeinerte Simpson (1750) in seinem Buch „The doctrine and application of fluxions“ die Aufgabenstellung, indem er den Eckpunkten des Dreiecks nicht mehr unbedingt gleiche Anziehungskräfte zuwies und eine geometrische Konstruktion der Lösung bewies. Dadurch war es möglich, jeden der Punkte mehr oder weniger zu gewichten als die anderen, jedoch nach wie vor in einer rein geometrischen Aufgabenstellung. Heutzutage wird meist Weber (1909) als derjenige genannt, der diese verallgemeinerte Problemstellung erstmals praktisch im Kontext der Optimierung von (Industrie-)Standorten formuliert und genutzt hat.

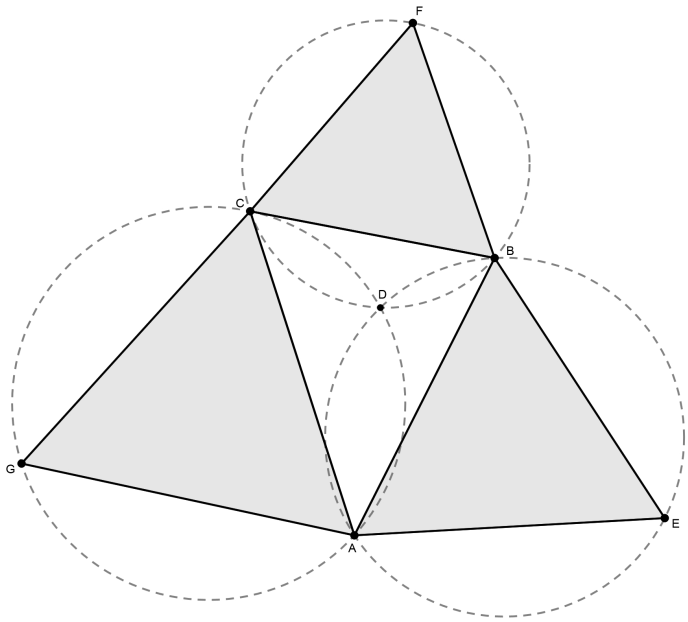

Abbildung 1.1: Torricellis geometrische Lösung des Fermat-Problems (eigene Darstellung mit GeoGebra)

Allerdings weisen beispielsweise Pinto (1977) in einer diesbezüglichen Abhandlung oder Laporte, Nickel und Saldanha da Gama (2015) in ihrem Standardwerk zur Standortoptimierung darauf hin, dass Launhardt (1887, 1900) bereits lange vor Weber optimale Standorte in der Industrie unter Berücksichtigung gewichteter Transportkosten untersucht hat. Da Weber aber außerdem als Erster eine Formalisierung der Problemstellung vorschlug und diese auch auf mehr als drei Standorte ausweitete, scheint die heute übliche Bezeichnung des Optimierungsproblems als Fermat-Weber-Problem oder schlicht Weber-Problem dennoch nicht ungerechtfertigt.

Für drei Standorte ermittelte Launhardt ebenfalls eine geometrische Lösung des gewichteten Problems, indem er neben dem Dreieck ABC der Standorte ein weiteres Dreieck ABO nutzte, dessen Kantenlängen proportional zu den Wichtungen der entsprechenden Punkte sind und welches daher später von Weber als Gewichts-Dreieck bezeichnet wurde. Der Schnittpunkt zwischen dem Umkreis dieses Dreiecks und der Geraden durch O und C ergibt den gesuchten optimalen Standort D. Abbildung 1.2 zeigt diese Konstruktion für dasselbe Ausgangsdreieck ABC wie in Abbildung 1.1 und macht deutlich, wie sehr sich die Lage des Punktes D durch die Gewichtung der Eckpunkte verschiebt.

Trotz einiger weiterer Untersuchungen und Veröffentlichungen in den folgenden Dekaden gelangte die Standortoptimierung erst in den 1960er Jahren mit der Entwicklung des Computers und der damit einhergehenden, stärker algorithmisch geprägten Denkweise zum Durchbruch. Zu den ersten wichtigen Arbeiten der folgenden Blütezeit der Standortoptimierung gehörten die Veröffentlichung von Kuhn und Kuenne

(1962), welche einen bereits von Weiszfeld (1937) entwickelten Algorithmus zur Lösung des verallgemeinerten Fermat-Weber-Problems „wiederentdeckten“, die Untersuchung optimaler Netzwerkverbindungen durch Miehle (1958) sowie die Arbeiten von Hakimi (1964, 1965) zur optimalen Positionierung von einer bzw. mehreren neuen Schaltanlagen in Netzwerken. Cooper (1963) untersuchte in seinem Übersichtsartikel unter anderem auch die optimale simultane Platzierung mehrerer neuer Standorte und warf damit eine ganz neue Problemklasse auf, welche auch heutzutage noch ungebrochenes Forschungsinteresse erfährt, zum Beispiel bei Brimberg und Drezner (2013) oder Drezner, Brimberg, Mladenović und Salhi (2015).

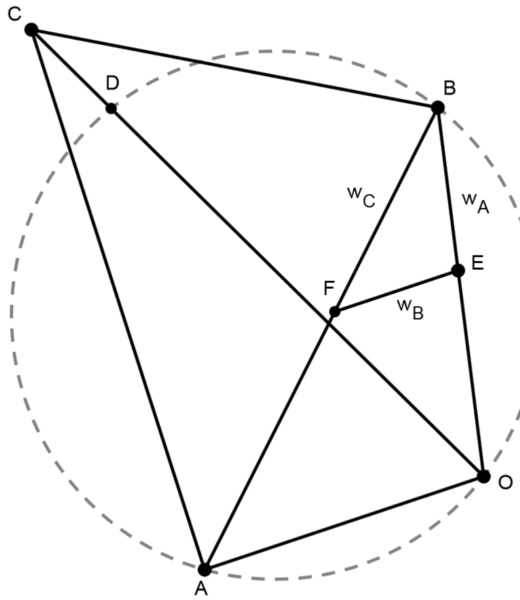

Abbildung 1.2: Launhardts geometrische Lösung der gewichteten Problemstellung (eigene Darstellung mit GeoGebra)

Seitdem ist die Entwicklung der Standortoptimierung vor allem von zwei Bestrebungen geprägt: Einerseits versucht man, praktische Problemstellungen so genau wie möglich zu modellieren und theoretischen Betrachtungen zugänglich zu machen und andererseits sollen die dabei entstehenden, immer komplexer werdenden Problemstellungen natürlich auch gelöst werden, sei es vermittels der Herleitung von Optimalitätsbedingungen oder durch exakt bzw. heuristisch agierende Algorithmen.

Für eine tiefere und detailliertere Darstellung der geschichtlichen Entwicklung der Standortoptimierung und ihrer mannigfaltigen Facetten, welche den Rahmen dieser Arbeit sprengen würde, sei der interessierte Leser auf die Vielzahl an diesbezüglichen Übersichtswerken verwiesen. Insbesondere die Bücher von Love, Morris und Wesolowsky (1988), Drezner (1995), Drezner und Hamacher (2002), Nickel und Puerto (2005) sowie Laporte, Nickel und Saldanha da Gama (2015) sind seit vielen Jahren anerkannte Standardwerke und liefern einen breiten Überblick sowie eine Fülle an Referenzen zu diesem Thema. Weiter bietet Wagner (2015) in ihrer Dissertation eine umfangreiche Literaturübersicht und -analyse. Speziell der Geschichte des Fermat-Weber-Problems widmet sich zudem Wesolowsky (1993) in aller Ausführlichkeit.

Für die vorliegende Arbeit sind unter den verschiedenen Herangehensweisen zur verbesserten Modellierung praktischer Problemstellungen drei Konzepte von elementarer Bedeutung: Die Wahl anderer Abstandsmaße für die einzelnen Abstände, andere Verknüpfungen dieser Einzelabstände in der Zielfunktion sowie die Einführung abstoßender Standorte.

Während im klassischen Sinne der Abstandsbegriff beinahe „naiv“ durch die Anschauung als Luftlinie zwischen zwei Objekten festgesetzt wurde, ist dies längst nicht für alle Standortoptimierungsprobleme ein geeignetes Maß. Innerhalb von Städten kann es beispielsweise deutlich sinnvoller sein, Abstände auf dem Verlauf des Straßennetzes zu basieren. Automatische Fertigungsanlagen und Industrieroboter bewegen sich hingegen meist auf mehreren Achsen gleichzeitig, so dass hier der Betrag des längsten Weges auf einer einzelnen Achse ein sinnvolles Abstandsmaß darstellt. Wir werden ebenfalls verschiedene Abstandsmaße nutzen und diese dafür in Kapitel 2 noch mathematisch exakt formulieren.

Die so gewonnenen einzelnen Abstände können nun wiederum auf unterschiedliche Art und Weise miteinander in Relation gesetzt werden. Im Fermat-Weber-Problem werden die Einzelabstände des neuen Standortes zu den gegebenen Punkten aufaddiert, von dieser Summe wird das Minimum gesucht. Praktischen Nutzen haben aber auch viele andere Konzepte, wie beispielsweise die Minimierung des Maximums der Einzelabstände oder der komponentenweise Vergleich. Die daraus resultierenden Modelle und mögliche Lösungsalgorithmen unterscheiden sich in vielen Aspekten fundamental voneinander. In dieser Arbeit untersuchen wir mit den drei soeben genannten Ansätzen sowie der Centdian-Zielfunktion die vier wichtigsten Zielfunktionen und werden dieses Thema in Unterkapitel 1.2 noch einmal ausführlich aufgreifen.

Der letzte für uns relevante Punkt bei der realitätsnäheren Modellierung von Standortoptimierungsproblemen besteht darin, dass bereits existierende Standorte nicht immer vorteilhaft für einen neuen Standort sein müssen. Oft genug ist das Gegenteil der Fall und der neue Standort soll, aus unterschiedlichen Gründen, möglichst weit fernab eines oder mehrerer dieser Standorte realisiert werden.

Für das Fermat-Weber-Problem wurde diese Erweiterung der Aufgabenstellung erstmals von Tellier (1985) formuliert und für drei Standorte gelöst. Den abstoßenden Einfluss von einigen der existierenden Standorte realisierte er dabei, indem er für diese in der Formulierung der Problemstellung auch negative Gewichtungen zuließ. Da hierdurch im Allgemeinen jedoch die Konvexität der Zielfunktion verloren geht, macht diese Vorgehensweise die Lösung der Aufgabe zwar deutlich schwieriger, gleichzeitig aber lassen sich nur so viele praktisch relevante Probleme sinnvoll modellieren. Weitere Forschungen etablierten in den folgenden Jahren verschiedene exakte und heuristische Verfahren zur Lösung dieser verallgemeinerten Problemstellungen (siehe z. B. Drezner und Wesolowsky (1991), Chen, Hansen, Jaumard und Tuy (1992), Maranas und Floudas (1994) und Nickel und Dudenhöffer (1997)). Carrizosa und Plastria (1999) geben in ihrem Artikel eine sehr gute und kritische Übersicht über die verschiedenen Forschungsansätze in diesem Zeitraum.

Ein Musterbeispiel für die praktische Relevanz dieser Untersuchungen ist die Stadtplanung. Bei der Erschließung neuer Wohnflächen steht immer mehr die Optimierung der Lebensqualität im Vordergrund und so gilt es, sowohl positive Standortfaktoren (z.B. Einkaufsmöglichkeiten, Nahverkehrsanbindung, Arbeitsplätze, Schulen, Arztpraxen etc.) als auch negative Standortfaktoren (z.B. Fabriken, Gefängnisse, Flughäfen, Lärmquellen, stark befahrene Straßen etc.) zu berücksichtigen. Einige dieser Elemente vereinigen dabei sogar beide Eigenschaften, eventuell auch mit unterschiedlichen Abstandsfunktionen. So muss zum Beispiel eine fußläufige Erreichbarkeit des Nahverkehrs anders modelliert werden als der von diesem ausgehende Geräuschpegel, beide Komponenten sind aber gegebenenfalls für die Planung zu beachten.

Ein anderes Beispiel sind automatisierte Produktionsstätten, bei denen verschiedene Produktionsstellen durch eine Maschine auf möglichst kurzen Wegen erreicht werden sollen, andererseits zu gewissen Punkten oder Bereichen aber möglichst große Abstände gehalten werden müssen, um die Sicherheit der Anlage zu gewährleisten und Konflikte mit Arbeitern oder anderen Anlagen auszuschließen.

Ein hochgradig aktuelles und durch die Lage Halles inmitten des mitteldeutschen Braunkohlerevieres auch lokal präsentes Anwendungsgebiet der Standortoptimierung ist durch den Energiewandel gegeben. Im Abschlussbericht der Kommission „Wachstum, Strukturwandel und Beschäftigung“ aus dem Bundesministerium für Wirtschaft und Energie (2019) heißt es dazu, „[...], dass erneute Strukturbrüche sowie soziale und demografische Verwerfungen für die Menschen in allen Revieren dringend zu vermeiden sind und Wertschöpfungsketten vor Ort erhalten bleiben müssen.“, (S. 51) und „Die Kommission empfiehlt deshalb, Maßnahmen zur Beschleunigung von Genehmigungsprozessen zur Errichtung neuer Gaskraftwerke insbesondere an bestehenden Kohlekraftwerksstandorten zu prüfen.“, (S. 68) sowie speziell über die mitteldeutsche Region „Das Revier gewinnt eine hohe Lebensqualität aus dem Ineinandergreifen und der Vernetzung städtischer und ländlicher Räume mit urbanvitalen Quartieren und einer vielseitigen Kulturlandschaft sowie Bergbaufolgelandschaft mit einer hohen Umwelt-, Lebens- und Wohnqualität,[...]“, (S.80). Allein diese drei Aussagen bieten bereits Grundlage für eine Vielzahl hochkomplexer Standortoptimierungsfragen.

Unsere einleitenden Beispiele belegen bereits gut die Vielfalt an Aufgabenstellungen, welche in dieser Arbeit untersucht werden sollen. Wir werden daher als Erstes verschiedene Möglichkeiten zur Modellierung von Standortoptimierungsproblemen mit sowohl anziehenden als auch abstoßenden Standorten (wir werden diese Probleme im Folgenden in Anlehnung an die englische Bezeichnung „semi-obnoxious“ als teil-abstoßend oder Probleme mit teilweiser Abstoßung bezeichnen) vorstellen und erläutern, wieso deren Untersuchung von so hohem wissenschaftlichem Interesse ist.

1.2 Modellierung und Klassifikation der Aufgabenstellungen

Zunächst ist es wichtig, die in dieser Arbeit betrachteten Aufgabenstellungen exakt zu definieren und schematisch einzuordnen. In der Theorie der Standortoptimierung mit ausschließlich anziehenden Standorten und Konvexität der Zielfunktion werden im Wesentlichen drei Modelle zur Verknüpfung der einzelnen Abstände genutzt: Das skalare Medianproblem (mit Summenzielfunktion), das skalare Centerproblem und der mehrkriterielle bzw. vektorielle Ansatz. Einen guten Überblick über verschiedene Modellierungen im konvexen Fall liefern Hamacher, Klamroth und Tammer (2008).

Wir werden diese drei Modelle so erweitern, dass sie einerseits unserer Problemstellung genügen und andererseits in die klassischen konvexen Modelle zurücküberführt werden können. Zusätzlich wollen wir mit der Centdian-Formulierung noch einen eher exotischeren Ansatz betrachten, welcher, wie der Name bereits vermuten lässt, Center- und Medianproblem miteinander verknüpft. Außerdem geben wir für jedes Modell mögliche praktische Anwendungen an und erläutern Stärken und Schwächen der jeweiligen Formulierung.

Zur Klassifikation dieser Probleme nutzen wir das von Hamacher (1995) vorgeschlagene Schema sowie dessen Erweiterungen in den folgenden Jahren durch Hamacher, Nickel und Schneider (1996) sowie Hamacher und Nickel (1998). Die Klassifikation von Problemstellungen ist in der Standortoptimierung wichtig, um diese schnell einordnen und geeignete Lösungsmethoden finden zu können. Auch Solver und Softwarepakete wie beispielsweise der Facility Location Optimizer (FLO) von Günther, Hillmann, Tammer und Winkler (2016) oder die von Hamacher, Maier, Gross und Heßler (2015) entwickelte Library of Location and Layout Algorithms (LoLoLa) greifen auf dieses Schema zurück und ordnen ihre Algorithmen entsprechend ein.

1.2.1 Klassifikation von Standortoptimierungsproblemen

Das Klassifikationsschema nach Hamacher besteht aus fünf Positionen der Form

$$\text{Pos. 1 / Pos. 2 / Pos. 3 / Pos. 4 / Pos. 5},$$

wobei die Positionen folgende Bedeutungen haben:

Pos. 1.: Anzahl und Art der neu zu platzierenden Standorte

Dieser Eintrag enthält eine natürliche Zahl $n \in \mathbb{N}$, welche die Anzahl der neuen Standorte angibt, sowie gegebenenfalls einen Buchstaben, der zusätzliche Angaben über deren Art (Linien, Kreise, etc.) codiert. In dieser Arbeit wird der Eintrag immer eine 1 sein, wir beschäftigen uns ausschließlich mit Problemen, bei welchen ein einzelner neuer Standort punktuell platziert werden soll.

Pos. 2.: Art des Standortoptimierungsproblems

Hier sollte generell zumindest der grobe Typus des Standortoptimierungsproblems vermerkt sein (stetig, diskret oder auf einem Netzwerk), meist enthält dieser Eintrag aber noch genauere Informationen wie ein P für planare (zweidimensionale) Probleme, $\mathbb{R}^n$ für Probleme im allgemeinen n-dimensionalen Vektorraum oder Ähnliches. Wir behandeln hier stets stetige Probleme im $\mathbb{R}^n$, entsprechend wird dieser Eintrag immer $\mathbb{R}^n$ sein.

Pos. 3.: Besonderheiten des Standortoptimierungsproblems

Diese Position ist wohl am vielfältigsten belegbar und enthält beispielsweise Informationen über zulässige Bereiche, verbotene Gebiete, positive und/oder negative Gewichte, Barrieren und vieles mehr. Für unsere Arbeit von besonderer Bedeutung sind hier die Eintragungen über die

Gewichtungen der einzelnen Standorte, eine Beschränkung des zulässigen Bereiches betrachten wir nicht.

Pos. 4.: Definition der Einzelabstände zwischen existierenden und neuen Standorten

Auf dieser Position wird angegeben, auf welche Art und Weise die Abstände zwischen den Standorten gemessen werden. Dies kann entweder mit einer eigenen Distanzfunktion für jede einzelne Paarung von Standorten geschehen oder uniform mittels einer einzigen Funktion für all diese Abstände. Gebräuchliche Maße und damit Eintragungen auf dieser Position sind durch Normen oder Gauges definierte Funktionen oder Kosten- bzw. Nutzenfunktionen. In dieser Arbeit werden wir stets durch ℓ_p-Normen induzierte Abstandsmaße verwenden, dabei kann aber durchaus jedem einzelnen existierenden Standort eine andere induzierende Norm zugeordnet werden. Neben dem allgemeinen Fall werden wir auch Problemstellungen mit einheitlichen Abstandsmaßen untersuchen und dann entsprechend stärkere Resultate herleiten können, dies wird jeweils durch die Eintragung ℓ_p^m im Falle gemischter Metriken bzw. nur ℓ_p im Falle einheitlicher Metriken an dieser Position gekennzeichnet.

Pos. 5.: Beschreibung der Zielfunktion

Die letzte Position gibt an, wie die an Position 4 festgelegten Einzelabstände zur Zielfunktion verknüpft werden. Möglichkeiten hierfür sind Summenzielfunktionen, Maximumsfunktionen, mehrkriterielle Probleme, Centdian-Funktionen und noch einige mehr. Wir werden in dieser Arbeit genau die vier soeben namentlich erwähnten Funktionen verwenden und mit einem einheitlichen Ansatz bearbeiten. Auf diese Position und die damit verbundene Modellierung werden die folgenden Sektionen allerdings für jede Zielfunktion nochmals gesondert eingehen.

Sind in einer Position keine Eintragungen nötig, so wird dies durch $\bullet$ markiert. Das in Unterkapitel 1.1 beschriebene klassische Weber-Problem hat damit beispielsweise das Schema $1/\mathrm{P}/\bullet/\ell_2/\Sigma$ - ein neuer Standort soll planar platziert werden, ohne zusätzliche Besonderheiten, mit der euklidischen Norm als Maß für die Einzelabstände und mit einer Summenzielfunktion. Für weiterführende Informationen zum Klassifikationsschema, insbesondere zu allen möglichen Eintragungen in den einzelnen Positionen, sei nochmals auf Hamacher, Nickel und Schneider (1996) verwiesen.

Abbildung 1.3 zeigt die schematische Darstellung eines teil-abstoßenden Standortoptimierungsproblems zur Planung von Wohnraum in Halle mit der Software FLO (Kartendaten von OpenStreetMap). Dabei sind 7 existierende Standorte zu berücksichtigen, von diesen sind 5 anziehende Standorte (blau gefärbt) und 2 abstoßende Standorte (rot gefärbt). Den Standorten wurden bereits entsprechende positive bzw. negative Gewichte zugewiesen, die hier einfach einer persönlichen Gewichtung entsprechen. Noch ausstehend sind in diesem Beispiel die Zuweisung von Einzelabständen zu den Standorten und die Wahl einer geeigneten Zielfunktion. Die Problemstellung hat in der Klassifikation daher momentan das Schema $1/\mathrm{P}/w_i \gtrless 0/\bullet/\bullet$, wobei die Positionen 4 und 5 noch konkret besetzt werden müssen.

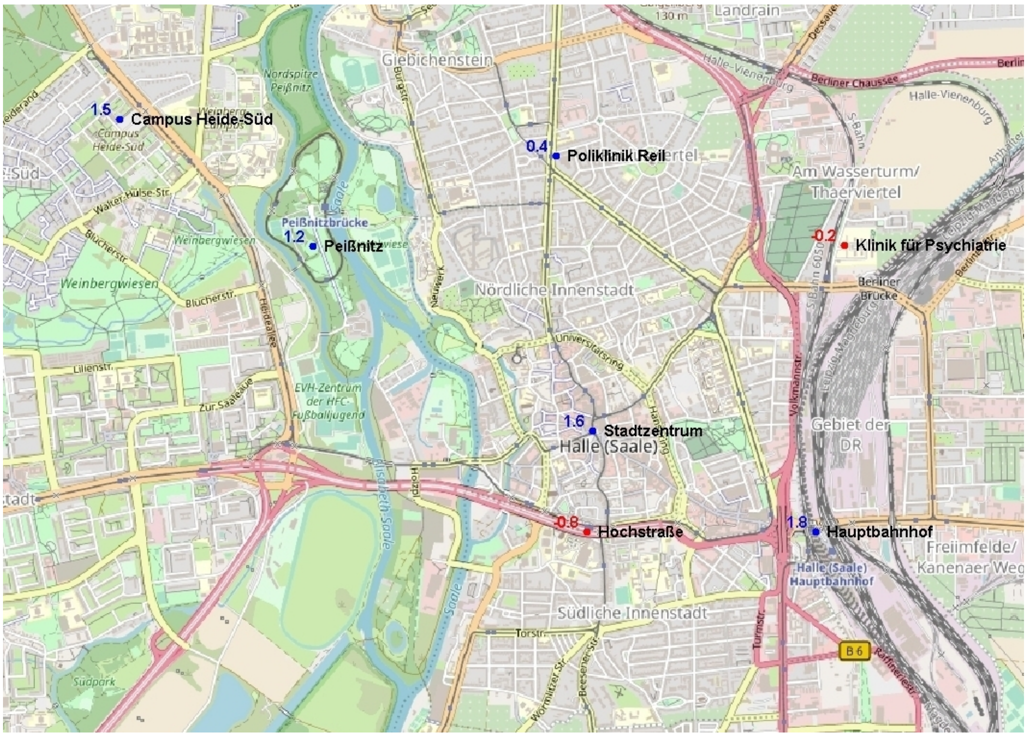

Abbildung 1.3: Ein teil-abstoßendes Standortoptimierungsproblem in Halle (eigene Darstellung mittels FLO)

Als Nächstes wollen wir uns nun mit den verschiedenen Modellierungsmöglichkeiten für Problemstellungen mit teilweiser Abstoßung durch unterschiedliche Zielfunktionen befassen und diese nach dem soeben eingeführten Schema klassifizieren.

1.2.2 Das skalare Medianproblem

Der Median- oder Summenzielfunktions-Ansatz ist die wahrscheinlich intuitivste Möglichkeit zur Bestimmung der Zielfunktion eines Standortoptimierungsproblems und zugleich der Ansatz, welcher auch beim Fermat-Weber-Problem zum Tragen kommt. Die Zielfunktion ist hier die gewichtete Summe der einzelnen Abstände des neuen Standortes zu allen bereits existierenden Standorten, diese Funktion bildet in die reellen Zahlen ab und wird dort minimiert.

Bei einem Problem mit teilweiser Abstoßung wird dabei analog zu Tellier (1985) anziehenden Standorten ein positives Gewicht und abstoßenden Standorten ein negatives Gewicht zugeordnet, wobei der Absolutbetrag des Gewichtes den Grad der Anziehung bzw. Abstoßung bestimmt. Wir betrachten bei dieser und allen folgenden Modellierungen bewusst keine Standorte mit Gewicht 0, da diese keinen Einfluss auf die Problemstellung haben und somit redundant sind oder die Berechnungen ohne Erbringung eines zusätzlichen Nutzens sogar erheblich schwieriger machen.

Die Zielfunktion des teil-abstoßenden Medianproblems ergibt sich durch Addition der anziehenden und der abstoßenden Komponenten als

$$f_w(x) = \sum_{i=1}^{r} w_i d_i(x, a^i) + \sum_{j=r+1}^{m} w_j d_j(x, a^j) \to \min_{x \in \mathbb{R}^n} .$$

Es bezeichnet $x \in \mathbb{R}^n$ den neu zu platzierenden Standort und $a^1, \dots, a^m \in \mathbb{R}^n$ die bereits existierenden Standorte als Punkte in $\mathbb{R}^n$. Dabei sollen die ersten r Standorte $a^1, \dots, a^r$ anziehend und die übrigen $m - r$ Standorte $a^{r+1}, \dots, a^m$ abstoßend sein, mit $0 \leq r \leq m$. Unsere Theorie umfasst als Spezialfälle für $r = m$ bzw. $r = 0$ also insbesondere auch Probleme mit ausschließlich anziehenden bzw. ausschließlich abstoßenden Standorten, die Zielfunktion geht dann in die klassische Zielfunktion eines solchen konvexen bzw. konkaven Problems über.

Weiterhin bezeichnet $w = (w_1, \dots, w_m) \in \mathbb{R}^m$ den Vektor, welcher den existierenden Standorten ihre Gewichte zuordnet, es ist $w_i > 0$ für $1 \leq i \leq r$ (die den anziehenden Standorten zugeordneten Gewichte) und $w_j < 0$ für $r + 1 \leq j \leq m$ (die den abstoßenden Standorten zugeordneten Gewichte). Analog zu den Gewichten ist jedem existierenden Standort a^k mit $1 \leq k \leq m$ auch eine zugehörige Abstandsfunktion $d_k \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ zugeordnet, welche den Einzelabstand zwischen diesem Standort a^k und x misst.

Die Zielfunktion lässt sich nun noch äquivalent umformulieren, um sie unserem Ansatz besser zugänglich zu machen:

$$\begin{aligned} f_w(x) &= \sum_{i=1}^{r} w_i d_i(x, a^i) + \sum_{j=r+1}^{m} w_j d_j(x, a^j) \\ &= \sum_{i=1}^{r} |w_i| d_i(x, a^i) + \sum_{j=r+1}^{m} (-|w_j|) d_j(x, a^j) \\ &= \sum_{i=1}^{r} |w_i| d_i(x, a^i) + \sum_{j=r+1}^{m} |w_j| (-d_j(x, a^j)) \to \min_{x \in \mathbb{R}^n} . \end{aligned} \tag{SP}$$

Dies entspricht also der Aufgabe, bei ausschließlich positiven Gewichten die Summe aus den gewichteten Abständen zu den anziehenden Standorten abzüglich der Summe der gewichteten Abstände zu den abstoßenden Standorten zu minimieren.

In der Klassifikation der Standortoptimierungsprobleme ergibt sich für das teilabstoßende Medianproblem das Schema $1/\mathbb{R}^n/w_k \gtrless 0/\bullet/\Sigma$. Wir wollen einen neuen Standort so platzieren, dass er optimal für bereits existierende Standorte in $\mathbb{R}^n$ liegt. Die existierenden Standorte sind dabei teilweise anziehend und teilweise abstoßend, jedem Standort ist eine eigene, noch nicht näher definierte Abstandsfunktion zugeordnet und diese Einzelabstände werden mittels einer Summenzielfunktion verknüpft. Wir nutzen in dieser Arbeit norminduzierte Abstandsmaße und werden in Kapitel 4 dementsprechend das schematisierte Problem $1/\mathbb{R}^n/w_k \gtrless 0/\ell_p^m/\Sigma$ untersuchen.

Der große Vorteil des Median-Ansatzes liegt neben seiner Einfachheit und Anschaulichkeit vor allem in der vielfältigen Anwendbarkeit, für viele Aufgabenstellungen ist die simple Formulierung als Medianproblem zugleich der beste Ansatz. Allerdings verleitet dies auch oft dazu, direkt eine Summenzielfunktion aufzustellen und andere Modellierungen gar nicht in Erwägung zu ziehen, obwohl diese möglicherweise besser passend sind.

Dieses Modell ist immer dann besonders geeignet, wenn alle Einzelabstände zugleich von Bedeutung sind und in die Zielfunktion einfließen sollen. Unter anderem ist dies im bereits beschriebenen Beispiel der Stadtplanung der Fall, aber auch bei der Optimierung von Produktionsanlagen oder der Allokation von Lagern. Den Anwendungen ist dabei gemein, dass der gesamte Abstand zu den existierenden Standorten optimiert werden soll, auch wenn dies zu Lasten eines oder mehrerer einzelner Standorte geht und die Abstände zu diesen besonders groß (bzw. besonders klein im Falle von abstoßenden Standorten) werden. Ist diese Voraussetzung nicht gegeben, so bietet sich vielmehr die Formulierung als Centerproblem oder mittels gemischtem Centdian-Ansatz an.

Des Weiteren ist es wichtig, dass bereits ein Gewichtungsvektor bekannt ist, sei es aus Erfahrungswerten, Umfragen oder anderen Quellen. Sind hingegen keine Gewichte gegeben, so ist die mehrkriterielle Formulierung zu bevorzugen. Diese liefert im Regelfall eine Menge an Lösungen, aus denen ein Entscheidungsträger dann die für ihn am besten geeignete auswählen kann.

1.2.3 Das skalare Centerproblem

Standortoptimierungsprobleme mit Center- oder Maximumszielfunktion basieren auf einem Ansatz, welcher sich deutlich von dem der Summenzielfunktionen unterscheidet. Bei klassischen Problemstellungen mit ausschließlich anziehenden Standorten ist das Ziel hierbei dadurch charakterisiert, das Maximum der (gewichteten) Einzelabstände zu minimieren, die übrigen Abstände sind für den Zielfunktionswert ohne Bedeutung. Damit ergibt sich in den meisten Fällen auch eine komplett andere Lösungsmenge.

Anschaulich entspricht die Aufgabenstellung ohne Gewichte dabei dem Problem, den Mittelpunkt eines Kreises (oder seines höherdimensionalen Äquivalents) zu finden, welcher bei möglichst kleinem Radius immer noch alle der existierenden Standorte überdeckt (smallest circle problem). Aus dieser Betrachtungsweise entspringt auch der Name Centerproblem.

Für die in dieser Arbeit betrachteten Problemstellungen mit teilweiser Abstoßung ordnen wir wiederum anziehenden Standorten ein positives Gewicht sowie abstoßenden Standorten ein negatives Gewicht zu und verknüpfen die beiden Komponentenfunktionen additiv. Als Zielfunktion ergibt sich also

$$f_w(x) = \max_{i \in \{1,\dots,r\}} w_i d_i(x, a^i) + \max_{j \in \{r+1,\dots,m\}} w_j d_j(x, a^j) \to \min_{x \in \mathbb{R}^n}.$$

Die additive Verknüpfung erscheint hier zunächst weniger naheliegend als beim Medianproblem. Da bei einem Centerproblem aber nur der maximale Einzelabstand für den Zielfunktionswert relevant ist, würde ein Zusammenfassen beider Zielfunktionsterme in einem Maximum den abstoßenden Teil der Problemstellung obsolet werden lassen und hätte damit nur eine höhere Komplexität des Problems ohne nutzbringende Eigenschaften zur Folge.

Wieder bezeichnet $x \in \mathbb{R}^n$ den neu zu platzierenden Standort und $a^1, \dots, a^m \in \mathbb{R}^n$ die bereits existierenden Standorte, wobei die ersten r Standorte $a^1, \dots, a^r$ mit $0 \leq r \leq m$, anziehend und die übrigen $m - r$ Standorte $a^{r+1}, \dots, a^m$ abstoßend sind. In den Spezialfällen ohne anziehende $(r = 0)$ oder abstoßende $(r = m)$ Standorte bettet sich die Zielfunktion erneut übergangslos in die klassische Theorie ein.

Im Falle ausschließlich abstoßender Standorte entspricht das dabei der Aufgabenstellung, das Minimum der (gewichteten) Einzelabstände zu maximieren, also den Mindestabstand zu den unerwünschten Standorten möglichst groß zu halten. Ohne Gewichte kommt dies wiederum dem geometrischen Problem gleich, den Mittelpunkt eines Kreises zu finden, der diesmal aber bei möglichst großem Radius keinen der existierenden Standorte beinhalten soll.

Des Weiteren wird auch hier jedem der existierenden Standorte a^k mit $1 \leq k \leq m$ sowohl eine Funktion $d_k \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ zum Messen des Abstandes zwischen a^k und x als auch mittels des Gewichtungsvektors $w = (w_1, \ldots, w_m) \in \mathbb{R}^m$ ein Gewicht zugeordnet, mit $w_i > 0$ für $1 \leq i \leq r$ (Gewichte, die den anziehenden Standorten zugeordnet sind) und $w_j < 0$ für $r+1 \leq j \leq m$ (Gewichte, die den abstoßenden Standorten zugeordnet sind).

Das teil-abstoßende Centerproblem wird als $1/\mathbb{R}^n/w_k \gtrless 0/\bullet/\max$ klassifiziert. Genau wie beim teil-abstoßenden Medianproblem soll ein neuer Standort optimal in $\mathbb{R}^n$ platziert werden, wobei die existierenden Standorte gemischt anziehend und abstoßend sein können und jeder Standort noch eine eigene Abstandsfunktion zugeordnet bekommt. Der große Unterschied besteht in der Verknüpfung dieser Einzelabstände zur Zielfunktion, für welche hier die Maximumsfunktion genutzt wird. In unserer Arbeit werden wir durch die Wahl norminduzierter Abstandsfunktionen für Position 4 des Schemas daher das Centerproblem $1/\mathbb{R}^n/w_k \gtrless 0/\ell_p^m/\max$ untersuchen.

Wir wollen zudem wieder eine äquivalente Formulierung der Zielfunktion nutzen, welche für unseren Ansatz besser geeignet ist, und für dieses Optimierungsproblem dann in Kapitel 5 notwendige Optimalitätsbedingungen herleiten:

$$\begin{aligned} f_w(x) &= \max_{i \in \{1,\ldots,r\}} w_i d_i(x, a^i) + \max_{j \in \{r+1,\ldots,m\}} w_j d_j(x, a^j) \\ &= \max_{i \in \{1,\ldots,r\}} |w_i| d_i(x, a^i) + \max_{j \in \{r+1,\ldots,m\}} -|w_j| d_j(x, a^j) \\ &= \max_{i \in \{1,\ldots,r\}} |w_i| d_i(x, a^i) + \max_{j \in \{r+1,\ldots,m\}} |w_j|(-d_j(x, a^j)) \to \min_{x \in \mathbb{R}^n} . \end{aligned} \tag{CP}$$

Die häufigste Anwendung erfahren klassische Centerprobleme in Standortplanungen zur (Notfall-)Versorgung. Soll beispielsweise eine neue Feuerwehrstation errichtet werden, so ist es unsinnig, darauf zu achten, dass die Gesamtsumme aller Fahrtwege der Feuerwehrfahrzeuge möglichst kurz ist, wenn dafür die weiter außerhalb liegenden Häuser vor Eintreffen eines Löschwagens niedergebrannt sind. Stattdessen wird man hier natürlich versuchen, den Standort so zu wählen, dass auch die längste Einsatzzeit noch möglichst kurz ist, also das Gebiet gleichmäßig abgedeckt ist. Weniger kritisch trifft diese Betrachtung unter anderem auch auf Versorgungsnetze oder Lieferdienste zu.

Die Formulierung als teil-abstoßendes Problem bietet dabei noch den zusätzlichen Vorteil, gleichzeitig eine bereits bestehende Versorgungslage einbeziehen zu können, indem man als abstoßende Standorte zum Beispiel die bereits existierenden Feuerwehrstationen wählt. Der Standort einer neu zu errichtenden Station kann bei geeigneter Wahl der Gewichte nun so optimiert werden, dass bei weiterhin guter Abdeckung der gewünschten Gebiete der Abstand zu den übrigen Stationen maximiert wird. Dadurch kann das zumeist hoch kostenintensive Versorgungsnetz breiter gestreut werden, ohne dass es zu nennenswerten Einbußen bei der Versorgungslage kommt. Ein Lieferdienst kann dementsprechend in seiner Planung bereits existierende eigene oder fremde Filialen berücksichtigen.

Da das Centerproblem aber gewissermaßen das Gegenstück zum Medianproblem darstellt, liegt sein größter Nachteil auch darin, dass die Summe der Einzelabstände keine Rolle spielt. Für solche Aufgabenstellungen ist dieses Modell daher ungeeignet und eine andere Zielfunktion sollte genutzt werden. Außerdem ist, ebenso wie bei den übrigen skalaren Problemen und in scharfer Abgrenzung zum mehrkriteriellen Fall, das Vorhandensein eines Gewichtungsvektors essentiell. Ohne diesen ist das Nutzen eines mehrkriteriellen Ansatzes sinnvoller, um so zunächst einen Überblick über mögliche Lösungen zu gewinnen.

1.2.4 Das skalare Centdianproblem

Centdianprobleme sind eine eher selten genutzte Möglichkeit bei der Modellierung von Standortoptimierungsproblemen und vereinen, nicht nur dem Namen nach, Center- und Medianprobleme in sich. Das erste Mal erschien diese Problemstellung bei Halpern (1976, 1978) in der parametrischen Optimierung von Konvexkombinationen aus Median- und Centerproblem auf Graphen. Der Grundgedanke dahinter ist es, durch eine Kombination dieser beiden Problemstellungen und die Steuerung über einen Parameter die jeweiligen Nachteile der Grundprobleme auszugleichen.

So beschreiben beispielsweise Pérez-Brito und Moreno Pérez (2000), dass aufgrund der unterschiedlichen Besiedlungsdichte in städtischen Gebieten und im Umland für die Platzierung von öffentlichen Einrichtungen weder Median- noch Centerprobleme komplett geeignet sind. Bei Verwendung eines Median-Ansatzes werden ländliche Gebiete durch die geringe Gewichtung oftmals benachteiligt, während beim Centerproblem die teils großen Entfernungen den Zielfunktionswert überproportional beeinflussen. Pérez-Brito und Moreno Pérez schlagen daher hierfür einen Centdian-Ansatz vor, um besser geeignete Modellierungen und Lösungen zu erhalten.

Im Falle von Problemen mit teilweiser Abstoßung kann dieser Grundgedanke weiterverfolgt werden, um auch hier für gewisse Aufgabenstellungen mit dem Centdian-Ansatz geeignetere Modellierungen zu schaffen. Analog zur klassischen Problemstellung von Halpern erhält man die Zielfunktion dabei durch die Konvexkombination von Medianproblem (SP) und Centerproblem (CP) als

$$\begin{aligned} f_{w,\alpha}(x) = \alpha &\left(\sum_{i=1}^{r} |w_i| d_i(x, a^i) + \sum_{j=r+1}^{m} |w_j| (-d_j(x, a^j)) \right) \\ &+ (1-\alpha) \left(\max_{i \in \{1,\dots,r\}} |w_i| d_i(x, a^i) + \max_{j \in \{r+1,\dots,m\}} |w_j| (-d_j(x, a^j)) \right) \to \min_{x \in \mathbb{R}^n} . \end{aligned} \tag{CDP}$$

Die einzelnen Bezeichnungen sind analog zum Median- oder Centerproblem zu sehen: Mit $x \in \mathbb{R}^n$ wird der neu zu platzierende Standort bezeichnet, mit $a^1, \dots, a^m \in \mathbb{R}^n$ die bereits existierenden Standorte. Wieder sind die ersten r Standorte $a^1, \dots, a^r$, $0 \leq r \leq m$, anziehend, dies entspricht dem jeweils vorderen Median- bzw. Centeranteil der Zielfunktion, und die übrigen $m - r$ Standorte $a^{r+1}, \dots, a^m$ abstoßend, was dann kongruent dem hinteren Median- bzw. Centeranteil entspricht. In den Spezialfällen $r = 0$ (keine anziehenden Standorte) und $r = m$ (keine abstoßenden

Standorte) führt auch das Centdianproblem (CDP) wieder auf die bekannten Aufgabenstellungen der konvexen bzw. konkaven Optimierung zurück.

Jedem der existierenden Standorte a^k mit $1 \leq k \leq m$ ist sowohl eine Abstandsfunktion $d_k \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ als auch durch den Gewichtungsvektor $w = (w_1, \ldots, w_m) \in \mathbb{R}^m$ ein Gewicht $w_i > 0$ für $1 \leq i \leq r$ bzw. $w_j < 0$ für $r + 1 \leq j \leq m$ zugeordnet. Die Wichtung von Median- und Centerproblem in der Zielfunktion wird mittels des Parameters $\alpha \in [0, 1]$ gesteuert, in den Grenzfällen $\alpha = 0$ bzw. $\alpha = 1$ geht die Aufgabenstellung dabei in ein reines Center- bzw. in ein reines Medianproblem über.

In der Klassifikation der Standortoptimierungsprobleme ändert sich im Vergleich zu den beiden anderen skalaren Problemstellungen wieder nur die Zielfunktion und damit der Eintrag an Position 5, für das teil-abstoßende Centdianproblem ergibt sich also das Schema $1/\mathbb{R}^n/w_k \gtrless 0/\bullet/$ CD. Es soll erneut ein neuer Standort optimal im $\mathbb{R}^n$ platziert werden, die existierenden Standorte können wiederum gemischt anziehend und abstoßend sein und jeweils eigene Abstandsfunktionen besitzen, diese werden nun aber mit einer Centdian-Zielfunktion verknüpft. Im Falle von norminduzierten Abstandsfunktionen ergibt sich entsprechend das Schema $1/\mathbb{R}^n/w_k \gtrless 0/\ell_p^m/$ CD, wir untersuchen diese Centdian-Probleme im Kapitel 6 der Arbeit.

Wie bereits beschrieben liegt das Ziel des Centdian-Ansatzes darin, durch Verknüpfung von Center- und Medianproblem die Nachteile der beiden Modelle abzuschwächen oder gar auszugleichen. Neben dem von Pérez-Brito und Moreno Pérez (2000) beschriebenen Stadt-Umland-Problem liegen gut geeignete Problemstellungen für die Centdian-Modellierung immer dann vor, wenn sich die existierenden Standorte auf eine größere und relativ nah zusammenliegende Menge sowie einige weiter entfernte „Ausreißer“ aufteilen. Diese Situation findet man beispielsweise auch bei der Verteilung von Einzelhandelsgeschäften vor: Während sich viele Geschäfte im Innenstadtbereich ballen, sind spezialisierte Geschäfte mit größerem Einzugsbereich (Autohäuser, Möbelläden, Baumärkte,...) oftmals nur in der städtischen Peripherie zu finden. Für die Platzierung eines Standortes mit optimaler Versorgungslage ist daher die Modellierung als Centdianproblem sinnvoll, die Lösung lässt sich dann auch parametrisch steuern. Über den Aspekt der teilweisen Abstoßung können zusätzlich noch unerwünschte Einflussfaktoren eingebunden werden.

Andererseits gibt es jedoch aufgrund der doch sehr speziellen Struktur dieser Problemklasse viele Aufgabenstellungen, in denen sich eine Modellierung als Centdianproblem gegenüber Median- und Centerproblem als nachteilhaft erweist und deren Schwächen sogar kombiniert und verstärkt. Zudem besitzen sowohl das Medianproblem (SP) als auch das Centerproblem (CP) bereits sehr viele freie Parameter, so dass eine Formulierung als Centdianproblem oftmals nur zu einer erhöhten Komplexität der Aufgabenstellung ohne entsprechenden Vorteil führt.

Eine sorgfältige Analyse der Problemstellung ist daher hier noch wichtiger, um zu bestmöglichen Ergebnissen zu gelangen. Weiterhin gilt für das Centdianproblem, wie auch für die übrigen skalaren Modelle, dass eine Gewichtung der Standorte vor der weiteren Untersuchung gegeben sein muss. Auch wenn dies hier teilweise noch durch die inneliegende Parametrisierung ausgeglichen werden kann, so ist doch bei einem völligen Fehlen von Ideen zu möglichen Gewichtungen die im Folgenden vorgestellte mehrkriterielle Formulierung vorzuziehen.

1.2.5 Mehrkriterielle Formulierung

Während sich die drei bisherigen Ansätze dadurch ähneln, dass sie alle skalare Zielfunktionen besitzen, geht man bei der mehrkriteriellen Formulierung einen anderen Weg. Anstatt die einzelnen Abstandsfunktionen zu den existierenden Standorten, welche jeweils Abbildungen in die reellen Zahlen sind, auf irgendeine Weise wieder zu einer reellwertigen Zielfunktion zu verknüpfen, fasst man sie nun stattdessen zu einem Vektor zusammen. Dieser fungiert dann als Zielfunktion für das mehrkriterielle Problem, seine Dimension entspricht dabei der Anzahl der existierenden Standorte.

Das verändert die gesamte Aufgabenstellung grundlegend. Da sich Vektoren über den reellen Zahlen, anders als die reellen Zahlen selbst, im Allgemeinen nicht direkt miteinander vergleichen lassen, benötigen wir zunächst einen neuen Lösungsbegriff. Wir werden dafür das Konzept der schwachen Pareto-Minimalität benutzen und suchen daher anschaulich nach Elementen, deren Abbild unter der Zielfunktion sich nicht in ausnahmslos allen Komponenten noch weiter verbessern lässt. In Sektion 2.1.4 werden wir darauf zurückkommen und diesen Lösungsbegriff mathematisch exakt formulieren.

Auch die Modellierung der Problemstellung weicht bei diesem Ansatz von denen der vorherigen Sektionen ab. Völlig anders als bei skalaren Problemen ist die genaue Gewichtung der Standorte abseits von positiv/negativ nämlich komplett unerheblich. In Kapitel 7 werden wir auch noch genau darauf eingehen, wieso dies in mehrkriteriellen Aufgabenstellungen der Fall ist. Es reicht hier also aus, abstoßende Standorte mit einem Minuszeichen zu versehen. Daher ordnen wir nun anziehenden Standorten den Faktor 1 und abstoßenden Standorten den Faktor -1 zu, um das Problem zu modellieren. Damit ergibt sich das zu lösende Problem als

$$f(x) = \begin{pmatrix} d_1(x, a^1) \\ \vdots \\ d_r(x, a^r) \\ -d_{r+1}(x, a^{r+1}) \\ \vdots \\ -d_m(x, a^m) \end{pmatrix} \to \operatorname*{v-wmin}_{x \in \mathbb{R}^n}. \tag{MP}$$

Es ist wieder $x \in \mathbb{R}^n$ der neu zu bestimmende Standort und $a^1, \dots, a^m \in \mathbb{R}^n$ sind die existierenden Standorte, die Zielfunktion f bildet damit also von $\mathbb{R}^n$ nach $\mathbb{R}^m$ ab. Die ersten r Standorte $a^1, \dots, a^r$ mit $0 \leq r \leq m$ sind wiederum die anziehenden Standorte (hier mit Faktor 1), die übrigen $m - r$ Standorte $a^{r+1}, \dots, a^m$ die abstoßenden Standorte mit Faktor -1. Außerdem ist wieder jedem der existierenden Standorte a^k mit $1 \leq k \leq m$ eine Abstandsfunktion $d_k \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ zugeordnet. In den Spezialfällen $r = m$ und $r = 0$ geht auch diese Zielfunktion in die jeweiligen Zielfunktionen der klassischen konvexen bzw. konkaven Theorie über.

Obwohl sich die mehrkriterielle Formulierung so stark von den bisherigen Ansätzen unterscheidet, ändern sich im Klassifikationsschema nur zwei Positionen. Einem mehrkriteriellen Standortoptimierungsproblem mit teilweiser Abstoßung wird das Schema $1/\mathbb{R}^n/w_k = \pm 1/\bullet/m{-}\sum_{wpar}$ zugeordnet. Es soll ein neuer Standort so in $\mathbb{R}^n$ platziert werden, dass dieser schwach Pareto-minimal unter der m-komponentigen Zielfunktion ist, wobei Anziehung und Abstoßung der existierenden Standorte sowie deren noch festzulegende Abstandsfunktionen berücksichtigt werden. Position 5 sagt insbesondere aus, dass die Vektorzielfunktion komponentenweise Summenzielfunktionen als Einträge besitzt, diese bestehen bei uns allerdings nur aus jeweils einem Element. Für ein mehrkriterielles Problem mit (möglicherweise verschiedenen) norminduzierten Abstandsfunktionen ergibt sich entsprechend die Klassifikation $1/\mathbb{R}^n/w_k = \pm 1/\ell_p^m/m - \sum_{wpar}$.

Im Unterschied zu den skalaren Problemstellungen ist die Lösungsmenge mehrkriterieller Standortoptimierungsprobleme im Allgemeinen überabzählbar. Darum bedarf es eines Entscheidungsträgers, der im Anschluss an die mathematische Lösung des mehrkriteriellen Problems aus dieser Menge noch die für ihn am besten geeignete Möglichkeit für den neuen Standort auswählt. Da andererseits keinerlei Gewichte mehr notwendig sind, verschiebt sich der externe Faktor weg von der Modellierung des Problems zur tatsächlichen Auswahl einer Lösung.

Dies stellt einen entscheidenden Vorteil des mehrkriteriellen Ansatzes dar: Gibt es keine Vorgabe an Gewichten oder ist ungewiss, welche Wichtung geeignet ist, dann kann mit der mehrkriteriellen Formulierung trotzdem sinnvoll eine Aufgabenstellung modelliert werden. Damit ist es möglich, einen Überblick über die Lösungsmenge zu gewinnen und gegebenenfalls auch Gewichte für eine andere Zielfunktion zu finden. Wir werden in Kapitel 7 bei der Untersuchung mehrkriterieller Problemstellungen sehen, dass in dieser Hinsicht sogar eine sehr enge Verknüpfung zwischen den Problemen (MP) und (SP) besteht.

Die entstehende Lösungsmenge stellt allerdings auch einen großen Nachteil dieser Modellierung dar, da in der Praxis zum tatsächlichen Festlegen eines einzigen neuen Standortes (fast) immer noch der Input eines Entscheidungsträgers notwendig ist. Oftmals sind diese Entscheidungsträger dann von der Vielzahl an Möglichkeiten schlicht überfordert oder wählen eine Option, die sich mit keiner der vorab formulierten möglichen Gewichtungen in Einklang bringen lässt.

Zudem sind von vornherein alle Standorte gleich wichtig und werden in der Zielfunktion gleichermaßen berücksichtigt. Entsprechend bieten mehrkriterielle Modelle in dieser Hinsicht die gleichen Vor- und Nachteile wie Medianprobleme und sind damit auch am besten für solche Problemstellungen geeignet, für die bei Vorgabe einer Gewichtung ein Medianproblem genutzt werden würde. Für andere Aufgabenstellungen ist eine Formulierung als Centerproblem oder Centdianproblem meist effektiver.

1.3 Aufbau der Arbeit

In diesem ersten Kapitel wurde die Zielsetzung der vorliegenden Arbeit vorgestellt: Wir wollen die bisher größtenteils nicht bekannten Limiting-Subdifferentiale für positive und negative ℓ_p-Normen konkret berechnen und anschließend mithilfe dieser Berechnungen sowie eines neuen und vereinheitlichten Ansatzes notwendige Optimalitätsbedingungen für eine Vielzahl praktisch hochrelevanter, nichtkonvexer Standortoptimierungsprobleme herleiten, für die solche Bedingungen bisher überhaupt nicht oder nur bruchstückhaft existieren. Die unterschiedlichen zu betrachtenden Aufgabenstellungen haben wir in diesem Kapitel dafür exakt formuliert und klassifiziert sowie in den historischen und mathematischen Kontext eingeordnet.

In Kapitel 2 steht zunächst eine Rekapitulation grundlegender, aber unverzichtbarer analytischer Begriffe wie Metriken und Normen oder Stetigkeit an. Da unsere Arbeit wesentlich auf dem approximierenden Subdifferential nach Ioffe und dem Limiting-Subdifferential nach Kruger und Mordukhovich aufbaut, werden wir dann zunächst die allgemeine Theorie der Subdifferentiale darlegen und dabei neben den genannten auch noch weitere Subdifferentiale präsentieren. Anschließend werden wir jedoch ausführen, warum sich aus der Vielzahl von Subdifferentialbegriffen für den nichtkonvexen Fall genau diese beiden für unsere Zwecke als besonders geeignet erweisen. Kapitel 2 wird fortgeführt durch die explizite Vorstellung des approximierenden und des Limiting-Subdifferentials sowie des jeweiligen Subdifferentialkalküls. Den Abschluss dieses Kapitels bildet ein Vergleich beider Subdifferentiale, insbesondere für die von uns betrachteten Aufgabenstellungen.

Beginnend mit dem folgenden Kapitel 3 präsentieren wir die im Rahmen dieser Promotion erarbeiteten eigenen Resultate. Während es für die vorgestellten Subdifferentiale für nichtkonvexe Funktionen zwar eine Vielzahl theoretischer Betrachtungen gibt, ist ihre praktische Anwendung aufgrund der Schwierigkeit der Berechnungen jedoch höchst selten möglich. Wir werden hingegen die approximierenden bzw. Limiting-Subdifferentiale von verschiedenen Normfunktionen und ihren Negativen konkret berechnen, was uns überhaupt erst den Zugang zum gewünschten subdifferentialbasierten Ansatz zur Herleitung von Optimalitätsbedingungen ermöglicht. Weiter stellen wir auch das Limiting-Subdifferential für die Maximumsfunktion und das Limiting-Subdifferential eines nichtlinearen Skalarisierungsfunktionals vor.

Auf diesen Ergebnissen basieren die notwendigen Optimalitätsbedingungen für die soeben dargestellten Modellierungen, welche den inhaltlichen Schwerpunkt der Arbeit bilden. Im nun folgenden Kapitel 4 stellen wir diese Optimalitätsbedingungen und ihre Herleitungen zunächst für das skalare Medianproblem vor. Dieses Kapitel ist nochmals zweigeteilt, wobei das erste Unterkapitel die Resultate für solche Medianprobleme enthält, bei denen spezielle Metriken als Maß für die Einzelabstände genutzt werden. Im zweiten Unterkapitel werden wir anschließend das wichtigste Resultat dieser Promotion in Form einer notwendigen Optimalitätsbedingung für Medianprobleme mit gemischten ℓ_p-Normen vorstellen und beweisen.

Ebenso werden wir in den Kapiteln 5, 6 und 7 neuartige notwendige Optimalitätsbedingungen für Centerprobleme, Centdianprobleme und mehrkriterielle Probleme unter der Nutzung verschiedener Abstandsfunktionen präsentieren. Allerdings stellen wir in diesen Kapiteln jeweils das Resultat zu Problemstellungen mit gemischten ℓ_p-Normen vorweg, welches sich durch unseren Ansatz stets auf die notwendige Optimalitätsbedingung eines zugehörigen Medianproblems zurückführen lässt. Von diesem allgemeinen Resultat gehen wir dann zu den spezielleren Betrachtungen über.

Im vorletzten Kapitel 8 diskutieren wir noch kurz die Verwendung der Resultate in Lösungsalgorithmen als möglichen Anwendungsbereich der erzielten Ergebnisse. Dabei werden wir auch sehen, dass sich hierdurch bereits existierende Algorithmen für konvexe Fälle harmonisch erweitern lassen.

Abgeschlossen wird die Arbeit mit Kapitel 9, in welchem die erreichten Resultate nochmals reflektiert sowie weitere offene Fragestellungen und mögliche Anknüpfungspunkte an diese Arbeit aufgezeigt werden.

2 Mathematische Grundlagen

Im folgenden Kapitel soll die mathematische Basis für unsere gesamte Arbeit geschaffen werden, um danach auf diesen Grundlagen aufbauend weiterführende Resultate erzielen zu können. Dazu werden zunächst alle benötigten Begriffe eingeführt bzw. wiederholt und einige Bezeichnungen festgesetzt. Wir werden dabei noch eine Zweiteilung vornehmen zwischen analytischem und algebraischem Basiswissen und der spezielleren, aber für diese Arbeit unerlässlichen, verallgemeinerten Differentiation.

2.1 Analytische und algebraische Grundlagen

Dieses Unterkapitel soll also zuerst eine kurze Wiederholung grundlegender Begriffe aus Analysis und (linearer) Algebra liefern, welche für diese Arbeit von größter Bedeutung sind. Das beinhaltet vor allem Metriken und Normen sowie die zugehörigen Räume, topologische Grundbegriffe, Eigenschaften der betrachteten Funktionale, insbesondere verschiedene Arten der Stetigkeit, sowie die Skalarisierung. Wir greifen für diese Definitionen auf verschiedene Grundlagenwerke beider Bereiche zurück und wenden sie in geeigneter Weise an.

2.1.1 Metrische und normierte Räume

Metrische und normierte Räume sind auf gleich zweierlei Art essentiell für unsere Arbeit: Einerseits werden sie zur Definition der verallgemeinerten Subdifferentiale in Unterkapitel 2.2 benötigt und andererseits liefern Metriken und Normen uns die Abstandsmaße in den zu untersuchenden Optimierungsproblemen (vgl. Unterkapitel 1.2). Wir werden uns schrittweise von allgemeinen zu immer spezielleren Räumen bewegen und beginnen dabei mit der Definition von Vektorräumen, welche wir an Definition 2.5.4 in Behrends (2004a) anlehnen.

Definition 2.1.1

Es seien K ein Körper, V eine Menge, $+\colon V \times V \to V$ eine innere Abbildung sowie $\cdot\colon K \times V \to V$ eine äußere Abbildung. Wir nennen das Tripel $(V, +, \cdot)$ einen ***Vektorraum über dem Körper K*** *oder einen* ***K-Vektorraum****, wenn für alle $u, v, w \in V$ und $\alpha, \beta \in K$ die folgenden Eigenschaften erfüllt sind:*

(V1) $u + v = v + u$ *(Kommutativgesetz)*

(V2) $u + (v + w) = (u + v) + w$ *(Assoziativgesetz)*

(V3) Es existiert ein Element $0_V \in V$ mit $v + 0_V = 0_V + v = v$ (neutrales Element oder Nullelement bezüglich $+$).

(V4) Es existiert ein Element $-v \in V$ mit $v + (-v) = 0$ (inverses Element bezüglich $+$).

(S1) $\alpha \cdot (u + v) = (\alpha \cdot u) + (\alpha \cdot v)$ *(Distributivgesetz)*

(S2) $(\alpha + \beta) \cdot v = (\alpha \cdot v) + (\beta \cdot v)$

(S3) $(\alpha \cdot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$

(S4) *Für das Einselement* $1 \in K$ *gilt* $1 \cdot v = v$.

Die gebräuchlichsten K-Vektorräume sind Vektorräume über den Körpern der reellen Zahlen $\mathbb{R}$ oder der komplexen Zahlen $\mathbb{C}$. Wenn aus dem Zusammenhang ersichtlich ist, welche Abbildungen den Vektorraum $(V, +, \cdot)$ definieren, so werden wir diesen nur mit V bezeichnen. Lässt sich zusätzlich noch eine Art „Maßfunktion“ etablieren, so erhalten wir die für uns so wichtigen metrischen und normierten Räume, diese findet man auch als Definition 3.1.1 und 3.1.2 bei Behrends (2004a).

Definition 2.1.2

Für eine beliebige Menge M *heißt die Abbildung* $d_M \colon M \times M \to \mathbb{R}$ ***Metrik*** *oder* ***Abstandsfunktion*** *auf* M*, wenn* d_M *für alle* $x, y, z \in M$ *die Bedingungen*

(M1) $d_M(x, y) \geq 0$ *und* $d_M(x, y) = 0$ *genau dann, wenn* $x = y$ *ist (Definitheit),*

(M2) $d_M(x, y) = d_M(y, x)$ *(Symmetrie) und*

(M3) $d_M(x, z) \leq d_M(x, y) + d_M(y, z)$ *(Dreiecksungleichung)*

erfüllt. Ist M *ein Vektorraum, so nennen wir das Tupel* (M, d_M) *einen* ***metrischen Raum****.*

Definition 2.1.3

Es sei X *ein Vektorraum über dem Körper* K *der reellen oder der komplexen Zahlen. Wir nennen* $\| \cdot \|_X \colon X \to \mathbb{R}_+ := \{x \in \mathbb{R} \mid x \geq 0\}$ *eine* ***Norm*** *auf* X *und* $(X, \| \cdot \|_X)$ *einen* ***normierten Raum****, wenn* $\| \cdot \|_X$ *für alle* $x, y \in X$ *und* $\alpha \in K$ *die folgenden Eigenschaften erfüllt:*

(N1) Es ist $\|x\|_X = 0$ *genau dann, wenn* $x = 0$ *ist (Definitheit),*

(N2) $\|\alpha x\|_X = |\alpha| \|x\|_X$ *(Homogenität) und*

(N3) $\|x + y\|_X \leq \|x\|_X + \|y\|_X$ *(Dreiecksungleichung).*

Zwischen Metriken und Normen besteht überdies ein wichtiger Zusammenhang:

Bemerkung 2.1.1

Jede Norm erzeugt mittels $d_X(x, y) := \|x - y\|_X$ *in eindeutiger Weise eine Metrik auf* X *bzw. auf Teilmengen von* X*, die* ***induzierte Metrik****.*

Für uns ist besonders die folgende Klasse von Normen auf endlich-dimensionalen Räumen von hohem Interesse, da deren induzierte Metriken als Abstandsfunktionen für in die Unterkapitel 1.2 modellierten Aufgabenstellungen fungieren.

Lemma 2.1.1

Für jedes $p \in \mathbb{R} \cup \{+\infty\}$ *mit* $p \geq 1$ *ist durch die Funktion* $\| \cdot \|_p \colon \mathbb{R}^n \to \mathbb{R}$ *mit*

$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \quad \text{für } p \in \mathbb{R} \qquad \text{bzw.} \qquad \|x\|_\infty := \max_{i \in \{1, \ldots, n\}} |x_i|$$

eine Norm auf $\mathbb{R}^n$ *definiert, die sogenannte* ℓ_p***-Norm****.*

Beweis : Es lässt sich leicht nachrechnen, dass für all diese Funktionen die Eigenschaften $(N1) - (N3)$ aus Definition 2.1.3 erfüllt sind. ■

Abbildung 2.1 zeigt für wichtige Spezialfälle der ℓ_p-Normen, nämlich für die Manhattan-Norm ℓ_1 (links), die euklidische Norm ℓ_2 (Mitte) und die Maximums-Norm ℓ_∞ (rechts), die Einheitssphären in der Ebene, also die Menge aller Punkte in $\mathbb{R}^2$ mit einer Norm von 1.

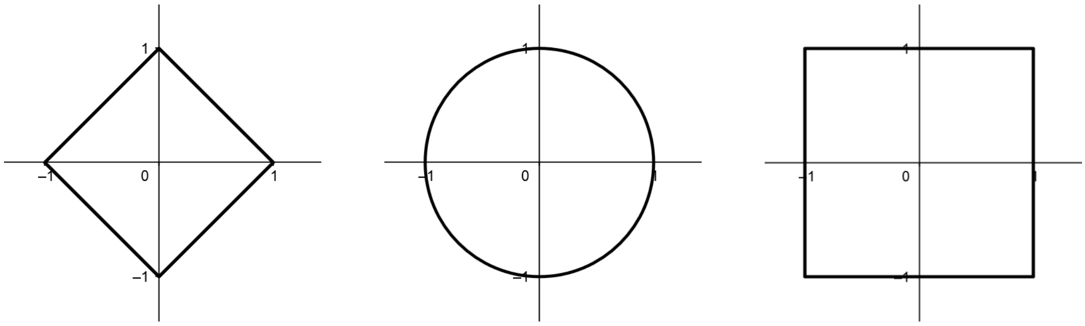

Abbildung 2.1: Einheitssphären von Manhattan-Norm, euklidischer Norm und Maximums-Norm in $\mathbb{R}^2$ (eigene Darstellung mit GeoGebra)

Die von den ℓ_p-Normen induzierten Metriken werden auch als Minkowski-Metriken bezeichnet und bilden in dieser Arbeit die Basis für die Messung der Einzelabstände in unseren Standortoptimierungsproblemen. Abbildung 2.1 verdeutlicht nochmals, welchen großen Einfluss die Wahl dieses Abstandsmaßes auf die zu untersuchenden Problemstellungen hat.

Die nächste von uns betrachtete Klasse von Räumen sind die Banachräume, deren Definition sich unter anderem als Definition I.1.2 in Werner (2018) findet.

Definition 2.1.4

*Ein **Banachraum** ist ein vollständiger normierter Raum $(X, \|\cdot\|_X)$, also ein normierter Raum, in dem jede Cauchy-Folge aus Elementen von X in der von der Norm induzierten Metrik konvergiert.*

Obgleich Banachräume in unserer Arbeit eine essentielle Rolle spielen, würde eine komplette Einführung des Begriffes der Vollständigkeit den Rahmen dieses Kapitels sprengen. Hierzu sei daher, ebenso wie für sonstige weiterführende Aussagen zu diesem Grundlagenkapitel, auf die Vielzahl an Büchern über Grundlagen der Analysis und Funktionalanalysis verwiesen, zum Beispiel Behrends (2004a,b), Göpfert und Riedrich (1994) oder Werner (2018).

Die von uns betrachteten Räume haben aufgrund des folgenden Resultates (Korollar I.2.6 in Werner (2018)) immer die Banach-Eigenschaft.

Satz 2.1.1

Jeder endlich-dimensionale normierte Raum ist ein Banachraum.

Wir werden in den nächsten beiden Sektionen mit den topologischen Vektorräumen und den Asplundräumen noch zwei weitere wichtige Arten von Räumen einführen. Dafür benötigen wir aber zunächst einige weitere Begrifflichkeiten.

2.1.2 Topologien, Stetigkeitseigenschaften und Konvexität

In der folgenden Sektion werden wir grundlegende Eigenschaften von Funktionen und Mengen definieren, auf die wir später fortwährend zurückgreifen werden. Von besonderer Wichtigkeit sind hierbei die unterschiedlichen Stetigkeitsbegriffe, welche für die Arbeit mit Subdifferentialen unerlässlich sind.

Dabei wird im Folgenden die Menge $\overline{\mathbb{R}}$ der (affin) erweiterten reellen Zahlen eine große Rolle spielen, diese besteht aus den reellen Zahlen zuzüglich zweier Elemente „minus unendlich" ($-\infty$) und „plus unendlich" ($+\infty$), also $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. Für die beiden zusätzlichen Elemente sollen die folgenden üblichen Rechenregeln gelten, wobei a stets eine beliebige (endliche) reelle Zahl darstellt:

Vergleiche und Negation:	$-\infty < a < +\infty$ $-(\pm\infty) = \mp\infty$
Addition/Subtraktion:	$(\pm\infty) + a = a + (\pm\infty) = \pm\infty$ $+\infty + (+\infty) = +\infty$ $-\infty + (-\infty) = -\infty$
Multiplikation:	$(\pm\infty) \cdot a = a \cdot (\pm\infty) = \pm\infty$ für $a > 0$ $(\pm\infty) \cdot a = a \cdot (\pm\infty) = \mp\infty$ für $a < 0$ $(\pm\infty) \cdot (+\infty) = (+\infty) \cdot (\pm\infty) = \pm\infty$ $(\pm\infty) \cdot (-\infty) = (-\infty) \cdot (\pm\infty) = \mp\infty$
Division:	$\frac{a}{\pm\infty} = 0$ $\frac{\pm\infty}{a} = \pm\infty$ für $a > 0$ $\frac{\pm\infty}{a} = \mp\infty$ für $a < 0$

Neben Ausdrücken wie $\frac{a}{0}$, welche bereits auf $\mathbb{R}$ undefiniert sind, gibt es auf $\overline{\mathbb{R}}$ leider zahlreiche weitere Operationen, die nicht definiert werden können ohne grundlegende Rechengesetze oder das Permanenzprinzip zu verletzen, so zum Beispiel $0 \cdot (\pm\infty)$ oder $(+\infty) + (-\infty)$.

Für Funktionen, die auf diese erweiterten reellen Zahlen abbilden, führen wir einige Begriffe ein, die uns im nächsten Unterkapitel 2.2 bei der Konstruktion der Subdifferentiale von Nutzen sein werden (vgl. S. 20 in Boţ, Grad und Wanka (2009)).

Definition 2.1.5

Ist X eine nichtleere Menge und $f\colon X \to \overline{\mathbb{R}}$, dann nennen wir die Menge $\operatorname{dom} f := \{x \in X \mid f(x) < +\infty\}$ *den* ***effektiven Definitionsbereich*** *oder* ***Domain*** *von f. Ist $f(x) > -\infty$ für alle $x \in X$ und* $\operatorname{dom} f \neq \emptyset$, *so heißt f* ***eigentlich****. Weiter ist mittels* $\operatorname{epi} f := \{(x, r) \in X \times \mathbb{R} \mid f(x) \leq r\}$ *der* ***Epigraph*** *von f definiert.*

Wir wollen die Definitionen und Aussagen in dieser Sektion sowie den folgenden Sektionen und Kapiteln stets so allgemein wie möglich formulieren. Deshalb ist es vielfach bereits eine zu starke Forderung, den zugrunde liegenden Raum als metrisch oder gar normiert vorauszusetzen. Aus diesem Grund führen wir zunächst noch die Klasse der topologischen Vektorräume ein, welche entsprechend schwächer strukturiert sind. Die dafür im Folgenden benötigten topologischen Grundbegriffe entnehmen wir den Definitionen 2.1.7 und 2.1.10 in Tammer und Weidner (2020).

Definition 2.1.6

*Es sei T eine nichtleere Menge und $\tau \subseteq 2^T$ eine Menge von Teilmengen von T. Wir nennen τ eine **Topologie** auf T, wenn*

(T1) die Grundmenge T und die leere Menge $\emptyset$ Elemente von τ sind,

(T2) jede Vereinigung von (endlich oder unendlich vielen) Mengen aus τ wieder zu τ gehört und

(T3) jeder Durchschnitt endlich vieler Mengen aus τ ebenfalls wieder zu τ gehört.

*Die Elemente von τ heißen **offene Mengen** und das Tupel (T, τ) **topologischer Raum**. Eine Menge $A \subseteq T$ heißt **abgeschlossen**, wenn $T \setminus A := \{x \in T \mid x \notin A\}$ offen ist.*

Wenn in dieser Arbeit topologische Eigenschaften auf metrischen Räumen genutzt werden, dann ist dies stets im Sinne der metrischen Topologie zu verstehen. Deren Definition mithilfe offener Kugeln entspricht den Definitionen 3.41 und 3.42 bei Manetti (2015).

Definition 2.1.7

*Es sei (M, d_M) ein metrischer Raum. Mit $B_r(x) := \{y \in M \mid d_M(x, y) < r\}$ bezeichnen wir die **offene Kugel** und mit $\overline{B_r(x)} := \{y \in M \mid d_M(x, y) \leq r\}$ die **abgeschlossene Kugel** mit Radius r um $x \in M$.*

*Die von der Metrik d_M auf M induzierte Topologie τ_M nennen wir die **metrische Topologie**. Dabei ist $O \in \tau_M$ für $O \subseteq M$ (O ist also eine offene Menge bezüglich der metrischen Topologie), wenn für jedes $x \in O$ ein $\varepsilon > 0$ existiert, so dass $B_\varepsilon(x) \subseteq O$ gilt.*

Weiterhin verlangen lokale Eigenschaften von Funktionen, wie zum Beispiel die lokale (Lipschitz-)Stetigkeit, ein mathematisch exaktes Verständnis von Umgebungen von Punkten. Wir nutzen dafür Definition 2.1.8 in Tammer und Weidner (2020).

Definition 2.1.8

*Es sei (T, τ) ein topologischer Raum. Eine Menge $U \subseteq T$ heißt **Umgebung** von $x \in T$, falls es eine offene Menge $O \in \tau$ mit $x \in O \subseteq U$ gibt.*

Auf topologischen Räumen können wir, in analoger Weise zu Definition 2.1.16 in Tammer und Weidner (2020), nun bereits eine der für uns wichtigsten Eigenschaften von Funktionen definieren, die Stetigkeit.

Definition 2.1.9

*Es seien (T, τ) und (S, σ) topologische Räume sowie $f : T \to S$. Wir sagen, dass f **stetig** in $x_0 \in T$ ist, wenn zu jeder Umgebung V von $f(x_0)$ eine Umgebung U von x_0 existiert, so dass $f(x) \in V$ für jedes $x \in U$ gilt.*

f heißt stetig (auf T), wenn f in jedem Punkt $x \in T$ stetig ist.

Indem wir für einen Vektorraum die Stetigkeit der ihn definierenden Abbildungen fordern und so eine Verknüpfung zwischen algebraischer und topologischer Struktur schaffen, gelangen wir zum Begriff der topologischen Vektorräume (vgl. Definition 2.1.19 in Tammer und Weidner (2020)).

Definition 2.1.10

*Es sei $(V, +, \cdot)$ ein K-Vektorraum, welcher zudem mit einer Topologie ν versehen ist. Wir nennen das Quadrupel $(V, +, \cdot, \nu)$ einen **topologischen K-Vektorraum**, wenn $+$ und $\cdot$ stetige Operationen bezüglich der Topologie ν sind.*

Genau wie bei den K-Vektorräumen aus Sektion 2.1.1 werden wir auch im Falle der topologischen K-Vektorräume auf die Angabe der Operationen bzw. der Topologie verzichten, wenn diese aus dem Kontext klar ersichtlich sind.

Für unsere Arbeit benötigen wir neben der „normalen“ Stetigkeit aus Definition 2.1.9 noch weitere Abstufungen dieses Begriffes. Diese Stetigkeitseigenschaften sichern uns vor allem die Anwendbarkeit vieler Rechenregeln für die im nächsten Unterkapitel folgenden Subdifferentiale. Die in diesem Zusammenhang wichtigste Art der Stetigkeit ist die Lipschitz-Stetigkeit, wir orientieren uns bei deren Definition an Definition 3.3.2 in Behrends (2004a) und Definition 3.3.4 in Tammer und Weidner (2020).

Definition 2.1.11

*Es seien (M, d_M) und (N, d_N) metrische Räume. Eine Funktion $f\colon M \to N$ heißt **Lipschitz-stetig** auf $M_0 \subseteq M$ mit $M_0 \neq \emptyset$, falls eine reelle Zahl $L \geq 0$ (eine Lipschitz-Konstante) existiert, so dass $d_N(f(x_1), f(x_2)) \leq L \cdot d_M(x_1, x_2)$ für alle $x_1, x_2 \in M_0$ gilt.*

*f heißt **lokal Lipschitz-stetig** auf M_0, wenn es zu jedem $x \in M_0$ eine Umgebung $U \subseteq M$ von x gibt, so dass f auf $M_0 \cap U$ Lipschitz-stetig ist.*

Lipschitz-Stetigkeit ist dabei eine stärkere Anforderung an eine Funktion als Stetigkeit, denn es gilt folgender Zusammenhang (siehe z. B. S. 71 bei Tammer und Weidner (2020) oder Satz 3.3.3 bei Behrends (2004a)):

Lemma 2.1.2

Es seien (M, d_M) und (N, d_N) metrische Räume. Jede Funktion $f\colon M \to N$, welche in einer Umgebung von $x \in M$ lokal Lipschitz-stetig ist, ist stetig in x. Insbesondere ist jede Lipschitz-stetige Funktion auch stetig.

Die Umkehrung dieses Resultates gilt im Allgemeinen nicht.

Die letzte Form der Stetigkeit, welche für uns von Interesse ist, ist die Unter- bzw. Oberhalbstetigkeit einer Funktion. Wir entnehmen diese der Definition 3.3.5 in Tammer und Weidner (2020).

Definition 2.1.12

*Es sei (T, τ) ein topologischer Raum. Eine Funktion $f\colon T \to \overline{\mathbb{R}}$ heißt **unterhalbstetig** in $x_0 \in T$, wenn entweder $f(x_0) = -\infty$ ist oder wenn für jedes $h \in \mathbb{R}$ mit $h < f(x_0)$ eine Umgebung $U \subseteq T$ von x_0 existiert, so dass $h < f(x)$ für alle $x \in U$ gilt.*

*f heißt **oberhalbstetig** in x_0, wenn $-f$ unterhalbstetig in x_0 ist, wenn also entweder $f(x_0) = +\infty$ ist oder wenn für jedes $h \in \mathbb{R}$ mit $h > f(x_0)$ eine Umgebung $U \subseteq T$ von x_0 existiert, so dass $h > f(x)$ für alle $x \in U$ gilt.*

Wir nennen die Funktion f unterhalbstetig bzw. oberhalbstetig, wenn f unterhalbstetig bzw. oberhalbstetig in jedem Punkt $x \in T$ ist.

Man sieht leicht den Zusammenhang zwischen Unter- bzw. Oberhalbstetigkeit und Stetigkeit, dies ist Lemma 3.3.7(b) in Tammer und Weidner (2020):

Lemma 2.1.3

Es sei (T, τ) ein topologischer Raum und $x \in T$. Die Funktion $f\colon T \to \overline{\mathbb{R}}$ ist genau dann stetig in x, wenn f sowohl unterhalbstetig als auch oberhalbstetig in x ist. f ist genau dann stetig, wenn f sowohl unterhalbstetig als auch oberhalbstetig ist.

Zu jeder der soeben eingeführten Arten von Stetigkeit gibt es mehrere äquivalente Charakterisierungen, auf welche wir in dieser Arbeit aber nicht weiter eingehen werden. Hierfür sei stattdessen auf die einschlägige Literatur zu diesem Thema verwiesen, beispielsweise auf die Bücher von Boţ, Grad und Wanka (2009) oder von Tammer und Weidner (2020).

Abbildung 2.2 veranschaulicht nochmals die verschiedenen Stetigkeitsbegriffe. Die links abgebildete Funktion $f(x) = \begin{cases} -1 & \text{für } x \leq 0 \\ 1 & \text{für } x > 0 \end{cases}$ ist unterhalbstetig, aber nicht oberhalbstetig und damit in ihrer Gesamtheit auch weder stetig noch Lipschitz-stetig. Sie ist allerdings stetig und lokal Lipschitz-stetig in jedem Punkt $x \neq 0$.

Hingegen ist die rechts abgebildete Funktion $f(x) = x^2$ zwar stetig (und damit auch unter- und oberhalbstetig nach Lemma 2.1.3), aber nicht Lipschitz-stetig, da ihr Anstieg nicht durch eine Lipschitz-Konstante beschränkt werden kann. Diese Funktion ist aber trotzdem auf jeder Teilmenge ihres Definitionsbereiches $\mathbb{R}$ lokal Lipschitz-stetig.

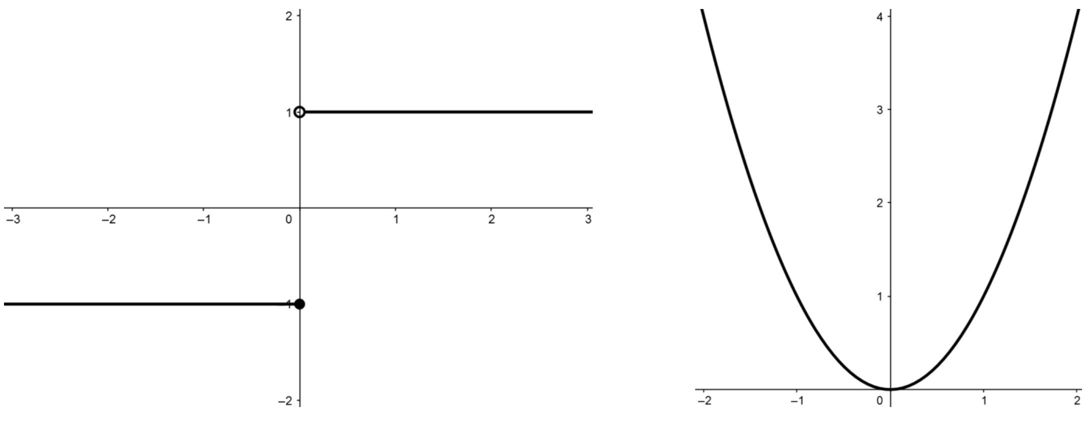

Abbildung 2.2: Beispielfunktionen zu Stetigkeitseigenschaften: Unterhalbstetigkeit (links) und lokale Lipschitz-Stetigkeit (rechts) (eigene Darstellung mit GeoGebra)

Da Metriken und Normen einen zentralen Bestandteil dieser Arbeit bilden, ist es natürlich von besonderem Interesse für uns, die Stetigkeitseigenschaften dieser Funktionen zu untersuchen.

Bemerkung 2.1.2

In jedem normierten Raum $(X, \|\cdot\|_X)$ *gilt für beliebige* $x, y \in X$ *die umgekehrte Dreiecksungleichung* $|\|x\|_X - \|y\|_X| \leq \|x-y\|_X$. *Auch in jedem metrischen Raum* (M, d_M) *gilt stets* $|d_M(x,y) - d_M(y,z)| \leq d_M(x,z)$ *für alle* $x, y, z \in M$. *Das liefert aber gerade, dass alle Abstandsfunktionen und Normen Lipschitz-stetig mit Lipschitz-Konstante* 1 *sind.*

Als letzte Eigenschaft von Funktionen wollen wir in dieser Sektion noch Konvexität definieren, denn diese nimmt, auch durch ihre Abwesenheit in unseren Aufgabenstellungen, in dieser Arbeit eine wesentliche Rolle ein. Wir orientieren uns dabei an Definition 7.2.1 in Behrends (2004b) und ergänzen die Definition einer konvexen Menge von S. 262 im selben Buch.

Definition 2.1.13

Eine Teilmenge U *eines* K*-Vektorraumes* $(V, +, \cdot)$ *heißt* ***konvex****, wenn für alle* $x, y \in U$ *und* $\lambda \in [0,1]$ *auch* $\lambda x + (1-\lambda)y \in U$ *gilt.*

Auf einer konvexen Menge $U \subseteq V$ *nennen wir die Funktion* $f\colon U \to \overline{\mathbb{R}}$ ***konvex****, wenn die Ungleichung* $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$ *für alle* $x, y \in U$ *und* $\lambda \in [0,1]$ *erfüllt ist.*

f *heißt* ***konkav****, falls* $-f$ *konvex ist.*

2.1.3 Asplund-Räume

Gewisse Eigenschaften der verallgemeinerten Subdifferentiale aus Unterkapitel 2.2 gelten nicht in allen Banachräumen, sondern nur in den noch spezielleren Asplund-Räumen. Um diese im Folgenden einführen zu können, wollen wir zunächst angeben, was unter Dichtheit (Definition 3.1.11 in Behrends (2004a)) zu verstehen ist.

Definition 2.1.14

Es sei (M, d_M) *ein metrischer Raum. Eine Teilmenge* D *von* M *heißt* ***dicht*** *in* M*, wenn* M *die kleinste abgeschlossene Menge ist, welche* D *enthält.*

Eine alternative Charakterisierung (vgl. S. 30 in Werner (2018)) ist, dass D genau dann dicht in M ist, wenn jede nichtleere offene Teilmenge von M mindestens ein Element von D enthält.

Des Weiteren benötigen wir noch den Begriff der Fréchet-Differenzierbarkeit, dieser findet sich z. B. als Definition 1.12 in Phelps (1993).

Definition 2.1.15

Es seien $(X, \|\cdot\|_X)$ *und* $(Y, \|\cdot\|_Y)$ *normierte Räume sowie* U *eine offene Teilmenge von* X. *Eine stetige Funktion* $f\colon U \to Y$ *heißt* ***Fréchet-differenzierbar*** *in* $x_0 \in U$*, wenn es eine stetige lineare Abbildung* $f'(x_0)\colon X \to Y$ *gibt, so dass zu jedem* $\varepsilon > 0$ *ein* $\delta > 0$ *existiert mit*

$$\|f(x_0 + x) - f(x_0) - f'(x_0)(x)\|_Y \leq \varepsilon \|x\|_X$$

für alle $x \in X$ *mit* $\|x\|_X < \delta$. *Die dadurch eindeutig bestimmte Funktion* $f'(x_0)$ *nennen wir die* ***Fréchet-Ableitung*** *von* f *in* x_0.

Damit können wir jetzt zunächst die geläufigste Definition der Asplund-Räume angeben, welche auch von Göpfert, Riedrich und Tammer (2017) als Definition 3.1.19 genutzt wird.

Definition 2.1.16

*Ein Banachraum $(X, \|\cdot\|_X)$ heißt **Asplund-Raum**, wenn jede konvexe stetige Funktion $f: U \to \mathbb{R}$, wobei U eine offene konvexe Teilmenge von X ist, Fréchet-differenzierbar auf einer dichten Teilmenge von U ist.*

Mit dieser Definition lässt sich allerdings eher schlecht arbeiten, weswegen wir noch eine weitere der vielen äquivalenten Darstellungen von Asplund-Räumen angeben wollen. Diese basiert auf der Betrachtung des stetigen Dualraumes, dessen Definition entspricht Definition 3.3 in Göpfert, Riedrich und Tammer (2009).

Definition 2.1.17

*Es sei $(X, \|\cdot\|_X)$ ein normierter Raum. Die Menge aller (stetigen) linearen Abbildungen auf X nennen wir den zu X gehörigen **(stetigen) Dualraum** X^*.*

Der stetige Dualraum X^* zu einem normierten Raum $(X, \|\cdot\|_X)$ wird üblicherweise mit der Operatornorm

$$\|x^*\|_* = \sup_{\substack{x \in X \\ \|x\|_X \leq 1}} |x^*(x)|$$

versehen. Dadurch wird $(X^*, \|\cdot\|_*)$ ebenfalls zu einem normierten Raum, sogar zu einem Banachraum (siehe Satz 3.4 in Göpfert, Riedrich und Tammer (2009)).

Für die folgenden Resultate benötigen wir außerdem Separabilität und führen diese daher analog zu Definition I.2.9 bei Werner (2018) ein.

Definition 2.1.18

*Ein topologischer Raum (T, τ) heißt **separabel**, wenn es eine höchstens abzählbare Teilmenge von T gibt, die in diesem Raum dicht liegt.*

Nun können wir Asplund-Räume wie gewünscht alternativ charakterisieren, diese Aussage findet sich z. B. als Theorem 2.34 in Phelps (1993).

Satz 2.1.2

Ein Banachraum $(X, \|\cdot\|_X)$ ist genau dann ein Asplund-Raum, wenn jeder separable abgeschlossene Unterraum W von X einen separablen stetigen Dualraum W^* besitzt.

Für separable Banachräume gilt sogar das folgende, etwas schärfere Resultat, welches ebenfalls im Buch von Phelps (1993) zu finden ist, dort als Theorem 2.19.

Satz 2.1.3

Ein separabler Banachraum $(X, \|\cdot\|_X)$ ist genau dann ein Asplund-Raum, wenn X^* separabel ist.

Die Zielfunktionen der in dieser Arbeit betrachteten und in Kapitel 1 präsentierten Aufgabenstellungen sind immer Abbildungen aus dem $\mathbb{R}^n$. Als Schlussfolgerung aus Satz 2.1.3 können wir nun direkt eine Aussage über dessen Asplund-Eigenschaft treffen.

Korollar 2.1.1

$\mathbb{R}^n$ ist ein Asplund-Raum.

Beweis: $\mathbb{R}^n$ ist separabel, da $\mathbb{Q}^n$ abzählbar ist und dicht in $\mathbb{R}^n$ liegt. Damit ist auch $(\mathbb{R}^n)^* = \mathbb{R}^n$ separabel und die Asplund-Eigenschaft folgt sofort aus Satz 2.1.3. ■

Interessant ist nun noch die Frage, welche der weiteren häufig untersuchten Banachräume die Asplund-Eigenschaft besitzen und welche nicht. Khan, Tammer und Zălinescu (2015) geben in Bemerkung 3.5.4 an, dass neben den von Satz 2.1.3 erfassten Banachräumen auch jeder reflexive Banachraum (das heißt, jeder Banachraum der seinem Bidualraum entspricht) ein Asplund-Raum ist. Beispielsweise sind die Folgenräume c_0 und l^p sowie die Funktionenräume $L^p[0,1]$ für $1 < p < +\infty$ Asplund-Räume, l^1 und l^∞ sind hingegen keine.

2.1.4 Lösungskonzepte und Skalarisierung

Diese letzte Sektion zu den algebraischen und analytischen Grundlagen beschäftigt sich vor allem mit den Begriffen, welche zur Behandlung des mehrkriteriellen Problems (MP) unverzichtbar sind. Da im vektoriellen Fall zwei Elemente im Allgemeinen nicht mehr direkt miteinander vergleichbar sind, benötigen wir zunächst einen geeigneten Lösungsbegriff. Außerdem wollen wir uns mit Skalarisierung beschäftigen, um Beziehungen zwischen Lösungen der mehrkriteriellen Aufgabe und solchen des skalaren Medianproblems herstellen zu können.

Im skalaren Fall unterscheiden wir im Wesentlichen nur zwischen lokalen und globalen Minimalstellen einer Funktion, deren Definitionen wir zunächst rekapitulieren wollen (vgl. S. 35/36 in Forst und Hoffmann (2010)).

Definition 2.1.19

*Es sei $(V, +, \cdot, \nu)$ ein topologischer K-Vektorraum und $f : V \to \overline{\mathbb{R}}$. Wir sagen, dass $x_0 \in V$ ein **globales Minimum** von f ist, wenn $f(x_0) \leq f(x)$ für alle $x \in V$ gilt.*

*Gibt es eine Umgebung $U \subseteq V$ von x_0 mit $f(x_0) \leq f(x)$ für alle $x \in U$, so nennen wir x_0 ein **lokales Minimum** von f.*

Für mehrkriterielle Problemstellungen gibt es hingegen deutlich mehr unterschiedliche Lösungsbegriffe. Am häufigsten wird dabei das Konzept der (Pareto-)Minimalität genutzt, beispielsweise auch von Gerth und Weidner (1990) in ihrer wegweisenden Arbeit zur Skalarisierung.

Zum Rechnen mit Mengen und mengenwertigen Funktionen benötigen wir zunächst noch eine Addition und eine Skalarmultiplikation und folgen dabei Definition 1.2 bei Jahn (2004). Sind also A und B zwei Teilmengen eines Vektorraumes $(V, +, \cdot)$ über dem Körper $\mathbb{R}$ sowie $t \in \mathbb{R}$, dann definieren wir $A + B := \{a + b \mid a \in A,\, b \in B\}$ (insbesondere $a + B := \{a + b \mid b \in B\}$ für jedes $a \in V$), $A - B := \{a - b \mid a \in A,\, b \in B\}$ und $t \cdot A := \{t \cdot a \mid a \in A\}$.

Die folgende Definition der Minimalität im vektoriellen Fall entspricht Definition 3.1 bei Gerth und Weidner (1990) oder Definition 2.4.1 bei Khan, Tammer und Zălinescu (2015).

Definition 2.1.20

Es seien $(W, \oplus, \odot, \omega)$ ein topologischer K-Vektorraum und $D \subset W$, $F \subseteq W$ nichtleere Mengen. Ein Punkt $y_0 \in F$ heißt ***minimales Element*** *von F bezüglich D, wenn für jeden Punkt $y \in F$ mit $y_0 \in y + D$ schon $y = y_0$ folgt. Die Menge aller minimalen Elemente von F bezüglich D bezeichnen wir mit $\mathrm{Min}(F, D)$.*

Ist auch $(V, +, \cdot, \nu)$ ein topologischer K-Vektorraum, $A \subseteq V$ und $f \colon V \to W$, dann nennen wir $x_0 \in A$ ein ***Minimum*** *von f über A bezüglich D, wenn $f(x_0) \in \mathrm{Min}(f(A), D)$ erfüllt ist.*

Wir wollen für diese Arbeit dennoch ein anderes Lösungskonzept nutzen, nämlich das der schwachen Minimalität. Gegenüber anderen Minimalitäts- oder Effizienzbegriffen liegt der Vorteil der schwachen Minimalität darin, dass die Menge aller schwachen Minima einer Funktion oft zusammenhängend ist oder andere nützliche Eigenschaften wie Abgeschlossenheit besitzt. Des Weiteren werden wir in Kapitel 7 sehen, dass genau die schwachen Minima des mehrkriteriellen Problems (MP) mittels nichtlinearer Skalarisierung wesentlich mit den notwendigen Optimalitätsbedingungen des skalaren Medianproblems (SP) verknüpft sind.

Vereinfacht dargestellt ist ein Element einer Menge genau dann schwach minimal, wenn kein anderes Element dieser Menge in jeder Komponente „besser“ ist. Eine exakte Definition des Begriffes entnehmen wir Definition 3.2 bei Gerth und Weidner (1990) oder Definition 2.4.2 bei Khan, Tammer und Zălinescu (2015).

Definition 2.1.21

Es seien $(W, \oplus, \odot, \omega)$ ein topologischer K-Vektorraum, $D \subset W$ eine Menge mit nichtleerem algebraischen Inneren sowie $F \subseteq W$ nichtleer. Ein Punkt $y_0 \in F$ heißt ***schwach minimales Element*** *von F bezüglich D, wenn kein Punkt $y \in F$ existiert, für welchen $y_0 \in y + \operatorname{int} D$ gilt. Die Menge aller schwach minimalen Elemente von F bezüglich D bezeichnen wir mit $\mathrm{WMin}(F, D)$.*

Ist auch $(V, +, \cdot, \nu)$ ein topologischer K-Vektorraum, $A \subseteq V$ und $f \colon V \to W$, dann nennen wir $x_0 \in A$ ein ***schwaches Minimum*** *von f über A bezüglich D, wenn $f(x_0) \in \mathrm{WMin}(f(A), D)$ erfüllt ist.*

Gilt in den obigen Definitionen $0 \notin \operatorname{int} D$ für die Ordnungsmenge D, dann folgt wegen $\operatorname{int} D \subseteq D \setminus \{0\}$ sofort, dass jedes minimale Element auch schwach minimal ist, insbesondere sind in der Menge aller schwach minimalen Elemente also stets alle minimalen Elemente enthalten. Die Forderung $0 \notin \operatorname{int} D$ ist dabei praktisch keine Einschränkung und wird von jeder relevanten Ordnungsmenge (z. B. auch von jedem Ordnungskegel) erfüllt. Das folgende Lemma fasst diese Aussage nochmals zusammen, vgl. Proposition 2.4.5 (iii) in Khan, Tammer und Zălinescu (2015).

Lemma 2.1.4

Ist $(W, \oplus, \odot, \omega)$ ein topologischer K-Vektorraum, $D \subset W$ eine Menge mit nichtleerem algebraischen Inneren und $0 \notin \operatorname{int} D$ sowie $F \subseteq W$ nichtleer, dann gilt $\mathrm{Min}(F, D) \subseteq \mathrm{WMin}(F, D)$.

Wir führen als Nächstes das nichtlineare Skalarisierungsfunktional ein, welches wir für die Herleitung von Optimalitätsbedingungen in Kapitel 7 nutzen werden. Dieses

Funktional wurde erstmals in der Dissertation von Gerstewitz (1984) zur Herleitung von Trennungssätzen für nicht notwendigerweise konvexe Mengen beschrieben. Zudem wurden in dieser Arbeit auch entsprechende Eigenschaften des Funktionals zur Skalarisierung von Vektoroptimierungsproblemen diskutiert und in späteren Veröffentlichungen vertieft, siehe dazu Gerstewitz und Iwanow (1985) und Gerth und Weidner (1990). Aus diesem Grund wird das nichtlineare Skalarisierungsfunktional in der Literatur häufig nur als Gerstewitz-Funktional bezeichnet. Auch der Name Pascoletti-Serafini-Skalarisierung wird teilweise verwendet, da Pascoletti und Serafini (1984) ein skalares Ersatzproblem diskutiert haben, welches in seinen Grundzügen mit der Skalarisierung übereinstimmt, die durch das Gerstewitz-Funktional beschrieben wird.

Die in unserer Arbeit verwendete Definition des nichtlinearen Skalarisierungsfunktionals findet sich z. B. auf S. 6 in Bao, Hillmann und Tammer (2017) oder auf S. 39/40 in Göpfert, Riahi, Tammer und Zălinescu (2003).

Definition 2.1.22

Es seien $(V, +, \cdot, \nu)$ ein topologischer $\mathbb{R}$-Vektorraum, A eine nichtleere Teilmenge von V und $k \in V \setminus \{0\}$ ein Richtungsvektor. Wir definieren die Funktion $\varphi_{A,k} \colon V \to \overline{\mathbb{R}}$ mittels

$$\varphi_{A,k}(y) := \inf \{t \in \mathbb{R} \mid y \in tk - A\} \tag{2.1}$$

und nennen sie ein ***nichtlineares Skalarisierungsfunktional*** *bezüglich der Menge A und der Richtung k.*

Bemerkenswert an der Definition des Skalarisierungsfunktionals ist, dass zunächst keine weiteren Anforderungen an die Menge A gestellt werden. Diese Menge kann also beispielsweise hochgradig nichtkonvex sein und das nichtlineare Skalarisierungsfunktional ist trotzdem wohldefiniert.

Allerdings werden wir später in Lemma 2.1.5 erkennen, dass viele wichtige Eigenschaften des Skalarisierungsfunktionals dennoch so stark mit Eigenschaften der Menge A korrelieren, dass gewisse Forderungen an A unter praktischen Gesichtspunkten unumgänglich sind.

Für einige Eigenschaften des Skalarisierungsfunktionals und für dessen Anwendung in den folgenden Kapiteln ist der Begriff des Kegels relevant, wir nutzen hier die auch von Khan, Tammer und Zălinescu (2015) in Definition 2.1.9 verwendete Charakterisierung.

Definition 2.1.23

Es sei $(V, +, \cdot)$ ein $\mathbb{R}$-Vektorraum. Eine nichtleere Menge $C \subseteq V$ heißt ***Kegel****, wenn für alle $c \in C$ und $\lambda \in \mathbb{R}_+$ auch $\lambda c \in C$ gilt.*

Wir nennen den Kegel C

a) ***konvex****, wenn aus $c_1, c_2 \in C$ stets auch $c_1 + c_2 \in C$ folgt,*

b) ***eigentlich****, wenn $C \neq \{0\}$ und $C \neq V$ ist und*

c) ***spitz****, wenn $C \cap -C = \{0\}$ gilt.*

Man sieht leicht, dass ein Kegel C genau dann konvex ist, wenn C eine konvexe Menge ist.

Abbildung 2.3 veranschaulicht das nichtlineare Skalarisierungsfunktional aus Definition 2.1.22 im Raum $\mathbb{R}^2$. Als Menge A, bezüglich welcher das Skalarisierungsfunktional definiert ist, fungiert hier mit $A = \mathbb{R}^2_+ := \{(a,b) \in \mathbb{R}^2 \mid a \geq 0, b \geq 0\}$ der gewöhnliche Ordnungskegel in $\mathbb{R}^2$ und als Richtungsvektor $k = (1,2)$.

Anschaulich erhält man dabei durch Verschiebung der Menge $-A$ entlang des Strahls tk verschiedene Niveaulinien des Funktionals, also Mengen von Punkten mit demselben Funktionswert. Davon sind in Abbildung 2.3 zwei Niveaulinien eingetragen: Die Punkte $\tilde{x}$, $\overline{x}$ und $\hat{x}$ liegen auf der Niveaulinie des Funktionswertes t_x und die Punkte $\tilde{y}$, $\overline{y}$ und $\hat{y}$ werden vom Skalarisierungsfunktional alle auf den Wert t_y abgebildet.

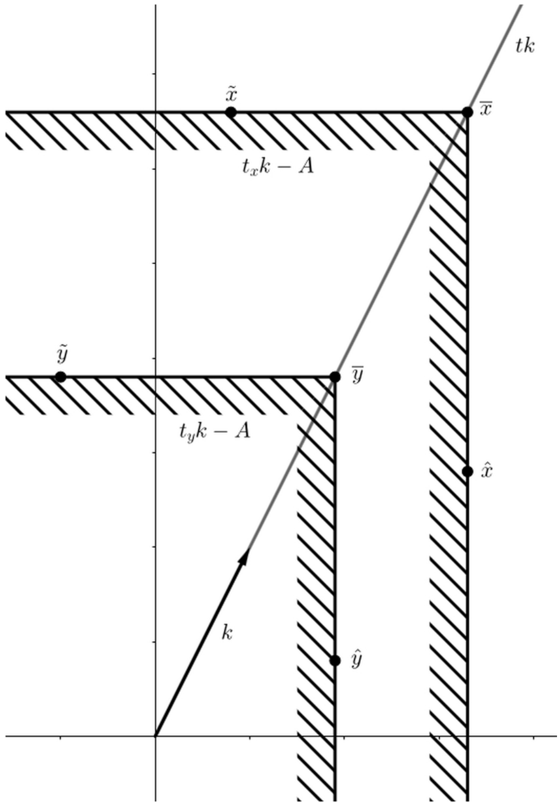

Abbildung 2.3: Verschiedene Niveaulinien des Skalarisierungsfunktionals in $\mathbb{R}^2$ bezüglich der Menge $\mathbb{R}^2_+$ und Richtung $k = (1,2)$ (eigene Darstellung mit GeoGebra)

Zum Abschluss dieser Sektion können wir nun noch eine Vielzahl nützlicher Eigenschaften des Skalarisierungsfunktionals angeben, diese finden sich als Theorem 2.3.1 in Göpfert, Riahi, Tammer und Zălinescu (2003).

Lemma 2.1.5 (*Eigenschaften des Skalarisierungsfunktionals*)

Es seien $(V, +, \cdot, \nu)$ ein topologischer $\mathbb{R}$-Vektorraum, $A \subset V$ nichtleer und abgeschlossen und $k \in V \setminus \{0\}$ ein Richtungsvektor, so dass der von k erzeugte Strahl in A liegt, das heißt, so dass

$$A + [0, +\infty) \cdot k \subseteq A \tag{2.2}$$

gilt.

Dann gelten für das mittels (2.1) definierte Funktional $\varphi_{A,k}$ die folgenden Aussagen:

a) *$\varphi_{A,k}$ ist unterhalbstetig mit Domain $\operatorname{dom}\varphi_{A,k} = \mathbb{R}k - A$, für alle $r \in \mathbb{R}$ sind die Niveaumengen des Funktionals gegeben durch*

$$\{y \in V \mid \varphi_{A,k}(y) \leq r\} = rk - A$$

und für Verschiebungen in Richtung k gilt $\varphi_{A,k}(y + \lambda k) = \varphi_{A,k}(y) + \lambda$ für beliebige $y \in V$ und $\lambda \in \mathbb{R}$.

b) *$\varphi_{A,k}$ ist konvex genau dann, wenn A konvex ist.*

c) *$\varphi_{A,k}$ ist positiv homogen, das heißt, es gilt $\varphi_{A,k}(\lambda y) = \lambda\varphi_{A,k}(y)$ für alle $y \in V$ und $\lambda > 0$ genau dann, wenn A ein Kegel ist.*

d) *$\varphi_{A,k}$ ist eigentlich genau dann, wenn A keine zu k parallelen Geraden enthält, wenn also für alle $y \in V$ ein $t \in \mathbb{R}$ mit $y + tk \notin A$ existiert.*

e) *$\varphi_{A,k}$ nimmt genau dann nur endliche Werte an, wenn A keine zu k parallelen Geraden enthält und zusätzlich $\mathbb{R}k - A = V$ gilt.*

f) *Für jede Menge $B \subset V$ ist $\varphi_{A,k}$ B-monoton, das heißt, aus $y_2 - y_1 \in B$ folgt stets $\varphi_{A,k}(y_1) \leq \varphi_{A,k}(y_2)$, genau dann, wenn $A + B \subseteq A$ gilt.*

g) *$\varphi_{A,k}$ ist subadditiv, das heißt, es gilt $\varphi_{A,k}(y_1 + y_2) \leq \varphi_{A,k}(y_1) + \varphi_{A,k}(y_2)$, genau dann, wenn $A + A \subseteq A$ gilt.*

Erfüllen A und k außerdem die Forderung

$$A + (0, +\infty) \cdot k \subseteq \operatorname{int} A, \tag{2.3}$$

so erhält man weiterhin:

h) *$\varphi_{A,k}$ ist stetig und die Niveaulinien des Funktionals sind für alle $r \in \mathbb{R}$ gegeben durch*

$$\{y \in V \mid \varphi_{A,k}(y) = r\} = rk - \operatorname{bd} A.$$

2.2 Verallgemeinerte Differentiation

Das zweite Unterkapitel zu den Grundlagen dieser Arbeit befasst sich nun mit verallgemeinerter Differentiation. Da unser ganzer Ansatz zur Herleitung von Optimalitätsbedingungen auf der Berechnung und Nutzung verallgemeinerter Subdifferentiale für nicht notwendigerweise konvexe Funktionen beruht, ist es unerlässlich, diese vorab gründlich zu studieren.

Dazu werden wir in Sektion 2.2.1 zunächst eine kurze Einführung in die Thematik geben und darlegen, was unter Subdifferentialen sowohl im konvexen als auch im nichtkonvexen Fall generell zu verstehen ist. Außerdem gehen wir darauf ein, wieso sich unter den vielen Subdifferentialkonzepten genau die beiden von Ioffe und Kruger/Mordukhovich als für unsere Zwecke am besten geeignet erwiesen haben.

Des Weiteren sollen dann natürlich diese beiden verallgemeinerten Subdifferentiale ausführlich vorgestellt werden. Neben der Herleitung und Konstruktion des jeweiligen Subdifferentials beinhalten die Sektionen 2.2.2 und 2.2.3 auch einen Überblick über deren geschichtliche Entstehung sowie wichtige Resultate des einhergehenden Subdifferentialkalküls. Eben dieser Kalkül ermöglicht es uns erst, die Berechnungen der Subdifferentiale auch in unseren Aufgabenstellungen anzuwenden.

Abschließen wird dieses Unterkapitel eine kurze Gegenüberstellung der beiden Subdifferentiale in Sektion 2.2.4, gerade auch im für uns besonders interessanten Vektorraum $\mathbb{R}^n$.

2.2.1 Konzept und Theorie der Subdifferentiale

Bereits in der Schule lernt man, über die Nullstellen der Ableitungsfunktion Minima und Maxima von Funktionen zu bestimmen. Viele Funktionen sind aber gar nicht, nicht überall oder nicht an den für die jeweilige Anwendung wichtigen Funktionswerten differenzierbar, womit diese Herangehensweise im Allgemeinen nicht mehr zielführend ist. Zu den nicht (überall) differenzierbaren Funktionen gehören dabei schon so einfache Funktionen wie die Betragsfunktion auf den reellen Zahlen, welche genau in ihrem globalen Minimum $x_0 = 0$ nicht differenzierbar ist.

Eine Verallgemeinerung des Ableitungsbegriffes ist daher unumgänglich und geschieht mittels der Subgradienten bzw. Subableitungen. Im klassischen Sinne sind diese dabei zunächst nur für konvexe Funktionen definiert, wir orientieren uns hier an Definition 3.1.13 in Göpfert, Riedrich und Tammer (2017).

Definition 2.2.1

Es seien $(X, \|\cdot\|_X)$ ein normierter Raum, $f\colon X \to \overline{\mathbb{R}}$ eine konvexe Funktion und $x_0 \in X$ mit $|f(x_0)| \neq +\infty$. Dann nennen wir die Menge

$$\partial f(x_0) := \{x^* \in X^* \mid \forall\, x \in X : x^*(x - x_0) \leq f(x) - f(x_0)\}$$

*das **Subdifferential** von f an der Stelle x_0 und die einzelnen Elemente von $\partial f(x_0)$ **Subgradienten** oder **Subableitungen** von f an der Stelle x_0.*

Ist hingegen $|f(x_0)| = +\infty$, so setzen wir $\partial f(x_0) := \emptyset$.

Dieses Subdifferential wird manchmal auch als „klassisches“ Subdifferential oder Fenchel-Subdifferential bezeichnet. Subgradienten wurden erstmals in den 1960er Jahren eingeführt, vorrangig von Moreau (1962, 1963) und Rockafellar (1963, 1964), eine sehr gute Übersicht über das Thema liefert auch Rockafellar (1970).

Abbildung 2.4 zeigt die Betragsfunktion $|\cdot|$ auf den reellen Zahlen in schwarz sowie rot gestrichelt einige Elemente ihres Subdifferentials an der Stelle $x_0 = 0$. Man sieht gut, warum die Bezeichnung Subgradienten sinnvoll ist: Während die Ableitung einer Funktion in einem Punkt x_0 die Tangente dieser Funktion an x_0 bildet und in x_0 denselben Anstieg wie die Funktion hat, bilden die Elemente des Subdifferentials gerade die Geraden, welche durch x_0 und ansonsten vollständig unterhalb des Graphen bzw. höchstens auf dem Graphen der gegebenen Funktion verlaufen (wie hier die Geraden mit Anstieg -1 bzw. 1).

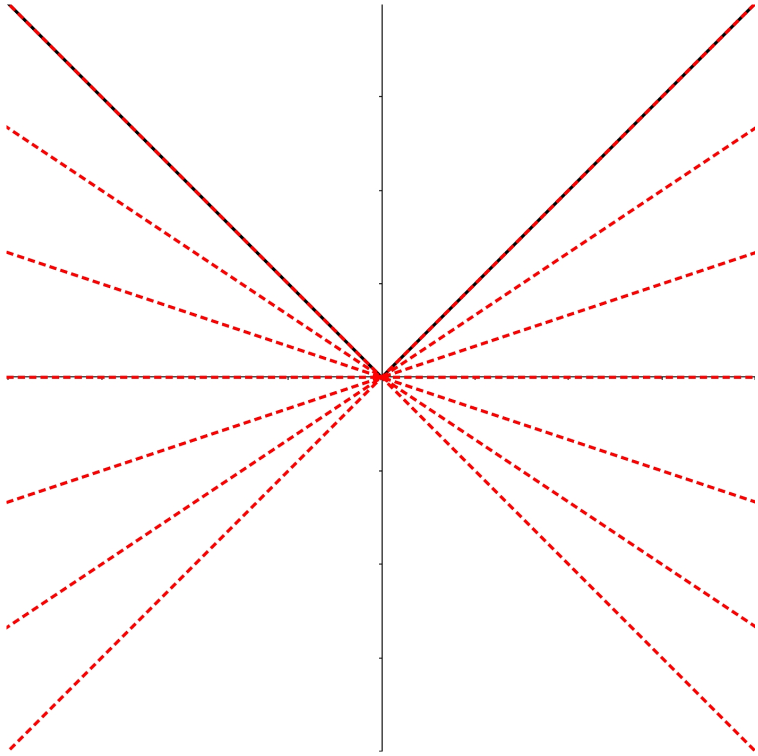

Abbildung 2.4: Subgradienten der Betragsfunktion in $\mathbb{R}$ an der Stelle $x_0 = 0$ (eigene Darstellung mit GeoGebra)

Der folgende Satz, zu finden als Theorem 25.1 in Rockafellar (1970), besagt, dass im Falle differenzierbarer Funktionen Subgradient und Gradient übereinstimmen und somit das Subdifferential eine natürliche Erweiterung des Differentials darstellt.

Satz 2.2.1

Es seien $(X, \|\cdot\|_X)$ ein normierter Raum, $f\colon X \to \overline{\mathbb{R}}$ konvex und $x_0 \in X$, so dass $f(x_0)$ endlich ist. Es ist f genau dann differenzierbar in x_0, wenn $\nabla f(x_0)$ der einzige Subgradient von f an der Stelle x_0 ist.

Es gibt für das Subdifferential für konvexe Funktionen einen umfangreichen Rechenkalkül, aus welchem für uns besonders die Summenregel wichtig ist, diese findet sich als Satz 3.1.2 im Buch von Göpfert, Riedrich und Tammer (2017).

Satz 2.2.2 (*Summenregel für konvexe Subdifferentiale*)

Es seien $(X, \|\cdot\|_X)$ ein normierter Raum und $f_1, \ldots, f_k\colon X \to \overline{\mathbb{R}}$ mit $k \geq 2$ konvexe Funktionale. Weiter existiere ein $x_0 \in X$, so dass $f_1, \ldots, f_k$ in x_0 nur endliche Werte annehmen und mindestens $k-1$ der Funktionale in x_0 stetig sind. Dann gilt

$$\partial \left(\sum_{i=1}^{k} f_i \right)(x) = \sum_{i=1}^{k} \partial f_i(x)$$

für alle $x \in X$.

Für die Optimierung sind Subdifferentiale deswegen so wichtig, weil sie das Extremalprinzip zur Bestimmung der Minima einer Funktion von der Klasse der differenzierbaren Funktionen auf die Klasse aller konvexen Funktionen verallgemeinern. Diese Aussage entspricht Satz 3.1.3 in Göpfert, Riedrich und Tammer (2017).

Satz 2.2.3

Es seien $(X, \|\cdot\|_X)$ ein normierter Raum und $f\colon X \to \overline{\mathbb{R}}$ eine eigentliche und konvexe Funktion. Dann ist $x_0 \in X$ genau dann eine (globale) Minimallösung des Problems $f(x) \to \min_{x \in X}$, wenn $0 \in \partial f(x_0)$ ist.

Leider gehören aber viele der Funktionen, welche besondere Relevanz für praktische Anwendungen haben, auch nicht zur Klasse der konvexen Funktionen. Zu diesen nichtkonvexen Funktionen zählen unter anderem die mit einem negativen Vorzeichen versehenen Abstandsfunktionen und Normen, welche auch in dieser Arbeit von höchster Bedeutung sind.

Deswegen wurde in den nachfolgenden Jahren zunehmend versucht, den Subdifferentialbegriff entsprechend zu erweitern. Die ersten verallgemeinerten Subgradienten etablierte Clarke (1973, 1975). Eine vollständige Übersicht über das Clarke-Subdifferential mitsamt dem zugehörigen Kalkül findet man in Clarke (1983), bis heute ist dies eines der am besten studierten und meistgenutzten Subdifferentiale im nichtkonvexen Fall. Wir werden aber noch darauf eingehen, wieso es für unseren Ansatz dennoch nicht geeignet ist.

In der Folge entwickelten sich bis heute mehrere Hundert verschiedene Subdifferentiale für den nichtkonvexen Fall, die sich teilweise nur geringfügig voneinander unterscheiden und deren Beziehungen untereinander fast unüberschaubar sind. Während es für konvexe Funktionen genau ein Subdifferential im konvexen Sinne gibt, ist im nichtkonvexen Szenario nicht einmal genau umrissen, was unter einem Subdifferential eigentlich zu verstehen ist. Selbst Ioffe, der seit Jahrzehnten zu den führenden Forschern auf diesem Gebiet gehört und der eines der wichtigsten Subdifferentiale geprägt hat, musste eingestehen: „It is probably hopeless (and likely unnecessary) to find an appropriate formal definition of the concept.“ (Ioffe (2017), S. 151).

Im Wesentlichen lassen sich die verallgemeinerten Subdifferentiale jedoch in zwei Gruppen einteilen: Die erste Gruppe besteht dabei aus den Subdifferentialen, die an ein verallgemeinertes Differential angelehnt sind und aus diesem meist durch Abschwächung einer Gleichheitsbedingung zu einer Ungleichung hervorgehen. Alle übrigen Subdifferentiale der zweiten Gruppe entspringen hingegen mehr oder weniger komplex konstruierten Normalenkegeln an den Epigraphen des Funktionals, Grenzwertbetrachtungen oder Kombinationen aus beidem.

Zu den wichtigsten Vertretern der erstgenannten Gruppe gehören, neben dem bereits erwähnten Clarke-Subdifferential, das Fréchet-Subdifferential, das Dini-Hadamard-Subdifferential (welches in der Literatur teils auch nur als Dini-Subdifferential oder Hadamard-Subdifferential zu finden ist) und das Michel-Pernot-Subdifferential. Der größte Vorteil all dieser Subdifferentiale ist es, dass sie sich relativ einfach berechnen lassen, meist sogar für die aus praktischen Problemstellungen entstehenden Zielfunktionen. Damit einher geht allerdings auch ihr wesentlicher Nachteil, denn für diese Berechnungen nutzt man nur sehr wenig von der Struktur der zugrunde liegenden Funktion. Darum sind diese Subdifferentiale meist schon für einfache Funktionen auf großen Teilen ihres Domains leer, wie Ioffe (2000) auf S. 533 anmerkt. Weiterhin fehlen diesen Subdifferentialen oftmals wichtige Rechenregeln, z. B. gelten üblicherweise nur unscharfe (fuzzy) und keine exakten Summenregeln. Zudem sind Subdifferentiale dieser ersten Gruppe stets konvexe Mengen, was unter anderem Ioffe (2017) in Kapitel 4.5 zeigt. Diese meist sehr vorteilhafte Eigenschaft verkehrt sich jedoch oft ins Negative, wenn man mittels Subdifferentialen notwendige Optimalitätsbedingungen herleiten will. Die Subdifferentiale sind dann nämlich in einem gewissen Sinne „sehr groß“ und die aus ihnen abgeleiteten Aussagen verlieren somit massiv an Nutzen, siehe dazu beispielsweise S. 61 ff. in Mordukhovich (2018).

Zu der zweiten, wesentlich größeren Gruppe verallgemeinerter Subdifferentiale gehören zuerst einmal natürlich die von uns verwendeten Subdifferentiale, das approximierende Subdifferential nach Ioffe und das Limiting-Subdifferential nach Kruger und Mordukhovich, aber beispielsweise auch das Bouligand-Subdifferential oder das Rubinov-Subdifferential. Ihre Berechnungen sind hochgradig technisch und komplex und damit zwar für theoretische Betrachtungen bemerkenswert, für praktische Problemstellungen im Allgemeinen aber nicht nutzbar bzw. wurden solche möglichen Anwendungen unseres Wissens nach in der Forschung bisher noch gar nicht thematisiert. Wir erschließen daher mit dieser Arbeit auch einen bisher unerforschten Anwendungsbereich dieser Subdifferentiale.

Für ausführlichere Betrachtungen sowohl zum historischen Hintergrund und zu den Eigenschaften aller hier erwähnten Subdifferentiale als auch zu weiteren Subdifferentialbegriffen verweisen wir an dieser Stelle auf die schier unermessliche Fülle an Werken zu Variationsrechnung und verallgemeinerter Differentiation, von denen insbesondere die Bücher von Rockafellar und Wets (1998), Mordukhovich (2006a,b, 2018) und Ioffe (2017), die Artikel von Borwein und Zhu (1999) und Ioffe (2000, 2012) sowie die Referenzen in diesen empfehlenswert sind.

In der modernen Theorie der Subdifferentiale ist man aufgrund der Vielzahl an Subdifferentialbegriffen größtenteils zum Konzept des abstrakten Subdifferentials übergegangen. Statt einzelner, spezieller Subdifferentiale untersucht man dabei vielmehr Klassen von Funktionalen mit gewissen Eigenschaften, welche man als grundlegend und charakterisierend für ein Subdifferential ansieht, die jedoch auch je nach Autor variieren können. Für diese Klassen von Funktionalen leitet man dann allgemein weitere Resultate her, die für jedes spezielle Subdifferential, welches die Charakteristiken des abstrakten Subdifferentials erfüllt, sofort Gültigkeit besitzen. Einen ausführlicheren Einblick in diese Vorgehensweise liefern beispielsweise Kapitel 5.6 in Göpfert, Riedrich und Tammer (2017) oder Kapitel 4.2 in Ioffe (2017). Dieses Modell hat jedoch schon aus dem naheliegenden Grund, dass ein abstraktes Konzept keine konkreten Berechnungen ermöglicht, für unsere Arbeit keine Relevanz.

Zum Abschluss dieser Sektion wollen wir noch kurz darlegen, weshalb wir in unserer Arbeit genau die Subdifferentiale nach Ioffe sowie Kruger und Mordukhovich verwenden. Zunächst ist für uns die Minimalität der verwendeten Subdifferentiale wichtig und zwar in dem Sinne, dass sie in einer Klasse von mengenwertigen Funktionalen mit gewissen charakteristischen Eigenschaften in allen anderen Funktionalen enthalten sind. Aussagen dieser Art findet man beispielsweise in Theorem 9.7 in Mordukhovich und Shao (1996), als Proposition 2.45 in Mordukhovich (2006a) oder auf S. 537/538 in Ioffe (2000). Diese Minimalitätsforderung macht unter anderem viele Subdifferentiale der ersten Gruppe obsolet, denn diese sind aufgrund ihrer Konvexität oft erheblich umfangreichere Mengen (so ist zum Beispiel das

Clarke-Subdifferential für lokal Lipschitz-stetige Funktionen stets der Abschluss der konvexen Hülle des Limiting-Subdifferentials, vgl. Theorem 3.57 in Mordukhovich (2006a)). Es würden sich zwar in gleicher Art und Weise Optimalitätsbedingungen für die betrachteten Problemstellungen herleiten lassen, diese Bedingungen wären aber geradezu trivial.

Die übrigen Subdifferentiale der ersten Gruppe sind für unsere Arbeit aufgrund der weiteren, oben benannten Nachteile ebenfalls nicht geeignet. Zum einen ist die Existenz einer exakten Summenregel für die Untersuchung der zusammengesetzten Zielfunktionen aus Kapitel 1 unverzichtbar und zum anderen werden wir in den Berechnungen von Kapitel 3 exemplarisch sehen, auf wievielen Punkten des Domains das Dini-Hadamard-Subdifferential nur der leeren Menge entspricht. Für unsere Betrachtungen können also nur Subdifferentiale der zweiten Gruppe in Frage kommen.

Aus dieser zweiten Gruppe sind das approximierende Subdifferential und das Limiting-Subdifferential diejenigen, welche am weitesten verbreitet sind und schon in der Vergangenheit großes Forschungsinteresse erfahren haben. Daher gibt es zu beiden Subdifferentialen sowohl einen breit gefächerten Rechenkalkül als auch eine Vielzahl an Resultaten über ihre Eigenschaften. Nicht zuletzt ist es überhaupt möglich, beide Subdifferentiale für unsere Aufgabenstellungen explizit zu berechnen, dies ist bei Weitem keine Selbstverständlichkeit und hebt sie in besonderem Maße von den übrigen Konstruktionen ab.

In Sektion 2.2.4 werden wir außerdem noch darauf eingehen, welche Zusammenhänge zwischen dem approximierenden Subdifferential und dem Limiting-Subdifferential bestehen, wo ihre Vor- und Nachteile liegen und vor allem, wieso wir in dieser Arbeit überhaupt mit zwei verschiedenen Subdifferentialbegriffen operieren.

2.2.2 Das approximierende Subdifferential nach Ioffe

Wir wollen uns jetzt ausführlich mit den Grundlagen der beiden von uns verwendeten Subdifferentiale beschäftigen und beginnen dabei mit dem approximierenden Subdifferential. Im Jahr 1981 fing Alexander D. Ioffe mit der Einführung einer neuen Konstruktion an, die Ioffe (1981a) selbst damals noch als „approximate M-subdifferential“ bezeichnete. Interessant ist, dass die Benennung „approximierend“ (approximate) dabei höchstwahrscheinlich auf einen Übersetzungsfehler aus dem Russischen und Französischen ins Englische zurückzuführen ist, wie beispielsweise Rockafellar und Wets (1998) auf S. 347 anmerken. Das Subdifferential selbst ist nämlich exakt und in keiner Weise approximierend, der Name sollte ursprünglich vielmehr, wie auch beim Limiting-Subdifferential, das Annähern im Sinne eines Grenzwertes zum Ausdruck bringen. Die Bezeichnung wurde allerdings nie korrigiert und hat sich gleichsam mit dem Subdifferential etabliert.

In späteren Arbeiten desselben Jahres verfeinerte Ioffe (1981b,c,d) die Definition des approximierenden Subdifferentials und arbeitete weitere Ergebnisse aus, aber erst einige Jahre später wurde die finalisierte Theorie von Ioffe (1984, 1986, 1989) in einer dreiteiligen Veröffentlichungsreihe zu „Approximate subdifferentials and applications“ publiziert. Das M-Subdifferential ist in diesen Artikeln durch eine Aufteilung zwischen zwei auf unterschiedliche Weise konstruierten Subdifferentialen ersetzt worden, dem analytischen A-Subdifferential und dem geometrischen G-Subdifferential.

Weitere Resultate wiesen in den 1990er-Jahren dann nach, dass das A-Subdifferential in dieser Form ebenfalls redundant ist bzw. vollständig vom G-Subdifferential abgedeckt wird, so dass heutzutage mit dem Term approximierendes Subdifferential stets das gemeint ist, was Ioffe in den Originalveröffentlichungen als Nukleus des G-Subdifferentials („nuclei of the corresponding G-subdifferentials, G-normal cones etc. or just G-nuclei", Ioffe (1989), S. 3) bezeichnet, vergleiche dazu Borwein und Ioffe (1996) sowie Ioffe (2000).

Auch unsere Arbeit orientiert sich in Bezug auf das approximierende Subdifferential an den modernen Veröffentlichungen von Ioffe (2000, 2012, 2017) und folgt deren dreischrittiger Vorgehensweise zur Definition desselben: Zuerst werden wir das approximierende Subdifferential ausschließlich für lokal Lipschitz-stetige Funktionen über Häufungspunkte des Dini-Hadamard-Subdifferentials definieren. Dieses werden wir anschließend nutzen, um den zugehörigen approximierenden Normalenkegel via des approximierenden Subdifferentials der Abstandsfunktion zu beschreiben. Zuletzt können wir dann das approximierende Subdifferential beliebiger Funktionen (mittels der allgemein gültigen Beziehung zwischen einem Subdifferential und dem mit ihm verbundenen Normalenkegel) über gewisse Elemente des Normalenkegels an ihrem Epigraphen definieren.

Für diese Konstruktionen benötigen wir also zunächst die Definition des Dini-Hadamard-Subdifferentials. Dieses erhält man, wie bereits weiter oben beschrieben, aus der Dini-Richtungsableitung durch Abschwächung der Gleichheit zu einer Ungleichung, vgl. dazu auch S. 123/124 in Borwein und Lewis (2006).

Definition 2.2.2

Es seien $(X, \|\cdot\|_X)$ ein normierter Raum und $f\colon X \to \overline{\mathbb{R}}$ eine Funktion, die in $x \in X$ lokal Lipschitz-stetig ist und für die $|f(x)| < +\infty$ gilt. Dann heißt

$$d^- f(x,h) := \liminf_{t \searrow 0} \frac{f(x+th) - f(x)}{t}$$

die ***(untere) Dini-Richtungsableitung*** von f *im Punkt* x *in Richtung* $h \in X$.

Definition 2.2.3

Es seien $(X, \|\cdot\|_X)$ ein normierter Raum und $f\colon X \to \overline{\mathbb{R}}$ lokal Lipschitz-stetig in $x \in X$. Gilt $|f(x)| < +\infty$, so nennen wir die Menge

$$\partial^- f(x) := \{x^* \in X^* \mid \forall\, h \in X\colon\ \langle x^*, h\rangle \leq d^- f(x,h)\}$$

das ***Dini-Hadamard-Subdifferential*** von f *in* x *und die einzelnen Elemente dieser Menge* ***Dini-Hadamard-Subgradienten*** von f *in* x.

Ist hingegen $|f(x)| = +\infty$, dann setzen wir $\partial^- f(x) := \emptyset$.

Im Falle eines unendlich-dimensionalen Banachraumes $(X, \|\cdot\|_X)$ wird das approximierende Subdifferential als Schnittmenge von Häufungspunkten der Dini-Hadamard-Subdifferentiale auf abgeschlossenen, endlich-dimensionalen Unterräumen definiert. Ist $\mathcal{L}$ ein solcher Unterraum von X, so setzen wir

$$\partial^-_{\mathcal{L}} f(x) := \{x^* \in X^* \mid \forall\, h \in \mathcal{L}\colon\ \langle x^*, h\rangle \leq d^- f(x,h)\}.$$

Damit können wir nun analog zu S. 534 in Ioffe (2000) das approximierende Subdifferential für lokal Lipschitz-stetige Funktionen definieren.

Definition 2.2.4

Es seien $(X, \|\cdot\|_X)$ ein Banachraum, $f\colon X \to \overline{\mathbb{R}}$ eine Funktion, die lokal Lipschitz-stetig in $x_0 \in X$ ist und für welche $|f(x_0)| < +\infty$ gilt, sowie $\mathcal{F}$ die Vereinigung aller endlich-dimensionalen Unterräume von X. Die Menge

$$\partial_a f(x_0) := \bigcap_{\mathcal{L}\in\mathcal{F}} \limsup_{x\to x_0} \partial^-_{\mathcal{L}} f(x) = \bigcap_{\mathcal{L}\in\mathcal{F}} \bigcap_{\varepsilon>0} \bigcup_{\|x-x_0\|_X<\varepsilon} \partial^-_{\mathcal{L}} f(x)$$

bezeichnen wir als das ***approximierende Subdifferential*** *von f im Punkt x_0.*

Für $x_0 \in X$ mit $|f(x_0)| = +\infty$ setzen wir auch hier $\partial_a f(x_0) := \emptyset$.

Ist $(X, \|\cdot\|_X)$ selbst ein endlich-dimensionaler Banachraum, wie es auch für unsere Aufgabenstellungen der Fall ist, so vereinfacht sich die Definition des approximierenden Subdifferentials zu

$$\partial_a f(x_0) := \limsup_{x\to x_0} \partial^- f(x) = \bigcap_{\varepsilon>0} \bigcup_{\|x-x_0\|<\varepsilon} \partial^- f(x).$$

Wir wollen diesen Subdifferentialbegriff im Folgenden auf Funktionen ausdehnen, die nicht lokal Lipschitz-stetig sind. Dazu benötigen wir im zweiten Schritt unserer Konstruktionskette den zugehörigen Normalenkegel. Dies ist der erste Teil von Definition 7.1 in Ioffe (1989), der dort noch als Nukleus des G-Normalenkegels bezeichnete Kegel entspricht dem heutzutage verwendeten approximierenden Normalenkegel.

Definition 2.2.5

Ist $(X, \|\cdot\|_X)$ ein Banachraum, $A \subseteq X$ eine abgeschlossene Menge mit der zugehörigen Abstandsfunktion $d_A(x) := \inf_{y\in A} \|x-y\|$ und $x_0 \in A$, so definieren wir den ***approximierenden Normalenkegel*** *an A im Punkt x_0 durch*

$$N_a(x_0, A) := \bigcup_{\lambda\geq 0} \lambda \partial_a d_A(x_0).$$

Der approximierende Normalenkegel ist wohldefiniert, da Abstandsfunktionen stets Lipschitz-stetig sind (siehe Bemerkung 2.1.2) und die Definition unabhängig von der verwendeten Norm ist, solange äquivalente Normen zugrunde liegen (vgl. Theorem 2.2 sowie S. 27 in Ioffe (1989) oder S. 175 ff., besonders Korollar 4.63, in Ioffe (2017)). Im Gegensatz zu Ioffes ursprünglicher Definition nutzen wir in Definition 2.2.5 die Formulierung $\lambda \geq 0$ statt $\lambda > 0$, um die Kegeleigenschaft hervorzuheben. Korollar 4.64 in Ioffe (2017) bestätigt aber die Gleichheit beider Definitionen.

Jetzt können wir im finalen Schritt das approximierende Subdifferential für (fast) beliebige Funktionen mittels einer allgemeinen Beziehung zwischen Subdifferential und zugehörigem Normalenkegel definieren, dies entspricht nun dem zweiten Teil der Definition 7.1 in Ioffe (1989). Formel (1) auf S. 534 in Ioffe (2000) fasst die beiden letzten Konstruktionsschritte zusammen, ohne den approximierenden Normalenkegel explizit zu benennen.

Definition 2.2.6

Es seien $(X, \|\cdot\|_X)$ ein Banachraum und $f\colon X \to \overline{\mathbb{R}}$ eine unterhalbstetige Funktion. Wir nennen

$$\partial_a f(x_0) := \{x^* \in X^* \mid (x^*, -1) \in N_a((x_0, f(x_0)), \operatorname{epi} f)\}$$

das (verallgemeinerte) ***approximierende Subdifferential*** *von f in $x_0 \in X$.*

Für lokal Lipschitz-stetige Funktionen f stimmt diese Definition mit Definition 2.2.4 überein (siehe dazu Proposition 4.65 in Ioffe (2017)) und bildet daher eine sinnvolle Erweiterung des Begriffes.

Die Bezeichnung als Subdifferential ist auch gerechtfertigt, denn $\partial_a(\cdot)$ besitzt die typischen und wichtigen Eigenschaften, durch welche auch abstrakte Subdifferentiale charakterisiert werden. Ioffe (2000) listet unter anderem die folgenden Eigenschaften in Abschnitt 1.5 von Kapitel 2 auf:

Lemma 2.2.1 (*Eigenschaften des approximierenden Subdifferentials*)

Es seien $(X, \|\cdot\|_X)$ ein Banachraum und $f\colon X \to \overline{\mathbb{R}}$ eine unterhalbstetige Funktion.

(A1) Ist $x \notin \operatorname{dom} f$, so ist $\partial_a f(x) = \emptyset$.

(A2) Für jede konvexe Funktion f stimmt $\partial_a f$ mit dem Subdifferential für konvexe Funktionen ∂f aus Definition 2.2.1 überein.

(A3) Ist f in $x \in X$ lokal Lipschitz-stetig mit Konstante L, dann gilt $\|x^*\|_* \leq L$ für alle $x^* \in \partial_a f(x)$.

(A4) Ist f eigentlich und $x \in X$ ein lokales Minimum von f, dann gilt $0 \in \partial_a f(x)$.

Die letzte Eigenschaft (A4) ist hier von besonderer Bedeutung, denn diese liefert die grundlegende Aussage zur Herleitung aller notwendigen Optimalitätsbedingungen. (A2) stellt außerdem sicher, dass das approximierende Subdifferential eine harmonische Erweiterung der Subdifferentialtheorie auf nichtkonvexe Funktionen bildet. Für uns ist zusätzlich noch die Summenregel des approximierenden Subdifferentials wichtig, diese ist in Theorem 4.69 in Ioffe (2017) enthalten.

Satz 2.2.4 (*Summenregel für approximierende Subdifferentiale*)

Es seien $(X, \|\cdot\|_X)$ ein Banachraum und $f_1, \ldots, f_k$ mit $k \geq 2$ unterhalbstetige Funktionen, welche von X nach $\overline{\mathbb{R}}$ abbilden und von denen zumindest $k-1$ in $x \in \bigcap\limits_{i=1}^{k} \operatorname{dom} f_i$ lokal Lipschitz-stetig sind. Dann gilt

$$\partial_a(f_1 + \ldots + f_k)(x) \subseteq \partial_a f_1(x) + \ldots + \partial_a f_k(x).$$

Das approximierende Subdifferential verfügt auch noch über zahlreiche weitere Eigenschaften und Kalkülregeln, welche den Rahmen dieser Arbeit aber bei Weitem überschreiten würden. Daher sei für eine tiefergehende Beschäftigung mit dem approximierenden Subdifferential auf die in dieser Sektion aufgeführten Quellen verwiesen.

2.2.3 Das Limiting-Subdifferential nach Kruger und Mordukhovich

Das zweite nichtlineare Subdifferential, welches wir in dieser Arbeit intensiv nutzen werden, ist das von Mordukhovich und Kruger erdachte Limiting-Subdifferential (in der Literatur teils auch als „basic subdifferential“ zu finden). Dieses wird ebenso wie das approximierende Subdifferential mittels seines zugehörigen Normalenkegels definiert. Beide Begriffe, die des Limiting-Subdifferentials und des Limiting-Normalenkegels, sind das erste Mal bei Mordukhovich (1976) zu finden und wurden damit nur kurze Zeit nach Clarkes verallgemeinertem Gradienten entwickelt. In diesem Artikel wird der Limiting-Normalenkegel allerdings noch über Grenzwertbetrachtungen von Tangenten definiert und alle der zugrunde liegenden Räume sind endlich-dimensional.

Zum Limiting-Normalenkegel und dessen Anwendungsmöglichkeiten, vor allem in Problemstellungen der optimalen Steuerung, erschienen in den darauf folgenden Jahren die Publikationen von Mordukhovich und Kruger (1976), Mordukhovich (1977) sowie Kruger und Mordukhovich (1978). Die heute übliche Darstellung des Limiting-Normalenkegels über sequentielle Grenzwerte, mit der wir ebenfalls arbeiten werden, sowie die Erweiterung der Theorie auf möglicherweise unendlich-dimensionale Banachräume wurde von Kruger und Mordukhovich (1980a,b) begonnen, durch Kruger (1981a) in dessen Dissertation unter Mordukhovichs Betreuung fortgeführt und schließlich von Kruger (1981b, 1985) in abgeschlossener Form veröffentlicht. Mordukhovich (1988) stellte dann die erste komplette Übersicht über die bis dahin gewonnenen Resultate vor.

Überraschenderweise gab es jedoch zum Limiting-Subdifferential, trotz der heutzutage so intensiv betriebenen Forschung und den weit gefächerten Anwendungsmöglichkeiten, in den folgenden Jahren vorerst kaum weitere Publikationen. Einige zusätzliche, wichtige Eigenschaften entstammen den Artikeln von Mordukhovich (1980, 1984) sowie, unter verallgemeinerteren Bedingungen, denen von Kruger (1981b, 1985), doch erst das Übersichtswerk von Mordukhovich (1988), in welchem auch die gesamte Theorie des Limiting-Subdifferentials dargelegt wurde, erzeugte ein breiteres Forschungsinteresse. Einige der wichtigsten Resultate und die Verallgemeinerung der bisherigen Ergebnisse für Asplund-Räume wurden dabei von Mordukhovich und Shao (1995, 1996) publiziert.

Bei allen Definitionen und Eigenschaften, die wir in dieser Sektion vorstellen, orientieren wir uns an den aktuellen Standardwerken von Mordukhovich (2006a,b, 2018), in welchen unter anderem sämtliche Forschungsergebnisse zu Limiting-Normalenkegel und Limiting-Subdifferential detailliert dargestellt werden. Da wir uns bei der Auswahl der Eigenschaften und Rechenregeln nur auf diejenigen beschränken, welche für uns zwingend notwendig sind, sei daher sowohl für viele weitere Resultate und weiterführende Literatur als auch für einen Überblick über noch offene Problemstellungen auf diese Bücher verwiesen.

Wir werden im Folgenden das Limiting-Subdifferential in seiner allgemeinen Form mithilfe des Limiting-Normalenkegels definieren. Dafür benötigen wir jedoch noch einen speziellen Grenzwertbegriff für mengenwertige Funktionen, den man z. B. als (1.1) auf S. 3 bei Mordukhovich (2006a) findet.

Definition 2.2.7

Es sei $F\colon X \rightrightarrows X^*$ eine mengenwertige Abbildung zwischen einem Banachraum $(X, \|\cdot\|_X)$ und seinem Dualraum. Dann bilden wir den ***sequentiellen oberen Painlevé-Kuratowski-Grenzwert*** *bezüglich der Normtopologie auf X und der schwach*-Topologie auf X^* mittels*

$$\operatorname*{Lim\,sup}_{x \to x_0} F(x) := \Big\{ x^* \in X^* \mid \text{es existieren Folgen } x_k \to x_0 \text{ und } x_k^* \xrightarrow{w^*} x^* \\ \text{mit } x_k^* \in F(x_k) \text{ für alle } k \in \mathbb{N} \Big\}.$$

Nun können wir im Banachraum $(X, \|\cdot\|_X)$ an beliebigen Mengen $\Omega \subseteq X$ sogenannte ε-Normale definieren und erhalten dann durch den sequentiellen Grenzübergang unseren Limiting-Normalenkegel, diese nächsten beiden Definitionen entsprechen Definition 1.1 (i) und (ii) in Mordukhovich (2006a).

Definition 2.2.8

Es sei Ω eine nichtleere Teilmenge des Banachraumes $(X, \|\cdot\|_X)$. Für $x \in \Omega$ und $\varepsilon \geq 0$ bezeichnen wir

$$\widehat{N}_\varepsilon(x, \Omega) := \left\{ x^* \in X^* \mid \limsup_{\substack{u \in \Omega \\ u \to x}} \frac{\langle x^*, u - x \rangle}{\|u - x\|_X} \leq \varepsilon \right\}$$

als die ***Menge aller ε-Normalen*** *an Ω in x.*

Für $\varepsilon = 0$ verzichten wir auf den Index und nennen die Menge $\widehat{N}(x, \Omega)$ den ***Prä-Normalenkegel*** *an Ω in x, seine einzelnen Elemente nennen wir* ***Fréchet-Normale****.*

Ist $x \notin \Omega$, so setzen wir $\widehat{N}_\varepsilon(x, \Omega) := \emptyset$ für alle $\varepsilon \geq 0$.

Definition 2.2.9

Es sei Ω eine nichtleere Teilmenge des Banachraumes $(X, \|\cdot\|_X)$. Für $x_0 \in \Omega$ heißt $x^* \in X^*$ ***Basis-Normal*** *oder* ***Limiting-Normal*** *an Ω in x_0, wenn Folgen $\varepsilon_k \searrow 0$, $x_k \to x_0$ mit $x_k \in \Omega$ und $x_k^* \xrightarrow{w^*} x^*$ existieren, so dass $x_k^* \in \widehat{N}_{\varepsilon_k}(x_k, \Omega)$ für alle $k \in \mathbb{N}$ gilt. Die Vereinigung aller solcher Normalen*

$$N_L(x_0, \Omega) := \operatorname*{Lim\,sup}_{\substack{\varepsilon \searrow 0 \\ x \to x_0}} \widehat{N}_\varepsilon(x, \Omega)$$

nennen wir den ***Basis-*** *oder* ***Limiting-Normalenkegel*** *an Ω in x_0.*

Analog zu Definition 2.2.8 setzen wir auch $N_L(x_0, \Omega) := \emptyset$ für $x_0 \notin \Omega$.

Es lässt sich leicht nachweisen, dass zwar die Mengen der ε-Normalen für $\varepsilon > 0$ von der gewählten Norm auf X abhängen, jedoch sowohl der Prä-Normalenkegel als auch der Normalenkegel invariant unter äquivalenten Normen sind und diese Definitionen damit sinnvoll formuliert sind. Außerdem ist die Menge aller ε-Normalen für jedes $\varepsilon \geq 0$ stets konvex, der Limiting-Normalenkegel selbst ist im Allgemeinen aber schon für sehr einfach strukturierte Mengen Ω eine nichtkonvexe Menge (vgl. S. 5 in Mordukhovich (2006a)).

In endlich-dimensionalen Räumen lassen sich Limiting-Normalenkegel deutlich einfacher berechnen, man kann nämlich entweder auf die ε-Grenzwertbetrachtung verzichten und den sequentiellen Grenzübergang direkt vom Prä-Normalenkegel vollziehen oder alternativ euklidische Projektoren auf die Menge nutzen und von diesen ausgehend den Limiting-Normalenkegel bestimmen. Beide Formeln sind in Theorem 1.6 in Mordukhovich (2006a) zu finden, die Definition des euklidischen Projektors ist diesem Theorem auf derselben Seite vorangestellt.

Definition 2.2.10

Es seien $\Omega \subseteq \mathbb{R}^n$ nichtleer und $x \in \mathbb{R}^n$ beliebig, dann bildet die Menge

$$\Pi(x, \Omega) := \left\{ w \in \Omega \mid \|x - w\|_2 = \inf_{u \in \Omega} \|x - u\|_2 \right\}$$

*den **euklidischen Projektor** von x auf Ω.*

Satz 2.2.5 (*Limiting-Normalenkegel in $\mathbb{R}^n$*)

Es sei $\Omega \subseteq \mathbb{R}^n$ nichtleer und lokal abgeschlossen um $x_0 \in \Omega$, das heißt, es gibt eine Umgebung $U \subseteq \mathbb{R}^n$ von x_0, so dass $U \cap \Omega$ abgeschlossen ist. Dann gelten

$$N_L(x_0, \Omega) = \operatorname*{Lim\,sup}_{x \to x_0} \widehat{N}(x, \Omega) \text{ und}$$

$$\begin{aligned} N_L(x_0, \Omega) &= \operatorname*{Lim\,sup}_{x \to x_0} \left[\operatorname{cone}(x - \Pi(x, \Omega))\right] \\ &= \operatorname*{Lim\,sup}_{x \to x_0} \left\{\alpha w \in \mathbb{R}^n \mid \alpha \geq 0, w \in x - \Pi(x, \Omega)\right\}. \end{aligned}$$

In Asplund-Räumen vereinfacht sich die Bestimmung der Limiting-Normalenkegel ebenso, auch hier kann auf die ε-Grenzwertbetrachtung verzichtet werden. Dieser Zusammenhang ist sogar so stark, dass Asplund-Räume darüber vollständig charakterisiert werden können, siehe dazu auch Theorem 2.35 in Mordukhovich (2006a).

Satz 2.2.6 (*Limiting-Normalenkegel in Asplundräumen*)

Ein Banachraum $(X, \|\cdot\|_X)$ ist genau dann ein Asplund-Raum, wenn für jede abgeschlossene Menge $\Omega \subseteq X$ und jedes $x_0 \in \Omega$ die Beziehung

$$N_L(x_0, \Omega) = \operatorname*{Lim\,sup}_{x \to x_0} \widehat{N}(x, \Omega)$$

gilt.

Jetzt können wir für beliebige Funktionen das Limiting-Subdifferential konstruieren. Analog zur Vorgehensweise beim approximierenden Subdifferential erfolgt diese Konstruktion mittels des zugehörigen Normalenkegels am Epigraphen der Funktion, dies ist Definition 1.77 (i) in Mordukhovich (2006a).

Definition 2.2.11

Es seien $(X, \|\cdot\|_X)$ ein Banachraum, $f\colon X \to \overline{\mathbb{R}}$ und $x \in X$ mit $|f(x)| < +\infty$. Dann nennen wir die Menge

$$\partial_L f(x) := \{x^* \in X^* \mid (x^*, -1) \in N_L((x, f(x)), \operatorname{epi} f)\}$$

das ***Basis-*** *oder* ***Limiting-Subdifferential*** *von f in x und die Elemente dieser Menge* ***Basis-*** *bzw.* ***Limiting-Subgradienten*** *von f in x.*

In gewohnter Weise setzen wir $\partial_L f(x) := \emptyset$ für $|f(x)| = +\infty$ fest.

Natürlich besitzt auch das Limiting-Subdifferential die üblichen und wichtigen Eigenschaften, welche es als Subdifferential auszeichnen. Von diesen wollen wir hier nur die folgenden vier Eigenschaften vorstellen, da sie für unsere Arbeit von großer Bedeutung sind. Alle dieser Aussagen sind bei Mordukhovich (2006a) zu finden, die erste Eigenschaft als Anmerkung auf S. 112, die übrigen (in dieser Reihenfolge) in Theorem 1.93, Korollar 1.81 und Proposition 1.114.

Lemma 2.2.2 (*Eigenschaften des Limiting-Subdifferentials*)

Es seien $(X, \|\cdot\|_X)$ ein Banachraum und $f\colon X \to \overline{\mathbb{R}}$. Dann gelten:

(L1) Für $\lambda \geq 0$ ist $\partial_L(\lambda f)(x) = \lambda \partial_L f(x)$.

(L2) Für jede konvexe Funktion f stimmt $\partial_L f$ mit dem Subdifferential für konvexe Funktionen ∂f aus Definition 2.2.1 überein.

(L3) Ist f in $x \in X$ lokal Lipschitz-stetig mit Konstante L, dann gilt $\|x^*\|_* \leq L$ für alle $x^* \in \partial_L f(x)$.

(L4) Ist f eigentlich und $x \in X$ ein lokales Minimum von f, so gilt $0 \in \partial_L f(x)$.

Eigenschaft (L2) sichert, dass auch das Limiting-Subdifferential eine Erweiterung des konvexen Subdifferentials darstellt, welche sich lückenlos in die klassische Theorie einfügt. Mit (L4) wiederum werden wir später die gewünschten notwendigen Optimalitätsbedingungen herleiten.

Auch zum Limiting-Subdifferential existiert ein umfangreicher Kalkül, welcher die Rechnungen mit diesem Konstrukt vereinfacht. Anders als beim approximierenden Subdifferential gibt es in diesem Kalkül aber nur einige wenige Resultate, welche in beliebigen Banachräumen gelten. Eine wirkliche Vielzahl an Rechenregeln erreicht man erst, wenn der zugrunde liegende Raum ein Asplund-Raum ist. Ähnlich der Charakterisierung in Satz 2.2.6 sind Asplund-Räume auch mit den Limiting-Subdifferentialen so eng verknüpft, dass man sie direkt über eine wesentliche Eigenschaft dieser definieren kann (vgl. Korollar 2.25 in Mordukhovich (2006a)):

Satz 2.2.7

Ein Banachraum $(X, \|\cdot\|_X)$ ist genau dann ein Asplund-Raum, wenn $\partial_L f(x) \neq \emptyset$ für jede in $x \in X$ lokal Lipschitz-stetige Funktion $f\colon X \to \overline{\mathbb{R}}$ ist.

Aus dem Kalkül des Limiting-Subdifferentials in Asplund-Räumen sind für uns die Summenregel und die Kettenregel besonders wichtig. Wir wollen dabei für beide Aussagen Formulierungen nutzen, die lokale Lipschitz-Stetigkeit der zugrunde liegenden Funktionen voraussetzen und sich damit von den üblicherweise verwendeten Resultaten unterscheiden.

Der folgende Satz basiert auf Theorem 3.36 in Mordukhovich (2006a) und den diesem Theorem folgenden Anmerkungen. Außerdem liefern Theorem 1.26 und Korollar 1.81 im selben Buch die Gültigkeit der in Theorem 3.36 geforderten SNEC- und Qualifikationsbedingungen, wenn lokale Lipschitz-Stetigkeit anstelle dieser Bedingungen gegeben ist.

Satz 2.2.8 (*Summenregel für Limiting-Subdifferentiale*)

Es seien $(X, \|\cdot\|_X)$ ein Asplund-Raum und $f_1, \ldots, f_k \colon X \to \overline{\mathbb{R}}$ mit $k \geq 2$ unterhalbstetige Funktionen, von denen mindestens $k-1$ in $x \in \bigcap_{i=1}^{k} \operatorname{dom} f_i$ lokal Lipschitz-stetig sind. Dann gilt

$$\partial_L (f_1 + \ldots + f_k)(x) \subseteq \partial_L f_1(x) + \ldots + \partial_L f_k(x).$$

Der von uns verwendeten Kettenregel liegen in gleicher Weise Theorem 3.41 und Korollar 3.43 in Mordukhovich (2006a) zugrunde, die Forderung lokaler Lipschitz-Stetigkeit zusammen mit Theorem 1.26 und Korollar 1.81 ebenda liefert wiederum die in Korollar 3.43 eigentlich benötigten SNEC- und Qualifikationsbedingungen. Zudem betrachten wir nur Abbildungen auf endlich-dimensionale Räume, da in diesem Fall die von uns vorausgesetzte lokale Lipschitz-Stetigkeit äquivalent zur strikten Lipschitz-Stetigkeit aus Korollar 3.43 ist.

Satz 2.2.9 (*Kettenregel für Limiting-Subdifferentiale*)

Es seien $(X, \|\cdot\|_X)$ ein Asplund-Raum, $(Y, \|\cdot\|_Y)$ ein endlich-dimensionaler Asplund-Raum, $f \colon X \to Y$ lokal Lipschitz-stetig in $x \in X$ und $g \colon Y \to \overline{\mathbb{R}}$ lokal Lipschitz-stetig in $f(x)$. Dann gilt

$$\partial_L (g \circ f)(x) \subseteq \bigcup_{y^* \in \partial_L g(f(x))} \partial_L \langle y^*, f(x) \rangle .$$

Zum Abschluss dieser Sektion sei für viele weitere Eigenschaften und Rechenregeln des Limiting-Subdifferentials, die wir hier nicht alle explizit vorstellen können, nochmals auf Mordukhovich (2006a,b, 2018) verwiesen.

2.2.4 Vergleich von approximierendem und Limiting-Subdifferential

Nachdem wir in Sektion 2.2.1 bereits erörtert haben, wieso sich die soeben vorgestellten Subdifferentiale nach Ioffe und nach Kruger/Mordukhovich für unseren Ansatz am besten eignen, wollen wir unser Kapitel mit einer kurzen Gegenüberstellung dieser Subdifferentiale beschließen. Insbesondere wollen wir dabei auch darauf eingehen, wieso wir in dieser Arbeit zwei verschiedene Subdifferentiale verwenden, statt uns auf eines zu beschränken.

Beide der betrachteten Subdifferentiale verfügen sowohl über verschiedene Vorzüge als auch über gewisse Schwächen. Der größte Vorteil des Limiting-Subdifferentials gegenüber dem approximierenden Subdifferential ist sein deutlich umfangreicherer und besser zu handhabender Kalkül. Allerdings gelten die meisten dieser Resultate nur in Asplund-Räumen, weswegen das approximierende Subdifferential in Banachräumen ohne Asplund-Eigenschaft üblicherweise besser geeignet ist. Zudem ist dessen Berechnung, obwohl ebenfalls komplex und selten explizit möglich, im Allgemeinen immer noch leichter zu handhaben als die des Limiting-Subdifferentials.

Der wichtigste Zusammenhang zwischen beiden Subdifferentialen wird in Punkt (iii) des Theorems 3.59 bei Mordukhovich (2006a) beschrieben, dies ist die Minimalität des Limiting-Subdifferentials gegenüber dem approximierenden Subdifferential.

Satz 2.2.10

Es seien $(X, \|\cdot\|_X)$ ein Asplund-Raum und $f\colon X \to \overline{\mathbb{R}}$ unterhalbstetig in einer Umgebung U von $x \in \operatorname{dom} f$, dann gilt stets

$$\partial_L f(x) \subseteq \partial_a f(x).$$

Ist f sogar lokal Lipschitz-stetig in x und X schwach kompakt erzeugt, dann gilt in der Formel Gleichheit.

Wir nennen einen Banachraum $(X, \|\cdot\|_X)$ dabei schwach kompakt erzeugt, wenn es eine schwach kompakte Menge $K \subseteq X$ mit $X = \operatorname{cl}(\operatorname{span} K)$ gibt. Mordukhovich (2006a) listet auf S. 319/320 verschiedene Klassen schwach kompakt erzeugter Banachräume auf, dazu gehören unter anderem alle reflexiven und alle separablen Banachräume. Beliebige Asplund-Räume sind zwar im Allgemeinen nicht schwach kompakt erzeugt, dafür vererbt sich in schwach kompakt erzeugten Asplund-Räumen diese Eigenschaft auf Unterräume.

In all diesen Räumen stimmen die beiden betrachteten Subdifferentiale für lokal Lipschitz-stetige Funktionen also komplett überein. Es bieten sich damit stets mehrere Möglichkeiten zur Berechnung und Nutzung beider Subdifferentiale, von denen je nach gegebener Struktur des Raumes einige leichter zu realisieren sein können.

Da auch $\mathbb{R}^n$ ein separabler Banachraum ist, folgt direkt aus Satz 2.2.10 das folgende Korollar:

Korollar 2.2.1

Ist $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ lokal Lipschitz-stetig in $x \in \operatorname{dom} f$, so gilt $\partial_L f(x) = \partial_a f(x)$.

Dieses Resultat begründet zusammen mit der Lipschitz-Stetigkeit der von uns verwendeten Abstandsfunktionen (siehe Bemerkung 2.1.2) sofort die Verwendung beider Subdifferentiale. Wir haben hierdurch für unsere Anwendungen im $\mathbb{R}^n$ zwei gut erforschte Subdifferentialbegriffe mit einem jeweils umfangreichen Kalkül zur Hand, zwischen denen wir bei Bedarf einfach wechseln können, ohne dabei irgendwelche Nachteile zu erfahren. Die oben beschriebenen unterschiedlichen Charakteristiken der Subdifferentiale gereichen uns mit dieser Möglichkeit zur parallelen Nutzung beider Subdifferentiale also nur zum Vorteil.

In dieser Arbeit werden wir das Limiting-Subdifferential etwas häufiger verwenden und daher im Folgenden auch nur dieses für Benennungen nutzen, aufgrund der Gleichheit beider Subdifferentiale für unsere Aufgabenstellungen ist dabei aber auch immer das approximierende Subdifferential eingeschlossen.

Wir haben damit nun alle Grundlagen für die weitere Arbeit gelegt und werden diese im folgenden Kapitel 3 nutzen, um (Limiting-)Subdifferentiale für alle benötigten Funktionen explizit zu berechnen.

3 Limiting-Subdifferentiale spezieller Funktionen

In diesem Kapitel werden die Berechnungen der (Limiting-)Subdifferentiale aller Funktionen vorgestellt, die wir in den weiteren Kapiteln für die Herleitung der notwendigen Optimalitätsbedingungen benötigen. Diese Funktionen lassen sich dabei in drei Kategorien einteilen, welchen im Folgenden auch jeweils ein Unterkapitel gewidmet ist: ℓ_p-Normen, mit negativem Vorzeichen behaftete ℓ_p-Normen (welche wir auch häufig einfach als negative ℓ_p-Normen bezeichnen werden) sowie die benötigten Hilfsfunktionen, nämlich die Maximumsfunktion und das nichtlineare Skalarisierungsfunktional aus Sektion 2.1.4. Anzumerken ist hierbei, dass die beiden letztgenannten Unterkapitel 3.2 und 3.3, obgleich für diese Arbeit noch den Grundlagen zugeordnet, selbst schon bemerkenswerte neue Resultate enthalten.

3.1 Limiting-Subdifferentiale der ℓ_p-Normen

Zunächst wollen wir die Limiting-Subdifferentiale der „klassischen" ℓ_p-Normen berechnen, welche wir in Lemma 2.1.1 eingeführt haben. Diese fließen über ihre induzierten Metriken jeweils in den Teil der Zielfunktionen ein, welcher den existierenden anziehenden Standorten zugeordnet werden kann. Da alle ℓ_p-Normen konvexe Funktionen sind, gilt für sie nach Eigenschaft (L2) aus Lemma 2.2.2 die Gleichheit von Limiting-Subdifferential und konvexem Subdifferential, so dass die deutlich einfachere Berechnung des konvexen Subdifferentials jeweils ausreichend ist. Für beliebige Normen gilt dabei das folgende bekannte Resultat, dieses findet man zum Beispiel als Satz 5.15 bei Göpfert, Riedrich und Tammer (2009):

Satz 3.1.1 (***Fenchel-Subdifferential für Normen***)

Ist $(X, \|\cdot\|_X)$ ein reeller Banachraum, so ist die Norm $\|\cdot\|_X\colon X \to \mathbb{R}_+$ subdifferenzierbar und es gilt

$$\partial_L\|\cdot\|_X(x) = \partial\|\cdot\|_X(x) = \{x^* \in X^* \mid x^*(x) = \|x\|_X, \|x^*\|_* = 1\}$$

für jedes $x \in X \setminus \{0\}$ sowie

$$\partial_L\|\cdot\|_X(0) = \partial\|\cdot\|_X(0) = \{x^* \in X^* \mid \|x^*\|_* \leq 1\}.$$

Für den Raum $\mathbb{R}^n$ erhält man aus Satz 3.1.1 wegen der Selbstdualität $(\mathbb{R}^n)^* = \mathbb{R}^n$ sofort das folgende Korollar:

Korollar 3.1.1

Auf $\mathbb{R}^n$ ist jede Norm $\|\cdot\|\colon \mathbb{R}^n \to \mathbb{R}_+$ subdifferenzierbar und es gilt

$$\partial_L\|\cdot\|(x) = \partial\|\cdot\|(x) = \{x^* \in \mathbb{R}^n \mid x^*(x) = \|x\|, \|x^*\|_* = 1\} \text{ für } x \in \mathbb{R}^n \setminus \{0\},$$

$$\partial_L\|\cdot\|(0) = \partial\|\cdot\|(0) = \{x^* \in \mathbb{R}^n \mid \|x^*\|_* \leq 1\}.$$

Allerdings sind die Berechnungen der Subdifferentiale für die ℓ_p-Normen auf $\mathbb{R}^n$ trotzdem umfangreich und in der Literatur unseres Wissens nach kaum aufzufinden, deswegen werden wir alle Resultate dieses Unterkapitels mit vollständiger Beweisführung angeben. Die erste Aussage zum Limiting-Subdifferential der ℓ_1-Norm wurde bereits in der Masterarbeit von Hillmann (2013) genutzt und ist dort Satz 2.10.

Satz 3.1.2 (***Subdifferential der ℓ_1-Norm***)

Die ℓ_1-Norm $\|\cdot\|_1\colon \mathbb{R}^n \to \mathbb{R}$ mit $\|\cdot\|_1(x) = \sum\limits_{i=1}^{n} |x_i|$ besitzt das Limiting-Subdifferential

$$\partial_L\|\cdot\|_1(x) = \left\{ x^* \in \mathbb{R}^n \mid \forall\, i \in \{1,\ldots,n\}\colon x_i^* \in \begin{cases} \{1\} & \text{für } x_i > 0 \\ [-1,1] & \text{für } x_i = 0 \\ \{-1\} & \text{für } x_i < 0 \end{cases} \right\}.$$

Beweis: Für $x = 0$ gilt mit Korollar 3.1.1 und $(\|\cdot\|_1)_* = \|\cdot\|_\infty$

$$\partial_L\|\cdot\|_1(0) = \{x^* \in \mathbb{R}^n \mid \|x^*\|_\infty \leq 1\} = \{x^* \in \mathbb{R}^n \mid |x_i^*| \leq 1, i = 1,\ldots,n\}$$

und daher $\partial_L\|\cdot\|_1(0) = \{x^* \in \mathbb{R}^n \mid x_i^* \in [-1,1] \text{ für } x_i = 0,\, i = 1,\ldots,n\}$.

Sei also $x \in \mathbb{R}^n \setminus \{0\}$, dann liefert wiederum Korollar 3.1.1

$$\begin{aligned} \partial_L\|\cdot\|_1(x) &= \{x^* \in \mathbb{R}^n \mid x^*(x) = \|x\|_1, (\|x^*\|_1)_* = 1\} \\ &= \left\{ x^* \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i^* x_i = \sum_{i=1}^{n} |x_i|, \|x^*\|_\infty = 1 \right\} \\ &= \{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1,\ldots,n\}\colon x_i^* x_i = |x_i|, \|x^*\|_\infty = 1\}. \end{aligned}$$

Die letzte Gleichung erhält man, da für alle $i \in \{1,\ldots,n\}$ wegen $\|x^*\|_\infty = 1$ schon $|x_i^*| \leq 1$ ist und damit stets $x_i^* x_i \leq |x_i|$ gilt.

Für $x_i < 0$ gilt die Gleichheit genau für $x_i^* = -1$, für $x_i > 0$ genau für $x_i^* = 1$. Da mindestens eine Komponente von x von 0 verschieden ist, wird $\|x^*\|_\infty = 1$ dadurch schon erfüllt. Also folgt $x_i^* \in [-1,1]$ für $x_i = 0$ und damit genau die Behauptung. ■

Der nächste Satz ist in dieser Arbeit neu bewiesen, er trifft eine allgemeine Aussage zum Limiting-Subdifferential aller ℓ_p-Normen außerhalb der Spezialfälle $p = 1$ und $p = \infty$.

Satz 3.1.3 (***Subdifferential allgemeiner ℓ_p-Normen***)

Für beliebiges $p \in \mathbb{R}$ mit $p > 1$ ist das Limiting-Subdifferential der ℓ_p-Norm $\|\cdot\|_p\colon \mathbb{R}^n \to \mathbb{R}$ mit $\|\cdot\|_p(x) = \left(\sum\limits_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$ gegeben durch

$$\partial_L\|\cdot\|_p(x) = \left\{ \frac{1}{\|x\|_p^{p-1}} \left(\operatorname{sgn}(x_1)|x_1|^{p-1}, \ldots, \operatorname{sgn}(x_n)|x_n|^{p-1}\right) \right\} \text{ für } x \in \mathbb{R}^n \setminus \{0\},$$

$$\partial_L\|\cdot\|_p(0) = \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p}{p-1}} \leq 1 \right\}.$$

Beweis: Das Limiting-Subdifferential in $x = 0$ erhält man wegen $(\|\cdot\|_p)_* = \|\cdot\|_{\frac{p}{p-1}}$ sofort aus Korollar 3.1.1.

Sei also $x \in \mathbb{R}^n \setminus \{0\}$, dann liefert Korollar 3.1.1

$$\begin{aligned}\partial_L \|\cdot\|_p(x) &= \{x^* \in \mathbb{R}^n \mid x^*(x) = \|x\|_p, (\|x^*\|_p)_* = 1\} \\ &= \left\{x^* \in \mathbb{R}^n \mid \sum_{i=1}^n x_i x_i^* = \|x\|_p, \|x^*\|_{\frac{p}{p-1}} = 1\right\} \\ &= \left\{x^* \in \mathbb{R}^n \mid \sum_{i=1}^n x_i x_i^* = \|x\|_p \|x^*\|_{\frac{p}{p-1}}, \|x^*\|_{\frac{p}{p-1}} = 1\right\}.\end{aligned}$$

Nach der Hölderschen Ungleichung gilt stets

$$\sum_{i=1}^n |x_i x_i^*| \leq \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n |x_i^*|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} = \|x\|_p \|x^*\|_{\frac{p}{p-1}}$$

für alle $x, x^* \in \mathbb{R}^n$. Es ergibt sich die Ungleichungskette

$$\sum_{i=1}^n x_i x_i^* \leq \sum_{i=1}^n |x_i x_i^*| \leq \|x\|_p \|x^*\|_{\frac{p}{p-1}} = \sum_{i=1}^n x_i x_i^*,$$

woraus

$$\sum_{i=1}^n |x_i x_i^*| = \|x\|_p \|x^*\|_{\frac{p}{p-1}} \tag{3.1}$$

und

$$\sum_{i=1}^n x_i x_i^* = \sum_{i=1}^n |x_i x_i^*| \tag{3.2}$$

folgen. Gleichung (3.1) ist gerade die Höldersche Ungleichung mit Gleichheit, diese gilt genau dann, wenn $(|x_1|^p, \ldots, |x_n|^p)$ und $(|x_1^*|^{\frac{p}{p-1}}, \ldots, |x_n^*|^{\frac{p}{p-1}})$ linear abhängig sind. Man erhält also mit einem Faktor $r \in \mathbb{R}$ weiter

$$\begin{aligned}\partial_L \|\cdot\|_p(x) &= \left\{x^* \in \mathbb{R}^n \mid \sum_{i=1}^n x_i x_i^* = \|x\|_p \|x^*\|_{\frac{p}{p-1}}, \|x^*\|_{\frac{p}{p-1}} = 1\right\} \\ &\subseteq \left\{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1, \ldots, n\} \colon |x_i^*|^{\frac{p}{p-1}} = r|x_i|^p, \|x^*\|_{\frac{p}{p-1}} = 1\right\}.\end{aligned}$$

Weil $|x_i|$ sowie $|x_i^*|$ für alle $i \in \{1, \ldots, n\}$ nichtnegativ sind, ist auch r nichtnegativ. Damit liefert Einsetzen der ersten Nebenbedingung in die zweite Nebenbedingung

$$\begin{aligned}1 = \|x^*\|_{\frac{p}{p-1}} &= \left(\sum_{i=1}^n |x_i^*|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} = \left(\sum_{i=1}^n r|x_i|^p\right)^{\frac{p-1}{p}} = r^{\frac{p-1}{p}} \left(\left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}\right)^{p-1} \\ &= r^{\frac{p-1}{p}} \|x\|_p^{p-1}\end{aligned}$$

bzw.

$$r^{\frac{p-1}{p}} = \frac{1}{\|x\|_p^{p-1}}.$$

Mittels weiterer Umformungen folgt nun

$$\begin{aligned}\partial_L \|\cdot\|_p(x) &\subseteq \left\{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1, \ldots, n\} \colon |x_i^*|^{\frac{p}{p-1}} = r|x_i|^p, \|x^*\|_{\frac{p}{p-1}} = 1\right\} \\ &= \left\{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1, \ldots, n\} \colon |x_i^*| = r^{\frac{p-1}{p}} |x_i|^{p-1}, \|x^*\|_{\frac{p}{p-1}} = 1\right\} \\ &\subseteq \left\{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1, \ldots, n\} \colon |x_i^*| = \frac{|x_i|^{p-1}}{\|x\|_p^{p-1}}\right\} \\ &= \left\{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1, \ldots, n\} \colon x_i^* = \pm\frac{|x_i|^{p-1}}{\|x\|_p^{p-1}}\right\}.\end{aligned}$$

Allerdings muss jedes Element des Subdifferentials auch (3.2) erfüllen, diese Gleichung ist (zusammen mit dem Ergebnis des letzten Umformungsschrittes) äquivalent zu $\operatorname{sgn}(x_i^*) = \operatorname{sgn}(x_i)$ für alle $i \in \{1, \dots, n\}$. Damit folgt

$$\partial_L \|\cdot\|_p(x) \subseteq \left\{ x^* \in \mathbb{R}^n \mid \forall\, i \in \{1, \dots, n\} \colon x_i^* = \operatorname{sgn}(x_i) \frac{|x_i|^{p-1}}{\|x\|_p^{p-1}} \right\},$$

Einsetzen in die Ausgangsgleichung verifiziert, dass das einzige Element der rechtsseitigen Menge tatsächlich im Limiting-Subdifferential liegt und somit die Gleichheit gilt. Das war gerade die Behauptung. ■

Die (mit Abstand) wichtigste der ℓ_p-Normen in der Praxis ist die euklidische Norm ℓ_2, so dass wir diese nochmals separat betrachten.

Korollar 3.1.2 (*Subdifferential der ℓ_2-Norm*)

Die ℓ_2-Norm $\|\cdot\|_2 \colon \mathbb{R}^n \to \mathbb{R}$ mit $\|\cdot\|_2(x) = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}$ besitzt das Limiting-Subdifferential

$$\partial_L \|\cdot\|_2(x) = \left\{ \frac{x}{\|x\|_2} \right\} \text{ für } x \in \mathbb{R}^n \setminus \{0\} \text{ und}$$

$$\partial_L \|\cdot\|_2(0) = \{x^* \in \mathbb{R}^n \mid \|x^*\|_2 \leq 1\}.$$

Beweis: Die Aussage folgt durch Einsetzen von $p = 2$ sofort aus Satz 3.1.3. ■

In Satz 3.1.3 werden zwei spezielle Normen ausgenommen, für welche nicht alle Umformungen und Rechenschritte wohldefiniert gewesen wären: Die ℓ_1-Norm bzw. deren Limiting-Subdifferential hatten wir bereits in Satz 3.1.2 untersucht, der folgende Satz beschäftigt sich nun mit dem anderen Spezialfall, dem Limiting-Subdifferential der ℓ_∞-Norm. Ein Ergebnis dazu findet sich auch als Satz 2.12 bei Hillmann (2013), im Sinne einer besseren praktischen Anwendbarkeit in den folgenden Kapiteln wollen wir hier aber eine deutlich überarbeitete Formulierung präsentieren.

Satz 3.1.4 (*Subdifferential der ℓ_∞-Norm*)

Die ℓ_∞-Norm $\|\cdot\|_\infty \colon \mathbb{R}^n \to \mathbb{R}$ mit $\|\cdot\|_\infty(x) = \max_{i \in \{1,\dots,n\}} |x_i|$ hat das Limiting-Subdifferential

$$\begin{aligned}
&\partial_L \|\cdot\|_\infty(x) \\
&= \left\{ x^* \in \mathbb{R}^n \mid \forall\, i \in \{1, \dots, n\} \colon x_i^* x_i - |x_i^*| \max_{i \in \{1,\dots,n\}} |x_i| = 0, \sum_{i=1}^{n} |x_i^*| = 1 \right\} \\
&= \Bigg\{ x^* \in \mathbb{R}^n \mid \forall\, i \in \{1, \dots, n\} \colon x_i^*(|x_i| - \max_{i \in \{1,\dots,n\}} |x_i|) = 0, \\
&\qquad \forall\, x_i^* \neq 0 \colon \operatorname{sgn}(x_i^*) = \operatorname{sgn}(x_i), \sum_{i=1}^{n} |x_i^*| = 1 \Bigg\} \text{ für } x \in \mathbb{R}^n \setminus \{0\},
\end{aligned}$$

$$\partial_L \|\cdot\|_\infty(0) = \{x^* \in \mathbb{R}^n \mid \|x^*\|_1 \leq 1\}.$$

Beweis: Der zweite Teil des Satzes, also das Limiting-Subdifferential im Punkt $x = 0$, folgt mit $(\|\cdot\|_\infty)_* = \|\cdot\|_1$ wieder direkt aus Korollar 3.1.1.

Für $x \in \mathbb{R}^n$ mit $x \neq 0$ liefert Korollar 3.1.1 hingegen

$$\begin{aligned}
\partial_L \|\cdot\|_\infty(x) &= \{x^* \in \mathbb{R}^n \mid x^*(x) = \|x\|_\infty, (\|x^*\|_\infty)_* = 1\} \\
&= \{x^* \in \mathbb{R}^n \mid x^*(x) = \|x\|_\infty, \|x^*\|_1 = 1\} \\
&= \left\{x^* \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^* x_i = \max_{i \in \{1,\dots,n\}} |x_i|, \sum_{i=1}^n |x_i^*| = 1\right\} \\
&= \left\{x^* \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^* x_i = \sum_{i=1}^n (|x_i^*| \max_{i \in \{1,\dots,n\}} |x_i|), \sum_{i=1}^n |x_i^*| = 1\right\}.
\end{aligned}$$

Weil auch $x_i^* x_i \leq |x_i^*| \max_{i \in \{1,\dots,n\}} |x_i|$ für alle $i \in \{1,\dots,n\}$ gilt, folgt daraus direkt $x_i^* x_i = |x_i^*| \max_{i \in \{1,\dots,n\}} |x_i|$ für alle $i \in \{1,\dots,n\}$ und damit bereits die erste der beiden Gleichungen.

Weiterhin gilt für jedes Element $x^* \in \partial_L \|\cdot\|_\infty(x)$ die Ungleichungskette

$$\sum_{i=1}^n x_i^* x_i \leq \sum_{i=1}^n |x_i^*||x_i| \leq \sum_{i=1}^n (|x_i^*| \max_{i \in \{1,\dots,n\}} |x_i|) = \sum_{i=1}^n x_i^* x_i,$$

woraus sich

$$\sum_{i=1}^n |x_i^*||x_i| = \sum_{i=1}^n (|x_i^*| \max_{i \in \{1,\dots,n\}} |x_i|) \tag{3.3}$$

und

$$\sum_{i=1}^n x_i^* x_i = \sum_{i=1}^n |x_i^*||x_i| \tag{3.4}$$

ergeben.

Natürlich gilt $|x_i| \leq \max_{i \in \{1,\dots,n\}} |x_i|$ für alle $i \in \{1,\dots,n\}$, also muss in Gleichung (3.3) sogar komponentenweise Gleichheit gelten und es folgt die äquivalente Bedingung

$$\begin{aligned}
& |x_i^*||x_i| = |x_i^*| \max_{i \in \{1,\dots,n\}} |x_i| \\
\Leftrightarrow \quad & |x_i^*|(|x_i| - \max_{i \in \{1,\dots,n\}} |x_i|) = 0 \\
\Leftrightarrow \quad & x_i^*(|x_i| - \max_{i \in \{1,\dots,n\}} |x_i|) = 0
\end{aligned}$$

für alle $i \in \{1,\dots,n\}$.

Analog ist $x_i^* x_i \leq |x_i^*||x_i|$ für alle $i \in \{1,\dots,n\}$, also gilt auch in Gleichung (3.4) komponentenweise Gleichheit. Das ist aber äquivalent dazu, dass $x_i^* x_i = 0$ oder $\operatorname{sgn}(x_i^*) = \operatorname{sgn}(x_i)$ für alle $i \in \{1,\dots,n\}$ ist, womit auch die zweite Gleichung gilt. ■

3.2 Limiting-Subdifferentiale negativer ℓ_p-Normen

Im folgenden Unterkapitel wollen wir die Limiting-Subdifferentiale derselben Normen wie in Unterkapitel 3.1 berechnen, diesmal allerdings mit einem negativen Vorzeichen behaftet. In unseren Aufgabenstellungen gehen die induzierten Metriken dieser negativen ℓ_p-Normen jeweils in den Teil der Zielfunktionen ein, welcher zu den existierenden abstoßenden Standorten gehört.

Der Unterschied zwischen ℓ_p-Normen und ihren Negativen mag dabei nur äußerst geringfügig erscheinen, er hat aber für unsere Betrachtungen gravierende Auswirkungen, denn durch das negative Vorzeichen verlieren diese Funktionen ihre Konvexität. Damit ist jedoch die Voraussetzung zur Anwendung von Satz 3.1.1 nicht mehr gegeben und wir benötigen einen völlig anderen Beweisweg. Wir werden daher im Folgenden stets das für diese Funktionen etwas leichter zu berechnende approximierende Subdifferential herleiten und regen Gebrauch von Korollar 2.2.1 machen. Da die negativen ℓ_p-Normen immer noch Lipschitz-stetig sind, liefert uns dieses Korollar direkt die Gleichheit von approximierendem und Limiting-Subdifferential.

Zuerst untersuchen wir das Limiting-Subdifferential der negativen ℓ_1-Norm, das folgende Resultat findet sich als Satz 3.2 bereits bei Hillmann (2013). Wir präsentieren hier jedoch einen neuen und komplett elementaren Beweis dieser Aussage, so dass nicht mehr wie zuvor umfangreiche zusätzliche Resultate benutzt werden müssen.

Satz 3.2.1 (*Subdifferential der negativen ℓ_1-Norm*)

Die negative ℓ_1-Norm $-\|\cdot\|_1\colon \mathbb{R}^n \to \mathbb{R}$ mit $-\|\cdot\|_1(x) = -\sum_{i=1}^{n} |x_i|$ besitzt das Limiting-Subdifferential

$$\partial_L(-\|\cdot\|_1)(x) = \left\{ x^* \in \mathbb{R}^n \mid \forall\, i \in \{1,\dots,n\}\colon x_i^* \in \begin{cases} \{-1\} & \text{für } x_i > 0 \\ \{-1,1\} & \text{für } x_i = 0 \\ \{1\} & \text{für } x_i < 0 \end{cases} \right\}.$$

Beweis: Da $-\|\cdot\|_1$ Lipschitz-stetig ist und nur endliche Werte annimmt, berechnen wir das approximierende Subdifferential dieser Funktion gemäß Definition 2.2.4. Diese Berechnung basiert auf dem Dini-Hadamard-Subdifferential aus Definition 2.2.3, welches wir deshalb zuerst herleiten.

Sei also $x \in \mathbb{R}^n$ beliebig. Ist $x_k = 0$ für irgendein $k \in \{1,\dots,n\}$, so betrachten wir den Richtungsvektor $h = (0,\dots,0,1,0,\dots,0) \in \mathbb{R}^n$, welcher als k-ten Eintrag eine Eins und ansonsten nur Nullen enthält. Für jedes Element $x^* \in \partial^-(-\|\cdot\|_1)(x)$ müsste dann einerseits

$$\begin{aligned} x_k^* &= \langle x^*, h\rangle \\ &\le d^-(-\|\cdot\|_1)(x,h) \\ &= \liminf_{t\searrow 0} \frac{-\|x+th\|_1 + \|x\|_1}{t} \\ &= \liminf_{t\searrow 0} \frac{-\sum_{i=1}^{n} |x_i + th_i| + \sum_{i=1}^{n} |x_i|}{t} \\ &= \liminf_{t\searrow 0} \frac{-\sum_{i=1}^{n} |x_i| + |x_k| - |x_k + t| + \sum_{i=1}^{n} |x_i|}{t} \\ &= \liminf_{t\searrow 0} \frac{-|t|}{t} \\ &= -1 \end{aligned}$$

gelten, aber andererseits auch

$$\begin{aligned}
-x_k^* &= \langle x^*, -h\rangle \\
&\le d^-(-\|\cdot\|_1)(x,-h) \\
&= \liminf_{t\searrow 0} \frac{-\|x-th\|_1+\|x\|_1}{t} \\
&= \liminf_{t\searrow 0} \frac{-\sum\limits_{i=1}^n |x_i-th_i| + \sum\limits_{i=1}^n |x_i|}{t} \\
&= \liminf_{t\searrow 0} \frac{-\sum\limits_{i=1}^n |x_i| + |x_k| - |x_k-t| + \sum\limits_{i=1}^n |x_i|}{t} \\
&= \liminf_{t\searrow 0} \frac{-|-t|}{t} \\
&= -1.
\end{aligned}$$

Ein solches x^* kann aber nicht existieren, also gilt $\partial^-(-\|\cdot\|_1)(x) = \emptyset$ für jedes $x \in \mathbb{R}^n$ mit $x_k = 0$ für mindestens ein $k \in \{1,\dots,n\}$.

Sei daher nun $x \in \mathbb{R}^n$ mit $x_i \neq 0$ für alle $i \in \{1,\dots,n\}$. Wählen wir $k \in \{1,\dots,n\}$ beliebig und betrachten den Richtungsvektor $h = (0,\dots,0,x_k,0,\dots,0) \in \mathbb{R}^n$, dann folgen für jedes $x^* \in \partial^-(-\|\cdot\|_1)(x)$ die beiden Ungleichungen

$$\begin{aligned}
\pm x_k^* x_k &= \langle x^*, \pm h\rangle \\
&\le d^-(-\|\cdot\|_1)(x,\pm h) \\
&= \liminf_{t\searrow 0} \frac{-\|x\pm th\|_1+\|x\|_1}{t} \\
&= \liminf_{t\searrow 0} \frac{-\sum\limits_{i=1}^n |x_i\pm th_i| + \sum\limits_{i=1}^n |x_i|}{t} \\
&= \liminf_{t\searrow 0} \frac{-\sum\limits_{i=1}^n |x_i| + |x_k| - |x_k\pm tx_k| + \sum\limits_{i=1}^n |x_i|}{t} \\
&= \liminf_{t\searrow 0} \frac{|x_k| - (1\pm t)|x_k|}{t} \\
&= \liminf_{t\searrow 0} \frac{\mp t|x_k|}{t} \\
&= \mp|x_k|.
\end{aligned}$$

Zusammengefasst liefert das $x_k^* x_k = -|x_k| = -x_k \operatorname{sgn}(x_k)$ und wegen $x_k \neq 0$ ist dies äquivalent zu $x_k^* = -\operatorname{sgn}(x_k)$. Da k beliebig gewählt war, erhalten wir somit $\partial^-(-\|\cdot\|_1)(x) \subseteq \{(-\operatorname{sgn}(x_1),\dots,-\operatorname{sgn}(x_n))\}$. Es bleibt noch zu zeigen, dass das einzige Element der rechtsseitigen Menge tatsächlich ein Dini-Hadamard-Subgradient von $\partial^-(-\|\cdot\|_1)$ in x ist.

Wir nutzen dabei im Folgenden aus, dass für alle $i \in \{1,\dots,n\}$ und jedes $h \in \mathbb{R}^n$ wegen $x_i \neq 0$ stets $\operatorname{sgn}(x_i) = \operatorname{sgn}(x_i + th_i)$ gilt, wenn $t > 0$ nur hinreichend klein genug ist. Insbesondere gilt diese Gleichheit für $h \neq 0$ (der Fall $h = 0$ ist trivial) auf dem Intervall $\left[0, \frac{\min\limits_{i\in\{1,\dots,n\}} |x_i|}{\max\limits_{i\in\{1,\dots,n\}} |h_i|}\right)$ in jeder Komponente und somit auch im Grenzwert.

Sei also $h \in \mathbb{R}^n$ beliebig, dann gilt

$$\begin{aligned}
d^-(-\|\cdot\|_1)(x,h) &= \liminf_{t\searrow 0} \frac{-\|x+th\|_1 + \|x\|_1}{t} \\
&= \liminf_{t\searrow 0} \frac{-\sum\limits_{i=1}^n |x_i + th_i| + \sum\limits_{i=1}^n |x_i|}{t} \\
&= \liminf_{t\searrow 0} \frac{-\sum\limits_{i=1}^n \operatorname{sgn}(x_i+th_i)(x_i+th_i) + \sum\limits_{i=1}^n \operatorname{sgn}(x_i)x_i}{t} \\
&= \liminf_{t\searrow 0} \frac{-\sum\limits_{i=1}^n \operatorname{sgn}(x_i)(x_i+th_i) + \sum\limits_{i=1}^n \operatorname{sgn}(x_i)x_i}{t} \\
&= \liminf_{t\searrow 0} \frac{-\sum\limits_{i=1}^n \operatorname{sgn}(x_i)x_i - t\sum\limits_{i=1}^n \operatorname{sgn}(x_i)h_i + \sum\limits_{i=1}^n \operatorname{sgn}(x_i)x_i}{t} \\
&= \liminf_{t\searrow 0} -\sum_{i=1}^n \operatorname{sgn}(x_i)h_i \\
&= -\sum_{i=1}^n \operatorname{sgn}(x_i)h_i \\
&= \langle(-\operatorname{sgn}(x_1),\ldots,-\operatorname{sgn}(x_n)),h\rangle
\end{aligned}$$

und damit $(-\operatorname{sgn}(x_1),\ldots,-\operatorname{sgn}(x_n)) \in \partial^-(-\|\cdot\|_1)(x)$. Das liefert nun die Gleichheit $\partial^-(-\|\cdot\|_1)(x) = \{(-\operatorname{sgn}(x_1),\ldots,-\operatorname{sgn}(x_n))\}$.

Jetzt können wir das approximierende Subdifferential von $-\|\cdot\|_1$ als Menge der oberen Häufungspunkte des Dini-Hadamard-Subdifferentials bestimmen und dabei im endlich-dimensionalen Raum $\mathbb{R}^n$ die vereinfachte Form der Berechnung nutzen, für beliebiges $x \in \mathbb{R}^n$ gilt also

$$\partial_a(-\|\cdot\|_1)(x) = \limsup_{y\to x} \partial^-(-\|\cdot\|_1)(y) = \bigcap_{\varepsilon>0} \bigcup_{\|y-x\|_1<\varepsilon} \partial^-(-\|\cdot\|_1)(y).$$

Jeder approximierende Subgradient von $-\|\cdot\|_1$ in x ist insbesondere auch ein Dini-Hadamard-Subgradient von $-\|\cdot\|_1$ (in einem möglicherweise anderen Punkt $y \in \mathbb{R}^n$) und kann daher komponentenweise ebenfalls nur die Werte 1 oder -1 annehmen. Für jedes $i \in \{1,\ldots,n\}$ mit $x_i \neq 0$ gilt stets $\operatorname{sgn}(y_i) = \operatorname{sgn}(x_i)$ für alle $y \in \mathbb{R}^n$ mit $\|y-x\|_1 < |x_i|$. Jeder Dini-Hadamard-Subgradient, der für ein Element aus dieser offenen Kugel existiert, kann also als i-te Komponente sogar nur den Wert $-\operatorname{sgn}(x_i)$ haben. Zusammen liefert dies bereits

$$\partial_a(-\|\cdot\|_1)(x) \subseteq \left\{ x^* \in \mathbb{R}^n \mid \forall\, i \in \{1,\ldots,n\}\colon x_i^* \in \begin{cases} \{-\operatorname{sgn}(x_i)\} & \text{für } x_i \neq 0 \\ \{-1,1\} & \text{für } x_i = 0 \end{cases} \right\}.$$

Wir betrachten nun zu beliebigem, aber festem $\varepsilon > 0$ die Menge

$$Y_\varepsilon(x) := \left\{ y \in \mathbb{R}^n \mid \forall\, i \in \{1,\ldots,n\}\colon y_i \in \begin{cases} \{x_i\} & \text{für } x_i \neq 0 \\ \{-\frac{\varepsilon}{2n}, \frac{\varepsilon}{2n}\} & \text{für } x_i = 0 \end{cases} \right\}.$$

Für jedes $y \in Y_\varepsilon(x)$ gilt $\|y-x\|_1 < \varepsilon$ und jedes $y \in Y_\varepsilon(x)$ ist in jeder Komponente von Null verschieden. Damit existiert aber ein Dini-Hadamard-Subgradient von

$-\|\cdot\|_1$ zu jedem $y \in Y_\varepsilon(x)$ und da dieser jeweils sogar unabhängig von ε ist, folgt

$$\begin{aligned}\partial_a(-\|\cdot\|_1)(x) &= \bigcap_{\varepsilon>0} \bigcup_{\|y-x\|_1<\varepsilon} \partial^-(-\|\cdot\|_1)(y) \\ &\supseteq \bigcap_{\varepsilon>0} \bigcup_{y\in Y_\varepsilon(x)} \partial^-(-\|\cdot\|_1)(y) \\ &= \left\{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1,\ldots,n\} \colon x_i^* \in \begin{cases}\{-\operatorname{sgn}(x_i)\} & \text{für } x_i \neq 0 \\ \{-1,1\} & \text{für } x_i = 0\end{cases}\right\}.\end{aligned}$$

Mit Korollar 2.2.1 erhält man nun sofort die Aussage des Satzes. ■

Genau wie im vorhergehenden Unterkapitel 3.1 präsentieren wir im nächsten Satz das Limiting-Subdifferential der negativen ℓ_p-Normen für den Fall $1 < p < +\infty$, dabei handelt es sich wieder um ein komplett neues Resultat dieser Arbeit.

Satz 3.2.2 (*Subdifferential allgemeiner negativer ℓ_p-Normen*)

Für beliebiges $p \in \mathbb{R}$ mit $p > 1$ hat die negative ℓ_p-Norm $-\|\cdot\|_p \colon \mathbb{R}^n \to \mathbb{R}$ mit $(-\|\cdot\|_p)(x) = -\left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$ das Limiting-Subdifferential

$$\partial_L(-\|\cdot\|_p)(x) = \left\{-\frac{1}{\|x\|_p^{p-1}} \left(\operatorname{sgn}(x_1)|x_1|^{p-1}, \ldots, \operatorname{sgn}(x_n)|x_n|^{p-1}\right)\right\}$$

für $x \in \mathbb{R}^n \setminus \{0\}$ bzw.

$$\partial_L(-\|\cdot\|_p)(0) = \left\{x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p}{p-1}} = 1\right\}.$$

Beweis: Für jedes $p \in \mathbb{R}$ mit $p > 1$ ist $-\|\cdot\|_p$ Lipschitz-stetig und nimmt nur endliche Werte an, wir können also wieder das Dini-Hadamard-Subdifferential nach Definition 2.2.3 und das approximierende Subdifferential nach Definition 2.2.4 berechnen.

In $x = 0$ müsste jedes Element $x^* \in \partial^-(-\|\cdot\|_p)(0)$ des Dini-Hadamard-Subdifferentials für beliebiges $h \in \mathbb{R}^n$ sowohl

$$\begin{aligned}\langle x^*, h\rangle &\leq d^-(-\|\cdot\|_p)(x,h) \\ &= \liminf_{t\searrow 0} \frac{-\|x+th\|_p + \|x\|_p}{t} \\ &= \liminf_{t\searrow 0} \frac{-|t|\|h\|_p}{t} \\ &= -\|h\|_p\end{aligned}$$

als auch

$$\begin{aligned}-\langle x^*, h\rangle = \langle x^*, -h\rangle &\leq d^-(-\|\cdot\|_p)(x,-h) \\ &= \liminf_{t\searrow 0} \frac{-\|x-th\|_p + \|x\|_p}{t} \\ &= \liminf_{t\searrow 0} \frac{-|-t|\|h\|_p}{t} \\ &= -\|h\|_p\end{aligned}$$

erfüllen. Wegen $\|h\|_p > 0$ für alle $h \neq 0$ ist dies aber für kein $x^* \in \mathbb{R}^n$ möglich, daher folgt $\partial^-(-\|\cdot\|_p)(0) = \emptyset$.

Im Folgenden nutzen wir vielfach die Bernoullische Ungleichung $(1+x)^r \geq 1 + rx$ für $x \in \mathbb{R}$ mit $x > -1$ und $r \in \mathbb{R}$ mit $r \geq 1$ sowie die (Taylor-)Reihenentwicklung $(1+x)^r = \sum\limits_{j=0}^{\infty} \binom{r}{j} x^j = 1 + rx + g(x)$ mit $g(x) \in \mathcal{O}(x^2)$ für $x \in \mathbb{R}$ mit $|x| < 1$ und $r \in \mathbb{R}$.

Sei also $x \in \mathbb{R}^n \setminus \{0\}$ und $k \in \{1, \dots, n\}$ beliebig. Ist $x_k \neq 0$, so betrachten wir den Richtungsvektor $h = (0, \dots, 0, x_k, 0, \dots, 0) \in \mathbb{R}^n$. Für jedes $x^* \in \partial^-(-\| \cdot \|_p)(x)$ muss dann

$$\begin{aligned}
x_k^* x_k &= \langle x^*, h \rangle \\
&\leq d^-(-\| \cdot \|_p)(x, h) \\
&= \liminf_{t \searrow 0} \frac{-\|x + th\|_p + \|x\|_p}{t} \\
&= \liminf_{t \searrow 0} \frac{-\left(\sum\limits_{i=1}^{n} |x_i + th_i|^p \right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&= \liminf_{t \searrow 0} \frac{-\left(\sum\limits_{i=1}^{n} |x_i|^p - |x_k|^p + |x_k + tx_k|^p \right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&= \liminf_{t \searrow 0} \frac{-\left(\|x\|_p^p - |x_k|^p + (1+t)^p |x_k|^p \right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&\leq \liminf_{t \searrow 0} \frac{-\left(\|x\|_p^p - |x_k|^p + (1+pt) |x_k|^p \right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&= \liminf_{t \searrow 0} \frac{-\left(\|x\|_p^p + pt |x_k|^p \right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&= \liminf_{t \searrow 0} \frac{-\left(\|x\|_p^p \left(1 + \frac{pt|x_k|^p}{\|x\|_p^p} \right) \right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&= \liminf_{t \searrow 0} \frac{-\|x\|_p \left(1 + \frac{pt|x_k|^p}{\|x\|_p^p} \right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&= \liminf_{t \searrow 0} \frac{-\|x\|_p \left(1 + \frac{1}{p} \frac{pt|x_k|^p}{\|x\|_p^p} + g\left(\frac{pt|x_k|^p}{\|x\|_p^p} \right) \right) + \|x\|_p}{t} \\
&= \liminf_{t \searrow 0} \frac{-\|x\|_p \left(t \frac{|x_k|^p}{\|x\|_p^p} + g\left(t \frac{p|x_k|^p}{\|x\|_p^p} \right) \right)}{t} \\
&= -\|x\|_p \frac{|x_k|^p}{\|x\|_p^p} - \|x\|_p \liminf_{t \searrow 0} \frac{g\left(t \frac{p|x_k|^p}{\|x\|_p^p} \right)}{t} \\
&= -\frac{|x_k|^p}{\|x\|_p^{p-1}}
\end{aligned}$$

gelten. Völlig analog folgt mittels Betrachtung des Richtungsvektors $-h \in \mathbb{R}^n$, dass jeder Dini-Hadamard-Subgradient $x^* \in \partial^-(-\| \cdot \|_p)(x)$ auch

$$-x_k^* x_k \leq \frac{|x_k|^p}{\|x\|_p^{p-1}}$$

erfüllen muss. Insgesamt muss also

$$-x_k^* x_k = \frac{|x_k|^p}{\|x\|_p^{p-1}} \qquad \text{bzw.} \qquad x_k^* = -\frac{\operatorname{sgn}(x_k)|x_k|^{p-1}}{\|x\|_p^{p-1}}$$

für jedes $x^* \in \partial^-(-\|\cdot\|_p)(x)$ und jedes $k \in \{1,\dots,n\}$ mit $x_k \neq 0$ gelten.

Ist hingegen $x_k = 0$ für $k \in \{1,\dots,n\}$, dann liefern die obigen Betrachtungen nur die triviale Ungleichung $0 \leq 0$. Stattdessen untersuchen wir in diesen Fällen den Richtungsvektor $h = (0,\dots,0,1,0,\dots,0) \in \mathbb{R}^n$, welcher genau an der k-ten Position eine Eins und ansonsten nur Nullen enthält. Für jedes $x^* \in \partial^-(-\|\cdot\|_p)(x)$ gilt dann in ähnlicher Weise wie zuvor die Ungleichung

$$\begin{aligned}
x_k^* &= \langle x^*, h\rangle \\
&\leq d^-(-\|\cdot\|_p)(x,h) \\
&= \liminf_{t\searrow 0} \frac{-\|x+th\|_p + \|x\|_p}{t} \\
&= \liminf_{t\searrow 0} \frac{-\left(\sum\limits_{i=1}^n |x_i + th_i|^p\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&= \liminf_{t\searrow 0} \frac{-\left(\sum\limits_{i=1}^n |x_i|^p - |x_k|^p + |x_k + t|^p\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&= \liminf_{t\searrow 0} \frac{-\left(\|x\|_p^p + t^p\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&= \liminf_{t\searrow 0} \frac{-\|x\|_p \left(1 + \frac{t^p}{\|x\|_p^p}\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
&= \liminf_{t\searrow 0} \frac{-\|x\|_p \left(1 + \frac{1}{p}\frac{t^p}{\|x\|_p^p} + g\left(\frac{t^p}{\|x\|_p^p}\right)\right) + \|x\|_p}{t} \\
&= \liminf_{t\searrow 0} \frac{-\|x\|_p \left(t^p \frac{1}{p\|x\|_p^p} + g\left(t^p \frac{1}{\|x\|_p^p}\right)\right)}{t} \\
&= 0,
\end{aligned}$$

wobei sich der Grenzwert wegen $p > 1$ aus $\liminf\limits_{t\searrow 0} \frac{t^p}{t} = 0$ ergibt.

Weiterhin erhält man auch hier für jedes $x^* \in \partial^-(-\|\cdot\|_p)(x)$ durch Betrachtung des Richtungsvektors $-h \in \mathbb{R}^n$ die umgekehrte Implikation

$$-x_k^* \leq 0,$$

zusammengefasst liefert das

$$x_k^* = 0 = -\frac{\operatorname{sgn}(x_k)|x_k|^{p-1}}{\|x\|_p^{p-1}}$$

für jedes $x^* \in \partial^-(-\|\cdot\|_p)(x)$ und jedes $k \in \{1,\dots,n\}$ mit $x_k = 0$.

Ist für beliebiges $x \in \mathbb{R}^n \setminus \{0\}$ also $x^* \in \partial^-(-\|\cdot\|_p)(x)$, so gilt stets

$$x^* = -\frac{1}{\|x\|_p^{p-1}} \left(\operatorname{sgn}(x_1)|x_1|^{p-1}, \dots, \operatorname{sgn}(x_n)|x_n|^{p-1}\right).$$

Dieses Element ist auch wirklich ein Dini-Hadamard-Subgradient von $-\|\cdot\|_p$ in x und das Dini-Hadamard-Subdifferential damit in jedem von Null verschiedenen Punkt einelementig, denn für beliebiges $h \in \mathbb{R}^n$ gilt immer

$$
\begin{aligned}
& d^-(-\|\cdot\|_p)(x,h) \\
= & \liminf_{t\searrow 0} \frac{-\|x+th\|_p + \|x\|_p}{t} \\
= & \liminf_{t\searrow 0} \frac{-\left(\sum\limits_{i=1}^n |x_i+th_i|^p\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
= & \liminf_{t\searrow 0} \frac{-\left(\sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i+th_i|^p + \sum\limits_{\substack{i=1\\x_i=0}}^n |x_i+th_i|^p\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
= & \liminf_{t\searrow 0} \frac{-\left(\sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p \left(1+t\frac{h_i}{x_i}\right)^p + \sum\limits_{\substack{i=1\\x_i=0}}^n |th_i|^p\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
= & \liminf_{t\searrow 0} \frac{-\left(\sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p \left(1+pt\frac{h_i}{x_i} + g\left(t\frac{h_i}{x_i}\right)\right) + t^p \sum\limits_{\substack{i=1\\x_i=0}}^n |h_i|^p\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
= & \liminf_{t\searrow 0} \frac{-\left(\sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p + pt\sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p\frac{h_i}{x_i} + \sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p g\left(t\frac{h_i}{x_i}\right) + t^p \sum\limits_{\substack{i=1\\x_i=0}}^n |h_i|^p\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
= & \liminf_{t\searrow 0} \frac{-\left(\|x\|_p^p\left(1+\frac{pt}{\|x\|_p^p}\sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p\frac{h_i}{x_i} + r\right)\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
= & \liminf_{t\searrow 0} \frac{-\|x\|_p\left(1+\frac{pt}{\|x\|_p^p}\sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p\frac{h_i}{x_i} + r\right)^{\frac{1}{p}} + \|x\|_p}{t} \\
= & \liminf_{t\searrow 0} \frac{-\|x\|_p\left(1+\frac{1}{p}\left(\frac{pt}{\|x\|_p^p}\sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p\frac{h_i}{x_i} + r\right) + k\left(\frac{pt}{\|x\|_p^p}\sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p\frac{h_i}{x_i} + r\right)\right) + \|x\|_p}{t},
\end{aligned}
$$

wobei

$$
r = \frac{1}{\|x\|_p^p}\sum_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p g\left(t\frac{h_i}{x_i}\right) + \frac{t^p}{\|x\|_p^p}\sum_{\substack{i=1\\x_i=0}}^n |h_i|^p
$$

und $k(x) \in \mathcal{O}(x^2)$ sein sollen.

Wegen

$$\liminf_{t\searrow 0} \frac{r}{t} = 0 \text{ und } \liminf_{t\searrow 0} \frac{k(t)}{t} = 0$$

gilt weiterhin

$$\begin{aligned}
&d^-(-\|\cdot\|_p)(x,h) \\
&= \liminf_{t\searrow 0} \frac{-\|x\|_p \left(1 + \frac{1}{p}\left(\frac{pt}{\|x\|_p^p} \sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p \frac{h_i}{x_i} + r\right) + k\left(\frac{pt}{\|x\|_p^p} \sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p \frac{h_i}{x_i} + r\right)\right) + \|x\|_p}{t} \\
&= \liminf_{t\searrow 0} \frac{-\|x\|_p \left(\frac{t}{\|x\|_p^p} \sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p \frac{h_i}{x_i} + \frac{1}{p} r + k\left(\frac{pt}{\|x\|_p^p} \sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p \frac{h_i}{x_i} + r\right)\right)}{t} \\
&= \liminf_{t\searrow 0} \left(-\frac{1}{\|x\|_p^{p-1}} \sum_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p \frac{h_i}{x_i} + \frac{-\|x\|_p \left(\frac{1}{p} r + k\left(\frac{pt}{\|x\|_p^p} \sum\limits_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p \frac{h_i}{x_i} + r\right)\right)}{t} \right) \\
&= -\frac{1}{\|x\|_p^{p-1}} \sum_{\substack{i=1\\x_i\neq 0}}^n |x_i|^p \frac{h_i}{x_i} \\
&= -\frac{1}{\|x\|_p^{p-1}} \sum_{\substack{i=1\\x_i\neq 0}}^n \operatorname{sgn}(x_i)|x_i|^{p-1} h_i \\
&= -\frac{1}{\|x\|_p^{p-1}} \sum_{i=1}^n \operatorname{sgn}(x_i)|x_i|^{p-1} h_i \\
&= \left\langle -\frac{1}{\|x\|_p^{p-1}} \left(\operatorname{sgn}(x_1)|x_1|^{p-1}, \dots, \operatorname{sgn}(x_n)|x_n|^{p-1}\right), h \right\rangle.
\end{aligned}$$

Daraus folgt nun mit der Definition des Dini-Hadamard-Subdifferentials und weil $h \in \mathbb{R}^n$ beliebig gewählt war

$$-\frac{1}{\|x\|_p^{p-1}} \left(\operatorname{sgn}(x_1)|x_1|^{p-1}, \dots, \operatorname{sgn}(x_n)|x_n|^{p-1}\right) \in \partial^-(-\|\cdot\|_p)(x),$$

somit gilt

$$\partial^-(-\|\cdot\|_p)(x) = \left\{ -\frac{1}{\|x\|_p^{p-1}} \left(\operatorname{sgn}(x_1)|x_1|^{p-1}, \dots, \operatorname{sgn}(x_n)|x_n|^{p-1}\right) \right\}$$

für jedes $x \in \mathbb{R}^n \setminus \{0\}$.

Das approximierende Subdifferential von $-\|\cdot\|_p$ ergibt sich für $x \in \mathbb{R}^n$ im endlichdimensionalen Raum $\mathbb{R}^n$ nun definitionsgemäß als

$$\partial_a(-\|\cdot\|_p)(x) = \limsup_{y\to x} \partial^-(-\|\cdot\|_p)(y) = \bigcap_{\varepsilon>0} \bigcup_{\|y-x\|_p<\varepsilon} \partial^-(-\|\cdot\|_p)(y).$$

Wir unterscheiden im Folgenden wieder zwei Fälle. Sei zuerst $x \in \mathbb{R}^n \setminus \{0\}$, dann folgt sofort $\partial^-(-\|\cdot\|_p)(x) \subseteq \partial_a(-\|\cdot\|_p)(x)$. Sei weiter $\varepsilon > 0$ beliebig und $y \in \mathbb{R}^n \setminus \{0\}$ mit $\|x - y\|_p < \varepsilon$. Dann ist $\left(-\frac{1}{\|y\|_p^{p-1}}\left(\operatorname{sgn}(y_1)|y_1|^{p-1}, \ldots, \operatorname{sgn}(y_n)|y_n|^{p-1}\right)\right)$ das einzige Element in $\partial^-(-\|\cdot\|_p)(y)$ und es gilt

$$\begin{aligned}
&\left\| -\frac{1}{\|x\|_p^{p-1}}\left(\operatorname{sgn}(x_1)|x_1|^{p-1}, \ldots, \operatorname{sgn}(x_n)|x_n|^{p-1}\right)\right. \\
&\quad \left. - \left(-\frac{1}{\|y\|_p^{p-1}}\left(\operatorname{sgn}(y_1)|y_1|^{p-1}, \ldots, \operatorname{sgn}(y_n)|y_n|^{p-1}\right)\right)\right\|_{\frac{p}{p-1}} \\
&= \left(\sum_{i=1}^n \left|\frac{\operatorname{sgn}(y_i)|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{\operatorname{sgn}(x_i)|x_i|^{p-1}}{\|x\|_p^{p-1}}\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}.
\end{aligned} \tag{3.5}$$

Ist $x_i = 0$ für $i \in \{1, \ldots, n\}$, so gilt im Falle $y_i = 0$

$$\left|\frac{\operatorname{sgn}(y_i)|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{\operatorname{sgn}(x_i)|x_i|^{p-1}}{\|x\|_p^{p-1}}\right| = \left|\frac{|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{|x_i|^{p-1}}{\|x\|_p^{p-1}}\right|$$

und anderenfalls

$$\left|\frac{\operatorname{sgn}(y_i)|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{\operatorname{sgn}(x_i)|x_i|^{p-1}}{\|x\|_p^{p-1}}\right| = |\operatorname{sgn}(y_i)| \left|\frac{|y_i|^{p-1}}{\|y\|_p^{p-1}}\right| = \left|\frac{|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{|x_i|^{p-1}}{\|x\|_p^{p-1}}\right|.$$

Ist hingegen $x_i \neq 0$ für $i \in \{1, \ldots, n\}$, dann gilt $\operatorname{sgn}(x_i) = \operatorname{sgn}(y_i)$ für hinreichend kleines $\varepsilon > 0$. Damit folgt auch in diesem Fall

$$\left|\frac{\operatorname{sgn}(y_i)|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{\operatorname{sgn}(x_i)|x_i|^{p-1}}{\|x\|_p^{p-1}}\right| = |\operatorname{sgn}(x_i)| \left|\frac{|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{|x_i|^{p-1}}{\|x\|_p^{p-1}}\right| = \left|\frac{|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{|x_i|^{p-1}}{\|x\|_p^{p-1}}\right|,$$

also gilt stets

$$\left|\frac{\operatorname{sgn}(y_i)|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{\operatorname{sgn}(x_i)|x_i|^{p-1}}{\|x\|_p^{p-1}}\right| = \left|\frac{|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{|x_i|^{p-1}}{\|x\|_p^{p-1}}\right|.$$

Einsetzen in Formel (3.5) liefert daher

$$\begin{aligned}
&\left(\sum_{i=1}^n \left|\frac{\operatorname{sgn}(y_i)|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{\operatorname{sgn}(x_i)|x_i|^{p-1}}{\|x\|_p^{p-1}}\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \\
&= \left(\sum_{i=1}^n \left|\frac{|y_i|^{p-1}}{\|y\|_p^{p-1}} - \frac{|x_i|^{p-1}}{\|x\|_p^{p-1}}\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \\
&= \left(\sum_{i=1}^n \left|\frac{|y_i|^{p-1}\|x\|_p^{p-1} - |x_i|^{p-1}\|y\|_p^{p-1}}{\|y\|_p^{p-1}\|x\|_p^{p-1}}\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}
\end{aligned}$$

und es gilt mit $|\|x\| - \|y\|| \leq \|x - y\|_p < \varepsilon$ offenbar

$$\lim_{\varepsilon \searrow 0} \left(\sum_{i=1}^n \left|\frac{|y_i|^{p-1}\|x\|_p^{p-1} - |x_i|^{p-1}\|y\|_p^{p-1}}{\|y\|_p^{p-1}\|x\|_p^{p-1}}\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} = 0.$$

Damit folgt für $x \in \mathbb{R}^n \setminus \{0\}$ aber direkt

$$\partial^-(-\|\cdot\|_p)(x) = \limsup_{y \to x} \partial^-(-\|\cdot\|_p)(y) = \partial_a(-\|\cdot\|_p)(x).$$

Als zweiten Fall betrachten wir nun $x = 0$. Zu beliebigem $y \in \mathbb{R}^n \setminus \{0\}$ ist

$$\left(-\frac{1}{\|y\|_p^{p-1}}\left(\operatorname{sgn}(y_1)|y_1|^{p-1}, \ldots, \operatorname{sgn}(y_n)|y_n|^{p-1}\right)\right)$$

der einzige Dini-Hadamard-Subgradient von $-\|\cdot\|_p$ in y und es gilt

$$\begin{aligned}
&\left\|\left(-\frac{1}{\|y\|_p^{p-1}}\left(\operatorname{sgn}(y_1)|y_1|^{p-1}, \ldots, \operatorname{sgn}(y_n)|y_n|^{p-1}\right)\right)\right\|_{\frac{p}{p-1}} \\
&= \left(\sum_{i=1}^n \left|-\frac{1}{\|y\|_p^{p-1}} \operatorname{sgn}(y_i)|y_i|^{p-1}\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \\
&= \left(\sum_{i=1}^n \left|\frac{1}{\|y\|_p^{p-1}}\right|^{\frac{p}{p-1}} \cdot |-\operatorname{sgn}(y_i)|^{\frac{p}{p-1}} \cdot \left||y_i|^{p-1}\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \\
&= \left(\frac{1}{\|y\|_p^p} \sum_{i=1}^n |y_i|^p\right)^{\frac{p-1}{p}} \\
&= \left(\frac{1}{\|y\|_p^p}\|y\|_p^p\right)^{\frac{p-1}{p}} \\
&= 1.
\end{aligned}$$

Daraus folgt bereits

$$\partial_a(-\|\cdot\|_p)(0) = \limsup_{y \to 0} \partial^-(-\|\cdot\|_p)(y) \subseteq \left\{x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p}{p-1}} = 1\right\}.$$

Sei nun andererseits $x^* \in \mathbb{R}^n$ beliebig mit $\|x^*\|_{\frac{p}{p-1}} = 1$ und $\varepsilon > 0$ beliebig, aber fest, dann betrachten wir $y \in \mathbb{R}^n$ mit

$$y_i = -\frac{\varepsilon}{2} \operatorname{sgn}(x_i^*)|x_i^*|^{\frac{1}{p-1}}$$

für alle $i \in \{1, \ldots, n\}$. Aufgrund von $\|x^*\|_{\frac{p}{p-1}} = 1$ gilt

$$\begin{aligned}
\|y\|_p &= \left(\sum_{i=1}^n \left|-\frac{\varepsilon}{2} \operatorname{sgn}(x_i^*)|x_i^*|^{\frac{1}{p-1}}\right|^p\right)^{\frac{1}{p}} \\
&= \left(\sum_{i=1}^n \left|-\frac{\varepsilon}{2}\right|^p |\operatorname{sgn}(x_i^*)|^p \left||x_i^*|^{\frac{1}{p-1}}\right|^p\right)^{\frac{1}{p}} \\
&= \left(\left|\frac{\varepsilon}{2}\right|^p \sum_{i=1}^n |x_i^*|^{\frac{p}{p-1}}\right)^{\frac{1}{p}} \\
&= \left|\frac{\varepsilon}{2}\right| \left(\left(\sum_{i=1}^n |x_i^*|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}\right)^{\frac{1}{p-1}} \\
&= \frac{\varepsilon}{2}\|x^*\|_{\frac{p}{p-1}}^{\frac{1}{p-1}} \\
&= \frac{\varepsilon}{2}.
\end{aligned}$$

Zudem ist wegen $x^* \neq 0$ auch $y \neq 0$ und damit

$$\partial^-(-\|\cdot\|_p)(y) = \left\{ \left(-\frac{1}{\|y\|_p^{p-1}} \left(\operatorname{sgn}(y_1)|y_1|^{p-1}, \ldots, \operatorname{sgn}(y_n)|y_n|^{p-1} \right) \right) \right\}.$$

Weiter gilt

$$\begin{aligned} -\frac{1}{\|y\|_p^{p-1}} \operatorname{sgn}(y_i)|y_i|^{p-1} &= -\frac{\operatorname{sgn}\left(-\frac{\varepsilon}{2}\operatorname{sgn}(x_i^*)|x_i^*|^{\frac{1}{p-1}}\right)\left|-\frac{\varepsilon}{2}\operatorname{sgn}(x_i^*)|x_i^*|^{\frac{1}{p-1}}\right|^{p-1}}{\left|\frac{\varepsilon}{2}\right|^{p-1}} \\ &= -\frac{\operatorname{sgn}\left(-\operatorname{sgn}(x_i^*)\right)\left|-\frac{\varepsilon}{2}\right|^{p-1}|\operatorname{sgn}(x_i^*)|^{p-1}\left||x_i^*|^{\frac{1}{p-1}}\right|^{p-1}}{\left|\frac{\varepsilon}{2}\right|^{p-1}} \\ &= -(-\operatorname{sgn}(x_i^*))|x_i^*| \\ &= x_i^*, \end{aligned}$$

daher ist also

$$x^* \in \partial^-(-\|\cdot\|_p)(y).$$

Zu jedem $x^* \in \mathbb{R}^n$ mit $\|x^*\|_{\frac{p}{p-1}} = 1$ und jedem $\varepsilon > 0$ existiert also ein $y \in \mathbb{R}^n$ mit $\|y\|_p < \varepsilon$ und $x^* \in \partial^-(-\|\cdot\|_p)(y)$, daraus folgt nun

$$\left\{x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p}{p-1}} = 1\right\} \subseteq \bigcap_{\varepsilon>0} \bigcup_{\|y\|_p<\varepsilon} \partial^-(-\|\cdot\|_p)(y) = \partial_a(-\|\cdot\|_p)(0)$$

bzw. zusammen mit der obigen Implikation

$$\left\{x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p}{p-1}} = 1\right\} = \partial_a(-\|\cdot\|_p)(0)$$

Korollar 2.2.1 liefert jetzt sofort die Aussage des Satzes. ■

Aufgrund ihrer Bedeutung für Anwendungen geben wir für die negative ℓ_2-Norm das Limiting-Subdifferential auch hier explizit an:

Korollar 3.2.1 (*Subdifferential der negativen ℓ_2-Norm*)

Das Limiting-Subdifferential der negativen ℓ_2-Norm $-\|\cdot\|_2 : \mathbb{R}^n \to \mathbb{R}$ mit $(-\|\cdot\|_2)(x) = -\left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}$ ist gegeben durch

$$\partial_L(-\|\cdot\|_2)(x) = \left\{-\frac{x}{\|x\|_2}\right\} \text{ für } x \in \mathbb{R}^n \setminus \{0\} \text{ und}$$

$$\partial_L(-\|\cdot\|_2)(0) = \{x^* \in \mathbb{R}^n \mid \|x^*\|_2 = 1\}.$$

Beweis: Einsetzen von $p = 2$ in Satz 3.2.2 liefert sofort die Behauptung. ■

Analog zu Satz 3.1.3 aus dem vorherigen Unterkapitel 3.1 lässt auch Satz 3.2.2 die negative ℓ_1-Norm und die negative ℓ_∞-Norm als Spezialfälle aus. Das Limiting-Subdifferential der negativen ℓ_1-Norm wiederum wurde bereits in Satz 3.2.1 untersucht, wir müssen also nun noch das Limiting-Subdifferential der negativen ℓ_∞-Norm bestimmen. Ein Ergebnis dazu ist schon als Satz 3.4 bei Hillmann (2013) zu finden, wir wollen jedoch auch hier eine Umformulierung dieser Aussage sowie einen überarbeiteten und deutlich präziseren Beweis derselben präsentieren.

Satz 3.2.3 (*Subdifferential der negativen ℓ_∞-Norm*)

Die negative ℓ_∞-Norm $-\|\cdot\|_\infty : \mathbb{R}^n \to \mathbb{R}$ mit $(-\|\cdot\|_\infty)(x) = -\max_{i\in\{1,\dots,n\}} |x_i|$ hat das Limiting-Subdifferential

$$\begin{aligned}
&\partial_L(-\|\cdot\|_\infty)(x)\\
&= \Big\{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1,\dots,n\}\colon x_i^* x_i + |x_i^*| \max_{i\in\{1,\dots,n\}} |x_i| = 0,\\
&\qquad\qquad \sum_{i=1}^n |x_i^*| = 1, \|x^*\|_\infty = 1\Big\}\\
&= \Big\{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1,\dots,n\}\colon x_i^*(|x_i| - \max_{i\in\{1,\dots,n\}} |x_i|) = 0,\\
&\qquad\qquad \forall\, x_i^* \neq 0\colon x_i^* = -\operatorname{sgn}(x_i), \sum_{i=1}^n |x_i^*| = 1\Big\} \text{ für } x \in \mathbb{R}^n \setminus \{0\},\\
&\partial_L(-\|\cdot\|_\infty)(0) = \{x^* \in \mathbb{R}^n \mid \|x^*\|_1 = 1, \|x^*\|_\infty = 1\}.
\end{aligned}$$

Beweis: Auch $-\|\cdot\|_\infty$ ist Lipschitz-stetig und nimmt nur endliche Werte an, wir können daher in gewohnter Weise zunächst das Dini-Hadamard-Subdifferential nach Definition 2.2.3 und anschließend das approximierende Subdifferential nach Definition 2.2.4 berechnen.

Im Folgenden soll dabei zur besseren Notation für jedes $x \in \mathbb{R}^n$ mit J_x die Menge bezeichnet werden, welche die Indizes der betragsmäßig größten Komponenten von x enthält, also

$$J_x := \Big\{j \in \{1,\dots,n\} \mid |x_j| = \max_{i\in\{1,\dots,n\}} |x_i|\Big\} = \{j \in \{1,\dots,n\} \mid |x_j| = \|x\|_\infty\}\,.$$

In $x = 0$ müsste auch hier für jedes Element $x^* \in \partial^-(-\|\cdot\|_\infty)(0)$ des Dini-Hadamard-Subdifferentials und beliebiges $h \in \mathbb{R}^n$ sowohl

$$\begin{aligned}
\langle x^*, h\rangle &\le d^-(-\|\cdot\|_\infty)(x,h)\\
&= \liminf_{t\searrow 0} \frac{-\|x+th\|_\infty + \|x\|_\infty}{t}\\
&= \liminf_{t\searrow 0} \frac{-|t|\|h\|_\infty}{t}\\
&= -\|h\|_\infty
\end{aligned}$$

als auch

$$\begin{aligned}
-\langle x^*, h\rangle &= \langle x^*, -h\rangle\\
&\le d^-(-\|\cdot\|_\infty)(x,-h)\\
&= \liminf_{t\searrow 0} \frac{-\|x-th\|_\infty + \|x\|_\infty}{t}\\
&= \liminf_{t\searrow 0} \frac{-|-t|\|h\|_\infty}{t}\\
&= -\|h\|_\infty
\end{aligned}$$

gelten. Dies ist aufgrund von $\|h\|_\infty > 0$ für alle $h \neq 0$ jedoch für kein $x^* \in \mathbb{R}^n$ möglich, daher folgt $\partial^-(-\|\cdot\|_\infty)(0) = \emptyset$.

Es sei also nun $x \in \mathbb{R}^n$ mit $x \neq 0$. Ist für $k \in \{1, \dots, n\}$ zunächst $k \in J_x$, so betrachten wir den Richtungsvektor $h = (0, \dots, 0, x_k, 0, \dots, 0) \in \mathbb{R}^n$. Offenbar gilt auch $k \in J_{x+th}$ für jedes $t > 0$ und jedes $x^* \in \partial^-(-\|\cdot\|_\infty)(x)$ muss demzufolge

$$\begin{aligned}
x_k^* x_k &= \langle x^*, h \rangle \\
&\leq d^-(-\|\cdot\|_\infty)(x, h) \\
&= \liminf_{t \searrow 0} \frac{-\|x + th\|_\infty + \|x\|_\infty}{t} \\
&= \liminf_{t \searrow 0} \frac{-|x_k + tx_k| + |x_k|}{t} \\
&= \liminf_{t \searrow 0} \frac{-(1+t)|x_k| + |x_k|}{t} \\
&= \liminf_{t \searrow 0} \frac{-t|x_k|}{t} \\
&= -|x_k|
\end{aligned}$$

erfüllen. Wegen $x \neq 0$ und $k \in J_x$ ist auch $x_k \neq 0$, damit muss insbesondere

$$x_k^* x_k \leq -|x_k| < 0 \tag{3.6}$$

gelten.

Wir unterscheiden nun erneut zwei Fälle: Sei zunächst $|J_x| = 1$, dann ist $|x_k| > |x_i|$ für alle $i \in \{1, \dots, n\} \setminus \{k\}$. Für hinreichend kleines $t > 0$ ist dann immer auch $(1-t)|x_k| > |x_i|$ für alle $i \in \{1, \dots, n\} \setminus \{k\}$ und damit $k \in J_{x-th}$. Die Betrachtung des Richtungsvektors $-h$ liefert daher

$$\begin{aligned}
-x_k^* x_k &= \langle x^*, -h \rangle \\
&\leq d^-(-\|\cdot\|_\infty)(x, -h) \\
&= \liminf_{t \searrow 0} \frac{-\|x - th\|_\infty + \|x\|_\infty}{t} \\
&= \liminf_{t \searrow 0} \frac{-|x_k - tx_k| + |x_k|}{t} \\
&= \liminf_{t \searrow 0} \frac{-(1-t)|x_k| + |x_k|}{t} \\
&= \liminf_{t \searrow 0} \frac{t|x_k|}{t} \\
&= |x_k|.
\end{aligned}$$

Zusammen mit (3.6) ergibt sich für $k \in J_x$ mit $|J_x| = 1$ für jedes $x^* \in \partial^-(-\|\cdot\|_\infty)(x)$ also die Bedingung $x_k^* x_k = -|x_k|$ bzw. $x_k^* = -\operatorname{sgn}(x_k)$.

Im Falle $|J_x| \geq 2$ hingegen existiert ein $l \in \{1, \dots, n\} \setminus \{k\}$ mit $|x_l| = |x_k|$. Damit gilt $|x_l| > (1-t)|x_k|$ für jedes $t > 0$, also $k \notin J_{x-th}$, aber $l \in J_{x-th}$. Die Betrachtung des gleichen Richtungsvektors $-h$ führt daher hier zu

$$\begin{aligned}
-x_k^* x_k &= \langle x^*, -h \rangle \\
&\leq d^-(-\|\cdot\|_\infty)(x, -h) \\
&= \liminf_{t \searrow 0} \frac{-\|x - th\|_\infty + \|x\|_\infty}{t} \\
&= \liminf_{t \searrow 0} \frac{-|x_l| + |x_l|}{t} \\
&= 0.
\end{aligned}$$

Diese Ungleichung bildet aber sofort einen Widerspruch zu (3.6), somit folgt für $x \in \mathbb{R}^n \setminus \{0\}$ mit $|J_x| \geq 2$ immer $\partial^-(-\|\cdot\|_\infty)(x) = \emptyset$.

Ist hingegen schon $k \notin J_x$, so untersuchen wir stattdessen den Richtungsvektor $h = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^n$, welcher genau an der k-ten Position eine Eins und ansonsten nur Nullen enthält. Für hinreichend kleines $t > 0$ gilt dann sowohl $k \notin J_{x+th}$ als auch $k \notin J_{x-th}$ und wir erhalten die Ungleichungen

$$\begin{aligned}
\pm x_k^* &= \langle x^*, \pm h\rangle \\
&\leq d^-(-\|\cdot\|_\infty)(x, \pm h) \\
&= \liminf_{t \searrow 0} \frac{-\|x \pm th\|_\infty + \|x\|_\infty}{t} \\
&= \liminf_{t \searrow 0} \frac{-\max\limits_{i \in \{1,\ldots,n\}} |x_i \pm th_i| + \max\limits_{i \in \{1,\ldots,n\}} |x_i|}{t} \\
&= \liminf_{t \searrow 0} \frac{-|x_l| + |x_l|}{t} \\
&= 0.
\end{aligned}$$

Für jedes $x^* \in \partial^-(-\|\cdot\|_\infty)(x)$ muss also $x_k^* = 0$ für jedes $k \in \{1, \ldots, n\}$ mit $k \notin J_x$ gelten. Insgesamt kann das Dini-Hadamard-Subdifferential demzufolge nur für $x \in \mathbb{R}^n \setminus \{0\}$ mit $|J_x| = 1$ eine nichtleere Menge sein und jeder solche Dini-Hadamard-Subgradient $x^* \in \partial^-(-\|\cdot\|_\infty)(x)$ muss die Form $x_k^* = -\operatorname{sgn}(x_k)$ für $k \in J_x$ sowie $x_k^* = 0$ für $k \in \{1, \ldots, n\} \setminus J_x$ haben.

Es bleibt nun noch zu zeigen, dass dieser Vektor tatsächlich ein Dini-Hadamard-Subgradient von $-\|\cdot\|_\infty$ in x ist. Sei dafür $x \in \mathbb{R}^n \setminus \{0\}$ beliebig mit $|J_x| = \{k\}$ und $h \in \mathbb{R}^n$ beliebig, dann gilt für hinreichend kleines $t > 0$ stets $|x_k + th_k| > |x_i + th_i|$ für alle $i \in \{1, \ldots, n\} \setminus \{k\}$, also ist $k \in J_{x+th}$. Für hinreichend kleines $t > 0$ gilt wegen $x_k \neq 0$ weiterhin $\operatorname{sgn}(x_k + th_k) = \operatorname{sgn}(x_k)$, also folgt

$$\begin{aligned}
d^-(-\|\cdot\|_\infty)(x, h) &= \liminf_{t \searrow 0} \frac{-\|x + th\|_\infty + \|x\|_\infty}{t} \\
&= \liminf_{t \searrow 0} \frac{-|x_k + th_k| + |x_k|}{t} \\
&= \liminf_{t \searrow 0} \frac{-\operatorname{sgn}(x_k + th_k)(x_k + th_k) + \operatorname{sgn}(x_k)x_k}{t} \\
&= \liminf_{t \searrow 0} \frac{-\operatorname{sgn}(x_k)(x_k + th_k) + \operatorname{sgn}(x_k)x_k}{t} \\
&= \liminf_{t \searrow 0} \frac{-\operatorname{sgn}(x_k)th_k}{t} \\
&= -\operatorname{sgn}(x_k)h_k \\
&= \langle x^*, h\rangle .
\end{aligned}$$

Damit ist aber $x^* = (0, \ldots, 0, -\operatorname{sgn}(x_k), 0, \ldots, 0) \in \partial^-(-\|\cdot\|_\infty)(x)$, also gilt zusammenfassend $\partial^-(-\|\cdot\|_\infty)(x) = \{(0, \ldots, 0, -\operatorname{sgn}(x_k), 0, \ldots, 0)\}$ für jedes $x \in \mathbb{R}^n \setminus \{0\}$ mit $J_x = \{k\}$ und $\partial^-(-\|\cdot\|_\infty)(x) = \emptyset$ sonst.

Nun können wir das approximierende Subdifferential von $-\|\cdot\|_\infty$ bestimmen, im endlich-dimensionalen Raum $\mathbb{R}^n$ erhält man dieses für $x_0 \in \mathbb{R}^n$ wiederum mittels

$$\partial_a(-\|\cdot\|_\infty)(x) = \limsup_{y \to x} \partial^-(-\|\cdot\|_\infty)(y) = \bigcap_{\varepsilon > 0} \bigcup_{\|y - x\|_\infty < \varepsilon} \partial^-(-\|\cdot\|_\infty)(y).$$

Sei $x \in \mathbb{R}^n \setminus \{0\}$ beliebig, dann existiert zu jedem $\varepsilon > 0$ und jedem $k \in J_x$ ein $y \in \mathbb{R}^n$ mit $\|y - x\|_\infty < \varepsilon$ und $J_y = \{k\}$, hierfür kann man beispielsweise $y = (x_1, \ldots, x_{k-1}, x_k + 0{,}5 \cdot \operatorname{sgn}(x_k)\varepsilon, x_{k+1}, \ldots, x_n)$ wählen. Daraus folgt nun sofort

$$\{(0, \ldots, 0, -\operatorname{sgn}(x_k), 0, \ldots, 0)\} = \partial^-(-\|\cdot\|_\infty)(y) \subseteq \partial_a(-\|\cdot\|_\infty)(x)$$

für alle $k \in J_x$.

Andererseits gilt aber $J_y \subseteq J_x$ für alle $y \in \mathbb{R}^n$ mit $\|y - x\|_\infty < \varepsilon$, sofern $\varepsilon > 0$ nur hinreichend klein gewählt wird, da bei entsprechender Wahl von ε für alle $k \in J_x$ und alle $i \in \{1, \ldots, n\} \setminus J_x$ stets $|x_k| - \varepsilon > |x_i|$ gilt. Außerdem ist wegen $x \neq 0$ für hinreichend kleines $\varepsilon > 0$ auch $\operatorname{sgn}(y_k) = \operatorname{sgn}(x_k)$ für alle $y \in \mathbb{R}^n$ mit $\|y - x\|_\infty < \varepsilon$ und alle $k \in J_x$. Daher können keine weiteren als die soeben bestimmten Elemente in $\partial_a(-\|\cdot\|_\infty)(x)$ liegen, für $x \in \mathbb{R}^n \setminus \{0\}$ gilt unter Verwendung von Korollar 2.2.1 somit

$$\partial_L(-\|\cdot\|_\infty)(x) = \partial_a(-\|\cdot\|_\infty)(x) = \bigcup_{k \in J_x} \{(0, \ldots, 0, -\operatorname{sgn}(x_k), 0, \ldots, 0)\}. \tag{3.7}$$

Man sieht leicht, dass alle Limiting-Subgradienten aus (3.7) in den beiden in Satz 3.2.3 angegebenen Mengen liegen.

Ist andererseits $x^* \in \Big\{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1, \ldots, n\}\colon x_i^* x_i + |x_i^*| \max\limits_{i \in \{1,\ldots,n\}} |x_i| = 0,$ $\sum\limits_{i=1}^n |x_i^*| = 1, \|x^*\|_\infty = 1\Big\}$, so folgt bereits aus der ersten Bedingung $x_i^* = 0$ für alle $i \in \{1, \ldots, n\} \setminus J_x$ und $x_i^* = 0$ oder $\operatorname{sgn}(x_i^*) = -\operatorname{sgn}(x_i)$ für alle $i \in J_x$. Die dritte Bedingung liefert zudem, dass $|x_i^*| \leq 1$ für alle $i \in J_x$ und $|x_i^*| = 1$ für mindestens ein $i \in J_x$ gilt und aus der zweiten Bedingung folgt nun, dass $|x_k^*| = 1$ für genau ein $k \in J_x$ und $x_i^* = 0$ für alle $i \in \{1, \ldots, n\} \setminus \{k\}$ ist. Damit ist aber gerade beschrieben, dass x^* in der Menge aus (3.7) liegt.

Gilt hingegen zuletzt $x^* \in \Big\{x^* \in \mathbb{R}^n \mid \forall\, i \in \{1, \ldots, n\}\colon x_i^*(|x_i| - \max\limits_{i \in \{1,\ldots,n\}} |x_i|) = 0,$ $\forall\, x_i^* \neq 0\colon x_i^* = -\operatorname{sgn}(x_i), \sum\limits_{i=1}^n |x_i^*| = 1\Big\}$, dann folgt aus der dritten Bedingung zunächst $|x_i^*| \leq 1$ für alle $i \in \{1, \ldots, n\}$ und aus der zweiten Bedingung damit, dass $|x_k^*| = 1$ mit $x_k^* = -\operatorname{sgn}(x_k)$ für genau ein $k \in \{1, \ldots, n\}$ und $x_i^* = 0$ für alle übrigen Komponenten gilt. Die erste Bedingung liefert abschließend $k \in J_x$, womit wiederum x^* in der Menge aus (3.7) enthalten ist.

Es fehlt nun noch das approximierende Subdifferental im Punkt $x = 0$. Bezeichnet e_i den n-dimensionalen Vektor, welcher an der i-ten Stelle eine Eins und ansonsten nur Nullen enthält, so gilt natürlich

$$\partial_a(-\|\cdot\|_\infty)(0) \subseteq \bigcup_{i=1}^n \{e_i, -e_i\},$$

da in der rechten Menge alle möglichen Dini-Hadamard-Subgradienten von $-\|\cdot\|_\infty$ zu Elementen aus $\mathbb{R}^n$ enthalten sind.

Andererseits kann man für beliebiges $\varepsilon > 0$ stets die Menge

$$Y_\varepsilon := \bigcup_{i=1}^n \left\{\frac{\varepsilon}{2} \cdot e_i, -\frac{\varepsilon}{2} \cdot e_i\right\}$$

betrachten. Es ist $|J_y| = 1$ für alle $y \in Y_\varepsilon$, daher existieren Dini-Hadamard-Subgradienten von $-\|\cdot\|_\infty$ zu allen Elementen aus Y_ε und für alle $i \in \{1,\ldots,n\}$ ist gerade $\partial^-(-\|\cdot\|_\infty)\left(\pm\frac{\varepsilon}{2}\cdot e_i\right) = \{\mp e_i\}$. Weiter gilt $\|y-0\|_\infty < \varepsilon$ für alle $y \in Y_\varepsilon$, damit erhält man aber schon

$$\bigcup_{i=1}^{n}\{e_i,-e_i\} = \bigcup_{y\in Y_\varepsilon} \partial^-(-\|\cdot\|_\infty)(y) \subseteq \partial_a(-\|\cdot\|_\infty)(0).$$

Aus beiden Inklusionen und mit Korollar 2.2.1 folgt nun sofort

$$\bigcup_{i=1}^{n}\{e_i,-e_i\} = \partial_L(-\|\cdot\|_\infty)(0).$$

Für jedes Element $x^* \in \partial_L(-\|\cdot\|_\infty)(0)$ gilt offenbar $\|x^*\|_1 = \|x^*\|_\infty = 1$. Ist andererseits $x^* \in \mathbb{R}^n$ mit $\|x^*\|_1 = \|x^*\|_\infty = 1$, so folgt aus $\|x^*\|_\infty = 1$, dass $|x_i^*| \leq 1$ für alle $i \in \{1,\ldots,n\}$ und $|x_i^*| = 1$ für mindestens ein $i \in \{1,\ldots,n\}$ gilt. Mit $\|x^*\|_1 = 1$ folgt dann weiter, dass $|x_k^*| = 1$ für genau ein $k \in \{1,\ldots,n\}$ und $x_i^* = 0$ für alle übrigen Komponenten von x^* gilt. Damit gilt aber $x^* \in \partial_L(-\|\cdot\|_\infty)(0)$ und die letzte Aussage des Satzes ist gezeigt. ■

Bemerkung 3.2.1

In ihrer Dissertation untersuchte Wagner (2015) das skalare Medianproblem (SP) mit Abstandsfunktionen, welche von polyedrischen Gauges induziert werden. Unter Nutzung der von Toland (1978) und Singer (1979) erarbeiteten Dualitätstheorie leitete sie dabei auf geometrischen Abbildungen basierende Lösungsalgorithmen für diese Problemstellungen her, deren Fortentwicklung von Wagner, Martínez-Legaz und Tammer (2016) publiziert wurde. Da die zur ℓ_1-Norm und zur ℓ_∞-Norm gehörigen Einheitskugeln auch polyedrische Gauges sind, entsprechen die von diesen Normen induzierten Abstandsfunktionen gerade den von Wagner (2015) bzw. Wagner, Martínez-Legaz und Tammer (2016) untersuchten. Unsere in Satz 3.2.1 und Satz 3.2.3 gewonnenen Ergebnisse formalisieren nun genau diesen Teil der dabei erzielten Resultate.

3.3 Limiting-Subdifferentiale des Skalarisierungsfunktionals und der Maximumsfunktion

Zum Abschluss dieses Kapitels wollen wir nun noch die Limiting-Subdifferentiale zweier Funktionen angeben, welche grundlegend für den Umgang mit bestimmten Klassen der von uns betrachteten Standortoptimierungsprobleme sind: Die Maximumsfunktion ist ein zentraler Bestandteil des skalaren Centerproblems (CP) bzw. des skalaren Centdianproblems (CDP) und das nichtlineare Skalarisierungsfunktional aus Sektion 2.1.4 benötigen wir, um notwendige Optimalitätsbedingungen für die mehrkriterielle Formulierung (MP) erarbeiten zu können.

Zuerst werden wir das Limiting-Subdifferential der Maximumsfunktion auf einem Asplund-Raum $(X,\|\cdot\|_X)$ untersuchen, das heißt, einer Funktion der Form

$$\left(\max_{i\in\{1,\ldots,k\}} f_i\right)(x) := \max_{i\in\{1,\ldots,k\}} f_i(x)$$

mit Teilfunktionen $f_1, \ldots, f_k \colon X \to \overline{\mathbb{R}}$. Die Minimierung einer solchen Maximumsfunktion steht unter anderem auch in enger Verbindung zur Tschebyscheff-Skalarisierung, welche Tammer und Weidner (2020) in Sektion 7.3.1 beschreiben. Das folgende Resultat ist Punkt (ii) in Theorem 3.46 bei Mordukhovich (2006a).

Satz 3.3.1 (*Limiting-Subdifferential der Maximumsfunktion*)

Es seien $(X, \|\cdot\|_X)$ ein Asplund-Raum und $f_1, \ldots, f_k \colon X \to \overline{\mathbb{R}}$ mit $k \geq 2$ Funktionen, welche in $x \in \bigcap_{i=1}^{k} \operatorname{dom} f_i$ lokal Lipschitz-stetig sind. Dann gilt

$$\partial_L \left(\max_{i \in \{1,\ldots,k\}} f_i \right)(x) \subseteq \bigcup_{\lambda \in \Lambda_x} \partial_L \left(\sum_{i \in I_x} \lambda_i f_i \right)(x) = \bigcup_{\lambda \in \Lambda_x} \partial_L \left(\sum_{i=1}^{k} \lambda_i f_i \right)(x),$$

wobei

$$\Lambda_x := \left\{ \lambda \in \mathbb{R}^k_+ \mid \forall\, i \in \{1,\ldots,k\} \colon \lambda_i \left(f_i(x) - \left(\max_{i \in \{1,\ldots,k\}} f_i \right)(x) \right) = 0, \sum_{i=1}^{k} \lambda_i = 1 \right\}$$

und

$$I_x := \left\{ i \in \{1,\ldots,k\} \mid f_i(x) = \left(\max_{i \in \{1,\ldots,k\}} f_i \right)(x) \right\}$$

sein sollen.

Ein Resultat zum Limiting-Subdifferential des nichtlinearen Skalarisierungsfunktionals, welches in dieser Arbeit zur Bestimmung notwendiger Optimalitätsbedingungen des Problems (MP) in Kapitel 7 genutzt werden soll, wurde von uns bereits als Proposition 3.1 in Bao, Hillmann und Tammer (2017) veröffentlicht.

Der Beweis dieser Aussage basiert in hohem Maße auf der Nutzung von Ko-Ableitungen und Resultaten über diese. Da der folgende Satz jedoch auch das einzige Ergebnis unserer Arbeit ist, für welches Ko-Ableitungen eine Rolle spielen und zugleich die ihnen zugrunde liegende Theorie sehr umfangreich ist, wurde auf die Einführung von Ko-Ableitungen in dieser Arbeit bewusst verzichtet. Für den Beweis des folgenden Resultates sei daher nochmals auf Bao, Hillmann und Tammer (2017) verwiesen.

Satz 3.3.2 (*Limiting-Subdifferential des Skalarisierungsfunktionals*)

Es seien $(X, \|\cdot\|_X)$ ein Asplund-Raum, $A \subset X$ nichtleer und abgeschlossen sowie $k \in X \setminus \{0\}$ ein Richtungsvektor, so dass $A + (0, +\infty) \cdot k \subseteq \operatorname{int} A$ gilt (also (2.3) erfüllt ist). Dann hat das durch (2.1) definierte nichtlineare Skalarisierungsfunktional $\varphi_{A,k}$ in $x \in \operatorname{dom} \varphi_{A,k}$ das Limiting-Subdifferential

$$\partial_L \varphi_{A,k}(x) = \{x^* \in X^* \mid x^*(k) = 1, -x^* \in N_L(\varphi_{A,k}(x)k - x, \operatorname{bd} A)\}.$$

Damit stehen uns nun alle Resultate zur Verfügung, welche wir in dieser Arbeit zur Herleitung von notwendigen Optimalitätsbedingungen benötigen. Wir werden dazu in den nächsten Kapiteln die Ergebnisse dieses Kapitels mit den grundlegenden Eigenschaften der Subdifferentiale aus Unterkapitel 2.2 verknüpfen und außerdem die generelle Struktur der Problemstellungen für uns nutzen.

4 Notwendige Optimalitätsbedingungen für nichtkonvexe skalare Medianprobleme

In diesem und den nächsten drei Kapiteln präsentieren wir die primären Resultate unserer Arbeit in der Form von notwendigen Optimalitätsbedingungen für die in Kapitel 1 vorgestellten Klassen von Standortoptimierungsproblemen. Wir werden dabei im Folgenden der allgemein üblichen Benennung folgen und Problemstellungen mit nichtkonvexer Zielfunktion verkürzt als nichtkonvexe Probleme bezeichnen. Weiterhin werden wir für den restlichen Verlauf der Arbeit davon ausgehen, dass jedem der existierenden Standorte jede mögliche Abstandsfunktion höchstens einmal zugeordnet wird, dies stellt schon aus der reinen Anschauung eine sinnvolle Konvention dar.

Als erste Modellierung betrachten wir nun das nichtkonvexe skalare Medianproblem (SP) aus Sektion 1.2.2 als Verallgemeinerung des „klassischsten" Standortoptimierungsproblems. Die dabei erhaltenen Ergebnisse sind von besonderer Wichtigkeit, da alle notwendigen Optimalitätsbedingungen für die in den folgenden Kapiteln betrachteten Probleme eng mit diesen Resultaten verknüpft sind.

4.1 Notwendige Optimalitätsbedingungen für Medianprobleme mit speziellen Metriken

Wir wollen zunächst solche Aufgabenstellungen behandeln, in denen jedem der existierenden anziehenden und abstoßenden Standorte a^k mit $k = 1, \ldots, m$ die gleiche Minkowski-Metrik als Abstandsfunktion zum neu zu bestimmenden Standort x zugeordnet wird, es ist demnach $d_k(x, a^k) = d(x, a^k) = \|x - a^k\|_p$ für $k = 1, \ldots, m$ und $p \in \overline{\mathbb{R}}$ mit $p \geq 1$. Dieses Problem hat in der Klassifikation von Standortoptimierungsproblemen nach Sektion 1.2.1 das Schema $1/\mathbb{R}^n/w_k \gtrless 0/\ell_p/\Sigma$.

Zuerst untersuchen wir hierbei den Fall, dass allen Standorten als Abstandsfunktion die von der ℓ_1-Norm induzierte Metrik zugeordnet ist, also Problemstellungen des Schemas $1/\mathbb{R}^n/w_k \gtrless 0/\ell_1/\Sigma$. Dieser Satz ist der einzige in diesem Kapitel, welcher kein neues Ergebnis dieser Arbeit, sondern bereits als Korollar 4.1 bei Hillmann (2013) zu finden ist.

Satz 4.1.1 (***Notwendige Optimalitätsbedingung für $1/\mathbb{R}^n/w_k \gtrless 0/\ell_1/\Sigma$***)
Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (SP) mit $d_k(x, a^k) = \|x - a^k\|_1$ für alle $k \in \{1, \ldots, m\}$, dann gilt für alle $l \in \{1, \ldots, n\}$

$$0 \in \sum_{i=1}^{r} \begin{cases} \{|w_i|\} & \text{für } \overline{x}_l > a_l^i \\ \{-|w_i|\} & \text{für } \overline{x}_l < a_l^i \\ [-|w_i|, |w_i|] & \text{für } \overline{x}_l = a_l^i \end{cases} + \sum_{j=r+1}^{m} \begin{cases} \{-|w_j|\} & \text{für } \overline{x}_l > a_l^j \\ \{|w_j|\} & \text{für } \overline{x}_l < a_l^j \\ \{-|w_j|, |w_j|\} & \text{für } \overline{x}_l = a_l^j \end{cases}.$$

Beweis: Es sei $\overline{x}$ ein lokales Minimum von (SP) mit $d_k(x, a^k) = \|x - a^k\|_1$ für alle $k \in \{1, \ldots, m\}$. Nach Eigenschaft (L4) des Limiting-Subdifferentials aus Lemma 2.2.2 muss dann

$$\begin{aligned} 0 &\in \partial_L f_w(\overline{x}) \\ &= \partial_L \left(\sum_{i=1}^{r} |w_i| \|x - a^i\|_1 + \sum_{j=r+1}^{m} |w_j| (-\|x - a^j\|_1) \right) (\overline{x}) \end{aligned}$$

gelten.

Wir formen diese Bedingung mithilfe der Summenregel für Limiting-Subdifferentiale (Satz 2.2.8) und Eigenschaft (L1) aus Lemma 2.2.2, der positiven Homogenität des Limiting-Subdifferentials, weiter um und erhalten

$$\begin{aligned} 0 &\in \partial_L \left(\sum_{i=1}^{r} |w_i| \|x - a^i\|_1 + \sum_{j=r+1}^{m} |w_j| (-\|x - a^j\|_1) \right) (\overline{x}) \\ &\subseteq \left(\sum_{i=1}^{r} \partial_L(|w_i| \|x - a^i\|_1) + \sum_{j=r+1}^{m} \partial_L(|w_j| (-\|x - a^j\|_1)) \right) (\overline{x}) \\ &= \left(\sum_{i=1}^{r} |w_i| \partial_L(\|x - a^i\|_1) + \sum_{j=r+1}^{m} |w_j| \partial_L(-\|x - a^j\|_1) \right) (\overline{x}). \end{aligned}$$

Setzen wir nun unsere Ergebnisse zum Limiting-Subdifferential der ℓ_1-Norm aus Satz 3.1.2 und zum Limiting-Subdifferential der negativen ℓ_1-Norm aus Satz 3.2.1 ein, so folgt für die Bedingung

$$\begin{aligned} 0 &\in \left(\sum_{i=1}^{r} |w_i| \partial_L(\|x - a^i\|_1) + \sum_{j=r+1}^{m} |w_j| \partial_L(-\|x - a^j\|_1) \right) (\overline{x}) \\ &= \sum_{i=1}^{r} |w_i| \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\} \colon x^*_l \in \begin{cases} \{1\} & \text{für } \overline{x}_l - a^i_l > 0 \\ \{-1\} & \text{für } \overline{x}_l - a^i_l < 0 \\ [-1, 1] & \text{für } \overline{x}_l - a^i_l = 0 \end{cases} \right\} \\ &\quad + \sum_{j=r+1}^{m} |w_j| \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\} \colon x^*_l \in \begin{cases} \{-1\} & \text{für } \overline{x}_l - a^j_l > 0 \\ \{1\} & \text{für } \overline{x}_l - a^j_l < 0 \\ \{-1, 1\} & \text{für } \overline{x}_l - a^j_l = 0 \end{cases} \right\} \\ &= \sum_{i=1}^{r} \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\} \colon x^*_l \in \begin{cases} \{|w_i|\} & \text{für } \overline{x}_l > a^i_l \\ \{-|w_i|\} & \text{für } \overline{x}_l < a^i_l \\ [-|w_i|, |w_i|] & \text{für } \overline{x}_l = a^i_l \end{cases} \right\} \\ &\quad + \sum_{j=r+1}^{m} \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\} \colon x^*_l \in \begin{cases} \{-|w_j|\} & \text{für } \overline{x}_l > a^j_l \\ \{|w_j|\} & \text{für } \overline{x}_l < a^j_l \\ \{-|w_j|, |w_j|\} & \text{für } \overline{x}_l = a^j_l \end{cases} \right\}. \end{aligned}$$

Die Komponenten der Vektoren, welche in der Menge auf der rechten Seite dieser notwendigen Optimalitätsbedingung enthalten sind, sind stets alle voneinander unabhängig. Daher ergibt sich durch komponentenweise Betrachtung nun sofort die Aussage des Satzes. ∎

Bemerkung 4.1.1

In Satz 4.1.1 wurden die Absolutbeträge der Gewichte in der notwendigen Optimalitätsbedingung beibehalten, um die Struktur dieser Bedingung besser zu verdeutlichen. Natürlich lassen sich diese Absolutbeträge wegen $w_i > 0$ für alle $i \in \{1, \ldots, r\}$ und $w_j < 0$ für alle $j \in \{r+1, \ldots, m\}$ aber auch auflösen, eine äquivalente Formulierung der Optimalitätsbedingung in Satz 4.1.1 ist daher:

$$0 \in \sum_{i=1}^{r} \begin{cases} \{w_i\} & \text{für } \overline{x}_l > a_l^i \\ \{-w_i\} & \text{für } \overline{x}_l < a_l^i \\ [-w_i, w_i] & \text{für } \overline{x}_l = a_l^i \end{cases} + \sum_{j=r+1}^{m} \begin{cases} \{w_j\} & \text{für } \overline{x}_l > a_l^j \\ \{-w_j\} & \text{für } \overline{x}_l < a_l^j \\ \{-w_j, w_j\} & \text{für } \overline{x}_l = a_l^j \end{cases}.$$

Auch im weiteren Verlauf dieser Arbeit werden wir zum Teil auf Kompaktheit der Resultate verzichten, um dafür die Struktur der Optimalitätsbedingungen besser herausstellen zu können.

Als Nächstes wollen wir eine notwendige Optimalitätsbedingung für Problemstellungen mit dem Klassifikationsschema $1/\mathbb{R}^n/w_k \gtrless 0/\ell_{p,1<p<+\infty}/\Sigma$ präsentieren, also nichtkonvexe Medianprobleme des Typs (SP) untersuchen, bei welchen alle Abstandsfunktionen von derselben ℓ_p-Norm mit $p \in \mathbb{R}$ und $p > 1$ induziert werden. Dabei handelt es sich um ein komplett neues Ergebnis dieser Promotion.

Satz 4.1.2 (*Notw. Optimalitätsbedingung für* $\mathbf{1/\mathbb{R}^n/w_k \gtrless 0/\ell_{p,1<p<+\infty}/\Sigma}$)

Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (SP) mit $d_k(x, a^k) = \|x - a^k\|_p$ für ein $p \in \mathbb{R}$ mit $p > 1$ und für alle $k \in \{1, \ldots, m\}$, dann ist entweder $\overline{x} = a^k$ für ein $k \in \{1, \ldots, r\}$ und es gilt

$$\Bigg\| \sum_{j=r+1}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}} \right) - \sum_{\substack{i=1 \\ i \neq k}}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} \right) \Bigg\|_{\frac{p}{p-1}} \leq |w_k|$$

oder es ist $\overline{x} = a^k$ für ein $k \in \{r+1, \ldots, m\}$ und es gilt

$$\Bigg\| \sum_{\substack{j=r+1 \\ j \neq k}}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}} \right) - \sum_{i=1}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} \right) \Bigg\|_{\frac{p}{p-1}} = |w_k|$$

oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ und $\overline{x}$ löst das System nichtlinearer Gleichungen

$$0 = \sum_{i=1}^{r} |w_i| \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p-1}}{\|\overline{x} - a^i\|_p^{p-1}} - \sum_{j=r+1}^{m} |w_j| \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p-1}}{\|\overline{x} - a^j\|_p^{p-1}}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis : Es sei $p \in \mathbb{R}$ mit $p > 1$ beliebig, aber fest und $\overline{x}$ ein lokales Minimum von (SP) mit $d_k(x, a^k) = \|x - a^k\|_p$ für alle $k \in \{1, \dots, m\}$. Wir nutzen wieder Eigenschaft (L4) des Limiting-Subdifferentials aus Lemma 2.2.2, formen diese notwendige Optimalitätsbedingung mit der Summenregel für Limiting-Subdifferentiale (Satz 2.2.8) sowie Eigenschaft (L1) aus demselben Lemma 2.2.2 weiter um und erhalten so

$$\begin{aligned}
0 &\in \partial_L f_w(\overline{x}) \\
&= \partial_L \left(\sum_{i=1}^{r} |w_i| \|x - a^i\|_p + \sum_{j=r+1}^{m} |w_j| (-\|x - a^j\|_p) \right) (\overline{x}) \\
&\subseteq \left(\sum_{i=1}^{r} \partial_L (|w_i| \|x - a^i\|_p) + \sum_{j=r+1}^{m} \partial_L (|w_j| (-\|x - a^j\|_p)) \right) (\overline{x}) \\
&= \left(\sum_{i=1}^{r} |w_i| \partial_L (\|x - a^i\|_p) + \sum_{j=r+1}^{m} |w_j| \partial_L (-\|x - a^j\|_p) \right) (\overline{x}).
\end{aligned}$$

Nun können wir die drei Fälle aus dem Satz unterscheiden und jeweils unsere Ergebnisse für das Limiting-Subdifferential der ℓ_p-Norm aus Satz 3.1.3 und für das Limiting-Subdifferential der negativen ℓ_p-Norm aus Satz 3.2.2 einsetzen. Ist zunächst $\overline{x}$ einer der existierenden Standorte mit anziehender Wirkung, also $\overline{x} = a^k$ für ein $k \in \{1, \dots, r\}$, dann folgt

$$\begin{aligned}
0 &\in \left(\sum_{i=1}^{r} |w_i| \partial_L (\|x - a^i\|_p) + \sum_{j=r+1}^{m} |w_j| \partial_L (-\|x - a^j\|_p) \right) (a^k) \\
&= \sum_{\substack{i=1 \\ i \neq k}}^{r} |w_i| \left\{ \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \dots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} \right) \right\} \\
&\quad + |w_k| \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p}{p-1}} \leq 1 \right\} \\
&\quad + \sum_{j=r+1}^{m} |w_j| \left\{ \left(-\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}, \dots, -\frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}} \right) \right\}.
\end{aligned}$$

Das ist jedoch schon äquivalent zu

$$\begin{aligned}
&\sum_{j=r+1}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}, \dots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}} \right) \\
&\quad - \sum_{\substack{i=1 \\ i \neq k}}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \dots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} \right) \\
&\quad \in \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p}{p-1}} \leq |w_k| \right\}
\end{aligned}$$

bzw. zu

$$\begin{aligned}
&\Bigg\| \sum_{j=r+1}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}, \dots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}} \right) \\
&\quad - \sum_{\substack{i=1 \\ i \neq k}}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \dots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} \right) \Bigg\|_{\frac{p}{p-1}} \leq |w_k|.
\end{aligned}$$

Ist $\overline{x}$ hingegen einer der existierenden Standorte mit abstoßender Wirkung, also $\overline{x} = a^k$ für ein $k \in \{r+1, \ldots, m\}$, so liefert das Einsetzen in analoger Weise

$$
\begin{aligned}
0 \in & \left(\sum_{i=1}^{r} |w_i| \partial_L(\|x - a^i\|_p) + \sum_{j=r+1}^{m} |w_j| \partial_L(-\|x - a^j\|_p) \right) (a^k) \\
= & \sum_{i=1}^{r} |w_i| \left\{ \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} \right) \right\} \\
& + |w_k| \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p}{p-1}} = 1 \right\} \\
& + \sum_{\substack{j=r+1 \\ j \neq k}}^{m} |w_j| \left\{ \left(-\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}, \ldots, -\frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}} \right) \right\}.
\end{aligned}
$$

Dies lässt sich wieder äquivalent umformen, zuerst in

$$
\begin{aligned}
& \sum_{\substack{j=r+1 \\ j \neq k}}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}} \right) \\
& - \sum_{i=1}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} \right) \\
& \in \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p}{p-1}} = |w_k| \right\},
\end{aligned}
$$

und schließlich zu

$$
\begin{aligned}
& \Bigg\| \sum_{\substack{j=r+1 \\ j \neq k}}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}} \right) \\
& - \sum_{i=1}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} \right) \Bigg\|_{\frac{p}{p-1}} = |w_k|.
\end{aligned}
$$

Zuletzt betrachten wir noch den allgemeinsten Fall $\overline{x} \notin \{a^1, \ldots, a^m\}$. Hier liefert das Einsetzen der berechneten Limiting-Subdifferentiale die Bedingung

$$
\begin{aligned}
0 \in & \left(\sum_{i=1}^{r} |w_i| \partial_L(\|x - a^i\|_p) + \sum_{j=r+1}^{m} |w_j| \partial_L(-\|x - a^j\|_p) \right) (\overline{x}) \\
= & \sum_{i=1}^{r} |w_i| \left\{ \left(\frac{\operatorname{sgn}(\overline{x}_1 - a_1^i)|\overline{x}_1 - a_1^i|^{p-1}}{\|\overline{x} - a^i\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(\overline{x}_n - a_n^i)|\overline{x}_n - a_n^i|^{p-1}}{\|\overline{x} - a^i\|_p^{p-1}} \right) \right\} \\
& + \sum_{j=r+1}^{m} |w_j| \left\{ \left(-\frac{\operatorname{sgn}(\overline{x}_1 - a_1^j)|\overline{x}_1 - a_1^j|^{p-1}}{\|\overline{x} - a^j\|_p^{p-1}}, \ldots, -\frac{\operatorname{sgn}(\overline{x}_n - a_n^j)|\overline{x}_n - a_n^j|^{p-1}}{\|\overline{x} - a^j\|_p^{p-1}} \right) \right\}.
\end{aligned}
$$

Die Menge auf der rechten Seite der Gleichung ist nur einelementig, es muss also sogar die Gleichheit

$$
\begin{aligned}
0 = & \sum_{i=1}^{r} |w_i| \left(\frac{\operatorname{sgn}(\overline{x}_1 - a_1^i)|\overline{x}_1 - a_1^i|^{p-1}}{\|\overline{x} - a^i\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(\overline{x}_n - a_n^i)|\overline{x}_n - a_n^i|^{p-1}}{\|\overline{x} - a^i\|_p^{p-1}} \right) \\
& + \sum_{j=r+1}^{m} |w_j| \left(-\frac{\operatorname{sgn}(\overline{x}_1 - a_1^j)|\overline{x}_1 - a_1^j|^{p-1}}{\|\overline{x} - a^j\|_p^{p-1}}, \ldots, -\frac{\operatorname{sgn}(\overline{x}_n - a_n^j)|\overline{x}_n - a_n^j|^{p-1}}{\|\overline{x} - a^j\|_p^{p-1}} \right)
\end{aligned}
$$

gelten. Komponentenweise Betrachtung dieser notwendigen Optimalitätsbedingung liefert uns jetzt sofort die Behauptung. ■

Aufgrund ihrer praktischen Relevanz betrachten wir auch hier die ℓ_2-Norm noch einmal separat. Das folgende Korollar ist dabei bereits als Korollar 4.3 von Hillmann (2013) bewiesen worden, hier folgt es nun als Spezialfall aus dem deutlich umfänglicheren Satz 4.1.2.

Korollar 4.1.1 (*Notw. Optimalitätsbedingung für* $1/\mathbb{R}^n/w_k \gtrless 0/\ell_2/\Sigma$)

Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (SP) mit $d_k(x, a^k) = \|x - a^k\|_2$ für alle $k \in \{1, \ldots, m\}$, dann ist entweder $\overline{x} = a^k$ für ein $k \in \{1, \ldots, r\}$ und es gilt

$$\left\| \sum_{j=r+1}^{m} |w_j| \frac{a^k - a^j}{\|a^k - a^j\|_2} - \sum_{\substack{i=1 \\ i \neq k}}^{r} |w_i| \frac{a^k - a^i}{\|a^k - a^i\|_2} \right\|_2 \leq |w_k|$$

oder es ist $\overline{x} = a^k$ für ein $k \in \{r+1, \ldots, m\}$ und es gilt

$$\left\| \sum_{\substack{j=r+1 \\ j \neq k}}^{m} |w_j| \frac{a^k - a^j}{\|a^k - a^j\|_2} - \sum_{i=1}^{r} |w_i| \frac{a^k - a^i}{\|a^k - a^i\|_2} \right\|_2 = |w_k|$$

oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ und $\overline{x}$ löst das System nichtlinearer Gleichungen

$$0 = \sum_{i=1}^{r} |w_i| \frac{\overline{x}_l - a_l^i}{\|\overline{x} - a^i\|_2} - \sum_{j=r+1}^{m} |w_j| \frac{\overline{x}_l - a_l^j}{\|\overline{x} - a^j\|_2}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis : Die Behauptung folgt durch Einsetzen von $p = 2$ sofort aus Satz 4.1.2. ■

Bemerkung 4.1.2

Mithilfe der exakt gleichen Beweisschritte, wie sie soeben im Beweis von Satz 4.1.2 angewandt wurden, lassen sich auch allgemeinere Resultate für Aufgabenstellungen erzielen, bei welchen die Abstandsfunktionen nicht alle von derselben ℓ_p-Norm induziert werden. Für solche Problemstellungen des Schemas $1/\mathbb{R}^n/w_k \gtrless 0/\ell^m_{p,1<p<+\infty}/\Sigma$ können zwei Fälle unterschieden werden, welche wiederum Spezialfälle von Satz 4.2.1 aus dem folgenden Unterkapitel 4.2 sind.

(i) Sind die existierenden Standorte $a^1, \ldots, a^m$ alle paarweise verschieden, so ändern sich gegenüber Satz 4.1.2 nur die durch die induzierenden Normen bedingten Variablen und man gelangt mit den obigen Beweisschritten zu folgender Aussage:

Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (SP) und sind alle existierenden Standorte $a^1, \ldots, a^m$ paarweise verschieden sowie $d_k(x, a^k) = \|x - a^k\|_{p_k}$ mit $p_k \in \mathbb{R}$ und $p_k > 1$ für alle $k \in \{1, \ldots, m\}$, dann ist entweder $\overline{x} = a^k$ für ein $k \in \{1, \ldots, r\}$ und es gilt

$$\Bigg\| \sum_{j=r+1}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}} \right)$$
$$- \sum_{\substack{i=1\\ i\neq k}}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}} \right) \Bigg\|_{\frac{p_k}{p_k-1}} \leq |w_k|$$

oder es ist $\overline{x} = a^k$ *für ein* $k \in \{r+1, \ldots, m\}$ *und es gilt*

$$\Bigg\| \sum_{\substack{j=r+1\\ j\neq k}}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}} \right)$$
$$- \sum_{i=1}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}} \right) \Bigg\|_{\frac{p_k}{p_k-1}} = |w_k|$$

oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ *und* $\overline{x}$ *löst das System nichtlinearer Gleichungen*

$$0 = \sum_{i=1}^{r} |w_i| \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} |w_j| \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}$$

für alle $l \in \{1, \ldots, n\}$.

(ii) Ein existierender Standort, welchem mehr als eine Abstandsfunktion zugeordnet wird, entspricht in unserer Formulierung zwei oder mehr gleichen Standorten mit jeweils einer unterschiedlichen zugeordneten Abstandsfunktion. Wir fassen die Indizes dieser gleichen Standorte in Indexmengen K_s *mit* $s \in \{1, \ldots, u\}$ *zusammen, indem wir fordern, dass für diese*

(a) $\bigcup_{s=1}^{u} K_s = \{1, \ldots, m\}$,

(b) $a^g = a^h$ *für alle* $g, h \in K_s$ *und für alle* $s \in \{1, \ldots, u\}$ *sowie*

(c) $a^g \neq a^h$ *für alle* $g \in K_s, h \in K_t$ *mit* $s, t \in \{1, \ldots, u\}$ *und* $s \neq t$

gelten. Mit diesen Indexmengen lässt sich nun eine Aussage für den Fall treffen, dass die existierenden Standorte $a^1, \ldots, a^m$ *nicht alle paarweise verschieden sind. Wenn wir hierfür die Schritte aus dem Beweis von Satz 4.1.2 nachvollziehen, so ändert sich zusätzlich zu Fall (i) noch die notwendige Optimalitätsbedingung für die existierenden Standorte und wir erhalten folgende Aussage:*

Ist $\overline{x} \in \mathbb{R}^n$ *ein lokales Minimum des Problems (SP), die existierenden Standorte in Indexmengen* K_s *mit* $s \in \{1, \ldots, u\}$ *aufgeteilt, welche (a) - (c) genügen, sowie* $d_k(x, a^k) = \|x - a^k\|_{p_k}$ *mit* $p_k \in \mathbb{R}$ *und* $p_k > 1$ *für alle* $k \in \{1, \ldots, m\}$, *dann ist entweder* $\overline{x} = a^g$ *für ein* $g \in K_s$ *mit* $s \in \{1, \ldots, u\}$ *und es gilt*

$$\sum_{\substack{j=r+1\\ j\notin K_s}}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^g - a_1^j)|a_1^g - a_1^j|^{p_j-1}}{\|a^g - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(a_n^g - a_n^j)|a_n^g - a_n^j|^{p_j-1}}{\|a^g - a^j\|_{p_j}^{p_j-1}} \right)$$

$$- \sum_{\substack{i=1\\ i\notin K_s}}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^g - a_1^i)|a_1^g - a_1^i|^{p_i-1}}{\|a^g - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a_n^g - a_n^i)|a_n^g - a_n^i|^{p_i-1}}{\|a^g - a^i\|_{p_i}^{p_i-1}} \right)$$

$$\in \sum_{\substack{i=1\\ i\in K_s}}^{r} \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_i}{p_i-1}} \leq |w_i| \right\} + \sum_{\substack{j=r+1\\ j\in K_s}}^{m} \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_j}{p_j-1}} = |w_j| \right\}$$

oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ *und* $\overline{x}$ *löst das System nichtlinearer Gleichungen*

$$0 = \sum_{i=1}^{r} |w_i| \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} |w_j| \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}$$

für alle $l \in \{1, \ldots, n\}$.

Zum Abschluss dieses Unterkapitels untersuchen wir nun noch Aufgabenstellungen des Typs (SP), bei denen alle Abstandsfunktionen von der ℓ_∞-Norm induziert werden, also $d_k(x, a^k) = \|x - a^k\|_\infty$ für alle $k \in \{1, \ldots, m\}$ gilt. Dies entspricht dem Schema $1/\mathbb{R}^n/w_k \gtrless 0/\ell_\infty/\Sigma$ in der Klassifikation der Standortoptimierungsprobleme. Der folgende Satz ist dabei wiederum ein neues Resultat dieser Arbeit.

Satz 4.1.3 (*Notw. Optimalitätsbedingung für* $\mathbf{1/\mathbb{R}^n/w_k \gtrless 0/\ell_\infty/\Sigma}$)

Ist $\overline{x} \in \mathbb{R}^n$ *ein lokales Minimum des Problems (SP) mit* $d_k(x, a^k) = \|x - a^k\|_\infty$ *für alle* $k \in \{1, \ldots, m\}$, *dann ist entweder* $\overline{x} = a^k$ *für ein* $k \in \{1, \ldots, r\}$ *und es existieren* $x^{*_h} \in \mathbb{R}^n$ *für alle* $h \in \{1, \ldots, m\} \setminus \{k\}$, *welche*

$$\text{(1)} \quad \left\| \sum_{\substack{h=1\\ h\neq k}}^{m} |w_h| x^{*_h} \right\|_1 \leq |w_k|,$$

$$\text{(2)} \quad x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^i| = 0,$$

$$\text{(3)} \quad x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^j| = 0,$$

$$\text{(4)} \quad \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1$$

für alle $i \in \{1, \ldots, r\} \setminus \{k\}$, *für alle* $j \in \{r+1, \ldots, m\}$ *und für alle* $l \in \{1, \ldots, n\}$ *erfüllen oder es ist* $\overline{x} = a^k$ *für ein* $k \in \{r+1, \ldots, m\}$ *und es existieren* $x^{*_h} \in \mathbb{R}^n$ *für alle* $h \in \{1, \ldots, m\} \setminus \{k\}$, *welche*

$$\text{(1)} \quad \left\| \sum_{\substack{h=1\\ h\neq k}}^{m} |w_h| x^{*_h} \right\|_1 = |w_k|, \quad \left\| \sum_{\substack{h=1\\ h\neq k}}^{m} |w_h| x^{*_h} \right\|_\infty = |w_k|,$$

$$\text{(2)} \quad x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^i| = 0,$$

$$\text{(3)} \quad x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^j| = 0,$$

$$\text{(4)} \quad \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1$$

für alle $i \in \{1, \ldots, r\}$, *für alle* $j \in \{r+1, \ldots, m\} \setminus \{k\}$ *und für alle* $l \in \{1, \ldots, n\}$ *erfüllen oder es ist schließlich* $\overline{x} \notin \{a^1, \ldots, a^m\}$ *und es existieren* $x^{*_h} \in \mathbb{R}^n$ *für alle* $h \in \{1, \ldots, m\}$, *welche*

$$\begin{aligned}
&(1) \quad \sum_{h=1}^{m} |w_h| x^{*_h} = 0,\\
&(2) \quad x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^i| = 0,\\
&(3) \quad x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^j| = 0,\\
&(4) \quad \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1
\end{aligned}$$

für alle $i \in \{1, \ldots, r\}$, *für alle* $j \in \{r+1, \ldots, m\}$ *und für alle* $l \in \{1, \ldots, n\}$ *erfüllen.*

Beweis : Es sei $\overline{x}$ ein lokales Minimum von (SP) mit $d_k(x, a^k) = \|x - a^k\|_\infty$ für alle $k \in \{1, \ldots, m\}$. Aus Eigenschaft (L4) des Limiting-Subdifferentials aus Lemma 2.2.2 erhalten wir erneut die notwendige Optimalitätsbedingung, zur Umformung dieser nutzen wir auch hier die Summenregel für Limiting-Subdifferentiale aus Satz 2.2.8 und Eigenschaft (L1) aus Lemma 2.2.2. Damit ergibt sich

$$\begin{aligned}
0 &\in \partial_L f_w(\overline{x})\\
&= \partial_L \left(\sum_{i=1}^{r} |w_i| \|x - a^i\|_\infty + \sum_{j=r+1}^{m} |w_j| (-\|x - a^j\|_\infty) \right) (\overline{x})\\
&\subseteq \left(\sum_{i=1}^{r} \partial_L (|w_i| \|x - a^i\|_\infty) + \sum_{j=r+1}^{m} \partial_L (|w_j| (-\|x - a^j\|_\infty)) \right) (\overline{x})\\
&= \left(\sum_{i=1}^{r} |w_i| \partial_L (\|x - a^i\|_\infty) + \sum_{j=r+1}^{m} |w_j| \partial_L (-\|x - a^j\|_\infty) \right) (\overline{x}).
\end{aligned}$$

Wir setzen in diese Bedingung nun die in Satz 3.1.4 und Satz 3.2.3 berechneten Limiting-Subdifferentiale der positiven und negativen ℓ_∞-Norm ein und gehen dabei die drei Fälle des Satzes durch.

Zuerst sei $\overline{x} = a^k$ für ein $k \in \{1, \ldots, r\}$ einer der anziehenden Standorte, dann erhalten wir

$$\begin{aligned}
0 \in& \left(\sum_{i=1}^{r} |w_i| \partial_L (\|x - a^i\|_\infty) + \sum_{j=r+1}^{m} |w_j| \partial_L (-\|x - a^j\|_\infty) \right) (a^k)\\
=& \sum_{\substack{i=1\\ i \neq k}}^{r} |w_i| \Bigg\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\} \colon x_l^*(a_l^k - a_l^i) - |x_l^*| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^i| = 0,\\
&\sum_{l=1}^{n} |x_l^*| = 1 \Bigg\} + |w_k| \{x^* \in \mathbb{R}^n \mid \|x^*\|_1 \leq 1\} + \sum_{j=r+1}^{m} |w_j| \Bigg\{ x^* \in \mathbb{R}^n \mid \|x^*\|_\infty = 1,\\
&\sum_{l=1}^{n} |x_l^*| = 1, \forall\, l \in \{1, \ldots, n\} \colon x_l^*(a_l^k - a_l^j) + |x_l^*| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^j| = 0 \Bigg\}.
\end{aligned}$$

Das ist aber äquivalent zur Existenz eines $y^* \in \mathbb{R}^n$, für welches die Aussagen

$$y^* \in \sum_{\substack{i=1\\i\neq k}}^{r} |w_i| \Big\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1,\ldots,n\} \colon x_l^*(a_l^k - a_l^i) - |x_l^*| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^i| = 0,$$
$$\sum_{l=1}^{n} |x_l^*| = 1 \Big\} + \sum_{j=r+1}^{m} |w_j| \Big\{ x^* \in \mathbb{R}^n \mid \|x^*\|_\infty = 1, \sum_{l=1}^{n} |x_l^*| = 1,$$
$$\forall\, l \in \{1,\ldots,n\} \colon x_l^*(a_l^k - a_l^j) + |x_l^*| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^j| = 0 \Big\}$$

sowie

$$-y^* \in \{x^* \in \mathbb{R}^n \mid \|x^*\|_1 \leq |w_k|\}$$

gelten und dies ist äquivalent dazu, dass $x^{*_h} \in \mathbb{R}^n$ für alle $h \in \{1,\ldots,m\} \setminus \{k\}$ existieren, welche das folgende System von Bedingungen erfüllen:

(1) $\left\| \sum_{\substack{h=1\\h\neq k}}^{m} |w_h| x^{*_h} \right\|_1 \leq |w_k|,$

(2) $\forall\, i \in \{1,\ldots,r\} \setminus \{k\}\ \forall\, l \in \{1,\ldots,n\} \colon x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^i| = 0,$

(3) $\forall\, j \in \{r+1,\ldots,m\}\ \forall\, l \in \{1,\ldots,n\} \colon x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^j| = 0,$

(4) $\forall\, h \in \{1,\ldots,m\} \setminus \{k\} \colon \|x^{*_h}\|_1 = 1,$

(5) $\forall\, j \in \{r+1,\ldots,m\} \colon \|x^{*_j}\|_\infty = 1$

Ist hingegen einer der abstoßenden Standorte ein lokales Minimum der Problemstellung, also $\overline{x} = a^k$ für ein $k \in \{r+1,\ldots,m\}$, so liefert das Einsetzen der vorher berechneten Limiting-Subdifferentiale die notwendige Optimalitätsbedingung

$$0 \in \left(\sum_{i=1}^{r} |w_i| \partial_L(\|x - a^i\|_\infty) + \sum_{j=r+1}^{m} |w_j| \partial_L(-\|x - a^j\|_\infty) \right)(a^k)$$
$$= \sum_{i=1}^{r} |w_i| \Big\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1,\ldots,n\} \colon x_l^*(a_l^k - a_l^i) - |x_l^*| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^i| = 0,$$
$$\sum_{l=1}^{n} |x_l^*| = 1 \Big\} + |w_k| \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_1 = 1, \|x^*\|_\infty = 1 \right\} + \sum_{\substack{j=r+1\\j\neq k}}^{m} |w_j| \Big\{ x^* \in \mathbb{R}^n \mid$$
$$\sum_{l=1}^{n} |x_l^*| = 1, \|x^*\|_\infty = 1, \forall\, l \in \{1,\ldots,n\} \colon x_l^*(a_l^k - a_l^j) + |x_l^*| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^j| = 0 \Big\}.$$

Mit den gleichen Schritten wie zuvor ist dies äquivalent zur Existenz eines $y^* \in \mathbb{R}^n$, welches

$$y^* \in \sum_{i=1}^{r} |w_i| \Big\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1,\ldots,n\} \colon x_l^*(a_l^k - a_l^i) - |x_l^*| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^i| = 0,$$
$$\sum_{l=1}^{n} |x_l^*| = 1 \Big\} + \sum_{\substack{j=r+1\\j\neq k}}^{m} |w_j| \Big\{ x^* \in \mathbb{R}^n \mid \|x^*\|_\infty = 1, \sum_{l=1}^{n} |x_l^*| = 1,$$
$$\forall\, l \in \{1,\ldots,n\} \colon x_l^*(a_l^k - a_l^j) + |x_l^*| \max_{l\in\{1,\ldots,n\}} |a_l^k - a_l^j| = 0 \Big\}$$

und

$$-y^* \in \{x^* \in \mathbb{R}^n \mid \|x^*\|_1 = |w_k|, \|x^*\|_\infty = |w_k|\}$$

erfüllt bzw. äquivalent zur Existenz von $x^{*_h} \in \mathbb{R}^n$ für alle $h \in \{1, \ldots, m\} \setminus \{k\}$, die dem folgenden System von Bedingungen genügen:

(1) $\left\| \sum_{\substack{h=1 \\ h \neq k}}^{m} |w_h| x^{*_h} \right\|_1 = |w_k|, \quad \left\| \sum_{\substack{h=1 \\ h \neq k}}^{m} |w_h| x^{*_h} \right\|_\infty = |w_k|,$

(2) $\forall\, i \in \{1, \ldots, r\}\ \forall\, l \in \{1, \ldots, n\}\colon x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^i| = 0,$

(3) $\forall\, j \in \{r+1, \ldots, m\} \setminus \{k\}\ \forall\, l \in \{1, \ldots, n\}\colon x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^j| = 0,$

(4) $\forall\, h \in \{1, \ldots, m\} \setminus \{k\}\colon \|x^{*_h}\|_1 = 1,$

(5) $\forall\, j \in \{r+1, \ldots, m\} \setminus \{k\}\colon \|x^{*_j}\|_\infty = 1$

Sei zuletzt $\overline{x} \notin \{a^1, \ldots, a^m\}$, dann erhalten wir durch das Einsetzen unserer Ergebnisse aus den Sätzen 3.1.4 und 3.2.3

$$\begin{aligned}
0 \in &\left(\sum_{i=1}^{r} |w_i| \partial_L(\|x - a^i\|_\infty) + \sum_{j=r+1}^{m} |w_j| \partial_L(-\|x - a^j\|_\infty) \right)(\overline{x}) \\
= &\sum_{i=1}^{r} |w_i| \Big\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\}\colon x_l^*(\overline{x}_l - a_l^i) - |x_l^*| \max_{l \in \{1,\ldots,n\}} |\overline{x}_l - a_l^i| = 0, \\
&\sum_{l=1}^{n} |x_l^*| = 1 \Big\} + \sum_{j=r+1}^{m} |w_j| \Big\{ x^* \in \mathbb{R}^n \mid \|x^*\|_\infty = 1, |x_l^*| = 1, \\
&\sum_{l=1}^{n} \forall\, l \in \{1, \ldots, n\}\colon x_l^*(\overline{x}_l - a_l^j) + |x_l^*| \max_{l \in \{1,\ldots,n\}} |\overline{x}_l - a_l^j| = 0 \Big\}.
\end{aligned}$$

Dies ist nun direkt äquivalent zur Existenz von $x^{*_h} \in \mathbb{R}^n$ für alle $h \in \{1, \ldots, m\}$, so dass das System

(1) $\sum_{h=1}^{m} |w_h| x^{*_h} = 0,$

(2) $\forall\, i \in \{1, \ldots, r\}\ \forall\, l \in \{1, \ldots, n\}\colon x_l^{*_i}(\overline{x}_l - a_l^i) - |x_l^{*_i}| \max_{l \in \{1,\ldots,n\}} |\overline{x}_l - a_l^i| = 0,$

(3) $\forall\, j \in \{r+1, \ldots, m\}\ \forall\, l \in \{1, \ldots, n\}\colon x_l^{*_j}(\overline{x}_l - a_l^j) + |x_l^{*_j}| \max_{l \in \{1,\ldots,n\}} |\overline{x}_l - a_l^j| = 0,$

(4) $\forall\, h \in \{1, \ldots, m\}\colon \|x^{*_h}\|_1 = 1$ und

(5) $\forall\, j \in \{r+1, \ldots, m\}\colon \|x^{*_j}\|_\infty = 1$

erfüllt ist. ■

Bemerkung 4.1.3

Alle Ergebnisse dieses Unterkapitels führen im Spezialfall ausschließlich anziehender existierender Standorte, wenn also $r = m$ gilt, auf die bekannten Ergebnisse der konvexen Standortoptimierung zurück: Die notwendige Optimalitätsbedingung in Satz 4.1.1 liefert dann eine Aussage, welche äquivalent zu Korollar 2.2 bei Hamacher (1995) ist, die Fallunterscheidung in Korollar 4.1.1 geht genau in *die Testbedingung und die Rekursionsvorschrift des Weiszfeld-Algorithmus über (vgl. Satz 2.8 und Algorithmus 2.5 bei Hamacher (1995)) und Satz 4.1.2 reduziert sich zu Testbedingung und Rekursionsvorschrift des hyperbolischen Approximationsalgorithmus (vgl. Satz 2.9 und Algorithmus 2.6 bei Hamacher (1995)). Die von Hamacher (1995) für den planaren Fall $n = 2$ vorgestellten Resultate lassen sich dabei (mit Ausnahme der Aussage über den Zusammenhang von ℓ_1-Norm und ℓ_∞-Norm) ohne Einschränkung auf beliebige Dimensionen verallgemeinern.*

Unsere notwendigen Optimalitätsbedingungen stellen somit richtige und natürliche Fortsetzungen der klassischen konvexen Theorie dar.

4.2 Notwendige Optimalitätsbedingungen für Medianprobleme mit gemischten Minkowski-Metriken

Nach den notwendigen Optimalitätsbedingungen für Medianprobleme mit gleichen Abstandsfunktionen wollen wir eine solche notwendige Optimalitätsbedingung jetzt auch für das Problem (SP) herleiten, wenn die zugeordneten Metriken zu den existierenden Standorten a^k für alle $k \in \{1, \ldots, m\}$ von unterschiedlichen ℓ_{p_k}-Normen (mit $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$) induziert werden. Eine derartige Problemstellung hat in der Klassifikation der Standortoptimierungsprobleme das Schema $1/\mathbb{R}^n/w_k \gtrless 0/\ell_p^m/\Sigma$.

Es kann bei dieser Aufgabenstellung nun der Fall auftreten, dass demselben existierenden Standort mehrere verschiedene Abstandsfunktionen zugeordnet sind, beispielsweise kann im Kontext der Stadtplanung eine Fabrik aufgrund der zur Verfügung stehenden Arbeitsplätze sowohl anziehender Standort mit ℓ_1-Norm sein als auch gleichzeitig abstoßender Standort mit ℓ_2-Norm, da von dieser Fabrik Lärmbelästigung ausgeht oder Abgase verbreitet werden. In unserer Zielfunktion aus Sektion 1.2.2 wird dies durch zwei oder mehr gleiche existierende Standorte mit jeweils einer dieser Metriken als zugeordneter Abstandsfunktion modelliert.

Die existierenden Standorte seien dabei im Folgenden ohne Beschränkung der Allgemeinheit nach ihren zugeordneten Abstandsfunktionen geordnet, so dass also

- $d_k(x, a^k) = \|x - a^k\|_1$ für alle $k \in \{1, \ldots, m\}$ mit $1 \leq k \leq r_1$ oder $r+1 \leq k \leq r_3$,
- $d_k(x, a^k) = \|x - a^k\|_{p_k}$ mit $p_k \in \mathbb{R}$, $p_k > 1$ für alle $k \in \{1, \ldots, m\}$ mit $r_1 + 1 \leq k \leq r_2$ oder $r_3 + 1 \leq k \leq r_4$ sowie
- $d_k(x, a^k) = \|x - a^k\|_\infty$ für alle $k \in \{1, \ldots, m\}$ mit $r_2 + 1 \leq k \leq r$ oder $r_4 + 1 \leq k \leq m$

gilt, wobei $0 \leq r_1 \leq r_2 \leq r \leq r_3 \leq r_4 \leq m$ sein soll.

Für die Notation nutzen wir zudem Indexmengen K_s mit $s \in \{1, \ldots, u\}$, in denen die Indizes gleicher Standorte zusammengefasst werden (vgl. Bemerkung 4.1.2 (ii) im vorherigen Unterkapitel) und welche dafür die Eigenschaften

(a) $\bigcup\limits_{s=1}^{u} K_s = \{1, \ldots, m\}$,

(b) $a^g = a^h$ für alle $g, h \in K_s$ und für alle $s \in \{1, \ldots, u\}$ sowie

(c) $a^g \neq a^h$ für alle $g \in K_s, h \in K_t$ mit $s, t \in \{1, \ldots, u\}$ und $s \neq t$

erfüllen müssen. Damit können wir nun das wichtigste Resultat unserer Arbeit formulieren.

Satz 4.2.1 (*Notw. Optimalitätsbedingung für* $\mathbf{1}/\mathbb{R}^n/w_k \gtrless 0/\ell_p^m/\Sigma$)

Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (SP), die existierenden Standorte in Indexmengen K_s mit $s \in \{1, \ldots, u\}$ aufgeteilt, welche (a) - (c) genügen, sowie $d_k(x, a^k) = \|x - a^k\|_{p_k}$ mit $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ für alle $k \in \{1, \ldots, m\}$, dann ist entweder $\overline{x} = a^g$ für ein $g \in K_s$ mit $s \in \{1, \ldots, u\}$ und es existieren $x^{*_h} \in \mathbb{R}^n$ für alle $h \in \{1, \ldots, r_1, r_2 + 1, \ldots, r_3, r_4 + 1, \ldots, m\} \setminus K_s$, welche

(1)
$$\begin{aligned}
&\sum_{\substack{i=1\\ i\notin K_s}}^{r_1} |w_i| x^{*_i} + \sum_{\substack{i=r_2+1\\ i\notin K_s}}^{r} |w_i| x^{*_i} + \sum_{\substack{j=r+1\\ j\notin K_s}}^{r_3} |w_j| x^{*_j} + \sum_{\substack{j=r_4+1\\ j\notin K_s}}^{m} |w_j| x^{*_j} \\
&+ \sum_{\substack{i=r_1+1\\ i\notin K_s}}^{r_2} |w_i| \left(\frac{\operatorname{sgn}(a_1^g - a_1^i)|a_1^g - a_1^i|^{p_i-1}}{\|a^g - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a_n^g - a_n^i)|a_n^g - a_n^i|^{p_i-1}}{\|a^g - a^i\|_{p_i}^{p_i-1}} \right) \\
&- \sum_{\substack{j=r_3+1\\ j\notin K_s}}^{r_4} |w_j| \left(\frac{\operatorname{sgn}(a_1^g - a_1^j)|a_1^g - a_1^j|^{p_j-1}}{\|a^g - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(a_n^g - a_n^j)|a_n^g - a_n^j|^{p_j-1}}{\|a^g - a^j\|_{p_j}^{p_j-1}} \right) \\
\in &\sum_{\substack{i=1\\ i\in K_s}}^{r_1} \{x^* \in \mathbb{R}^n \mid \|x^*\|_\infty \leq |w_i|\} + \sum_{\substack{i=r_1+1\\ i\in K_s}}^{r_2} \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_i}{p_i-1}} \leq |w_i| \right\} \\
&+ \sum_{\substack{i=r_2+1\\ i\in K_s}}^{r} \{x^* \in \mathbb{R}^n \mid \|x^*\|_1 \leq |w_i|\} + \sum_{\substack{j=r_3+1\\ j\in K_s}}^{r_4} \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_j}{p_j-1}} = |w_j| \right\} \\
&+ \sum_{\substack{j=r+1\\ j\in K_s}}^{r_3} \{-w_j, w_j\}^n + \sum_{\substack{j=r_4+1\\ j\in K_s}}^{m} \{x^* \in \mathbb{R}^n \mid \|x^*\|_1 = |w_j|, \|x^*\|_\infty = |w_j|\},
\end{aligned}$$

(2) $x_l^{*_{\widetilde{i}}} \in \begin{cases} \{1\} & \text{für } a_l^g - a_l^{\widetilde{i}} > 0 \\ \{-1\} & \text{für } a_l^g - a_l^{\widetilde{i}} < 0 \\ [-1, 1] & \text{für } a_l^g - a_l^{\widetilde{i}} = 0 \end{cases}$,

(3) $x_l^{*_{\widetilde{j}}} \in \begin{cases} \{-1\} & \text{für } a_l^g - a_l^{\widetilde{j}} > 0 \\ \{1\} & \text{für } a_l^g - a_l^{\widetilde{j}} < 0 \\ \{-1, 1\} & \text{für } a_l^g - a_l^{\widetilde{j}} = 0 \end{cases}$,

(4) $x_l^{*_{\widehat{i}}}(a_l^g - a_l^{\widehat{i}}) - |x_l^{*_{\widehat{i}}}| \max\limits_{l \in \{1,\ldots,n\}} |a_l^g - a_l^{\widehat{i}}| = 0,$

(5) $x_l^{*_{\widehat{j}}}(a_l^g - a_l^{\widehat{j}}) + |x_l^{*_{\widehat{j}}}| \max\limits_{l \in \{1,\ldots,n\}} |a_l^g - a_l^{\widehat{j}}| = 0,$

(6) $\|x^{*_{\widehat{i}}}\|_1 = 1$ *sowie*

(7) $\|x^{*_{\widehat{j}}}\|_1 = 1$ *und* $\|x^{*_{\widehat{j}}}\|_\infty = 1$

für alle $l \in \{1, \ldots, n\}$, für alle $\widetilde{i} \in \{1, \ldots, r_1\} \setminus K_s$, für alle $\widehat{i} \in \{r_2+1, \ldots, r\} \setminus K_s$, *für alle* $\widetilde{j} \in \{r+1, \ldots, r_3\} \setminus K_s$ *und für alle* $\widehat{j} \in \{r_4+1, \ldots, m\} \setminus K_s$ *erfüllen oder es ist* $\overline{x} \notin \{a^1, \ldots, a^m\}$ *und für alle* $h \in \{1, \ldots, r_1, r_2+1, \ldots, r_3, r_4+1, \ldots, m\}$ *existieren Vektoren* $x^{*h} \in \mathbb{R}^n$, *welche*

$$\text{(1)} \quad \sum_{j=r_3+1}^{r_4} |w_j| \left(\frac{\operatorname{sgn}(\overline{x}_1 - a_1^j)|\overline{x}_1 - a_1^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(\overline{x}_n - a_n^j)|\overline{x}_n - a_n^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}} \right)$$
$$- \sum_{i=r_1+1}^{r_2} |w_i| \left(\frac{\operatorname{sgn}(\overline{x}_1 - a_1^i)|\overline{x}_1 - a_1^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(\overline{x}_n - a_n^i)|\overline{x}_n - a_n^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} \right)$$
$$= \sum_{i=1}^{r_1} |w_i| x^{*i} + \sum_{i=r_2+1}^{r} |w_i| x^{*i} + \sum_{j=r+1}^{r_3} |w_j| x^{*j} + \sum_{j=r_4+1}^{m} |w_j| x^{*j},$$

$$\text{(2)} \; x_l^{*\widetilde{i}} \in \begin{cases} \{1\} & \text{für } \overline{x}_l - a_l^{\widetilde{i}} > 0 \\ \{-1\} & \text{für } \overline{x}_l - a_l^{\widetilde{i}} < 0 \\ [-1,1] & \text{für } \overline{x}_l - a_l^{\widetilde{i}} = 0 \end{cases},$$

$$\text{(3)} \; x_l^{*\widetilde{j}} \in \begin{cases} \{-1\} & \text{für } \overline{x}_l - a_l^{\widetilde{j}} > 0 \\ \{1\} & \text{für } \overline{x}_l - a_l^{\widetilde{j}} < 0 \\ \{-1,1\} & \text{für } \overline{x}_l - a_l^{\widetilde{j}} = 0 \end{cases},$$

$$\text{(4)} \; x_l^{*\widehat{i}}(\overline{x}_l - a_l^{\widehat{i}}) - |x_l^{*\widehat{i}}| \max_{l \in \{1,\ldots,n\}} |\overline{x}_l - a_l^{\widehat{i}}| = 0,$$

$$\text{(5)} \; x_l^{*\widehat{j}}(\overline{x}_l - a_l^{\widehat{j}}) + |x_l^{*\widehat{j}}| \max_{l \in \{1,\ldots,n\}} |\overline{x}_l - a_l^{\widehat{j}}| = 0,$$

$$\text{(6)} \; \|x^{*\widehat{i}}\|_1 = 1 \text{ sowie}$$

$$\text{(7)} \; \|x^{*\widehat{j}}\|_1 = 1 \text{ und } \|x^{*\widehat{j}}\|_\infty = 1$$

für alle $l \in \{1, \ldots, n\}$, *für alle* $\widetilde{i} \in \{1, \ldots, r_1\}$, *für alle* $\widehat{i} \in \{r_2+1, \ldots, r\}$, *für alle* $\widetilde{j} \in \{r+1, \ldots, r_3\}$ *und für alle* $\widehat{j} \in \{r_4+1, \ldots, m\}$ *erfüllen.*

Beweis: Es sei $\overline{x}$ ein lokales Minimum von (SP) mit $d_k(x, a^k) = \|x - a^k\|_{p_k}$, $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ für alle $k \in \{1, \ldots, m\}$. Eigenschaft (L4) aus Lemma 2.2.2 gibt uns erneut die allgemein notwendige Optimalitätsbedingung vor, welche wir mit der Summenregel für Limiting-Subdifferentiale aus Satz 2.2.8 und der Homogenitäts-Eigenschaft (L1) aus Lemma 2.2.2 entsprechend umformen. Wir erhalten dann die notwendige Optimalitätsbedingung

$$\begin{aligned}
0 &\in \partial_L f_w(\overline{x}) \\
&= \partial_L \left(\sum_{i=1}^{r} |w_i| \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} |w_j| (-\|x - a^j\|_{p_j}) \right) (\overline{x}) \\
&\subseteq \left(\sum_{i=1}^{r} \partial_L(|w_i| \|x - a^i\|_{p_i}) + \sum_{j=r+1}^{m} \partial_L(|w_j| (-\|x - a^j\|_{p_j})) \right) (\overline{x}) \\
&= \left(\sum_{i=1}^{r} |w_i| \partial_L(\|x - a^i\|_{p_i}) + \sum_{j=r+1}^{m} |w_j| \partial_L(-\|x - a^j\|_{p_j}) \right) (\overline{x})
\end{aligned}$$

$$= \left(\sum_{i=1}^{r_1} |w_i| \partial_L(\|x - a^i\|_1) + \sum_{i=r_1+1}^{r_2} |w_i| \partial_L(\|x - a^i\|_{p_i}) \right.$$
$$+ \sum_{i=r_2+1}^{r} |w_i| \partial_L(\|x - a^i\|_\infty) + \sum_{j=r+1}^{r_3} |w_j| \partial_L(-\|x - a^j\|_1)$$
$$\left. + \sum_{j=r_3+1}^{r_4} |w_j| \partial_L(-\|x - a^j\|_{p_j}) + \sum_{j=r_4+1}^{m} |w_j| \partial_L(-\|x - a^j\|_\infty) \right) (\overline{x}).$$

Wir setzen in diese Bedingung nun unsere gesamten Berechnungen aus den Unterkapiteln 3.1 und 3.2 ein. Beginnen wir zuerst mit dem Fall $\overline{x} = a^g$ für ein $g \in K_s$ mit $s \in \{1, \ldots, u\}$, so liefert das Einsetzen

$$0 \in \left(\sum_{i=1}^{r_1} |w_i| \partial_L(\|x - a^i\|_1) + \sum_{i=r_1+1}^{r_2} |w_i| \partial_L(\|x - a^i\|_{p_i}) \right.$$
$$+ \sum_{i=r_2+1}^{r} |w_i| \partial_L(\|x - a^i\|_\infty) + \sum_{j=r+1}^{r_3} |w_j| \partial_L(-\|x - a^j\|_1)$$
$$\left. + \sum_{j=r_3+1}^{r_4} |w_j| \partial_L(-\|x - a^j\|_{p_j}) + \sum_{j=r_4+1}^{m} |w_j| \partial_L(-\|x - a^j\|_\infty) \right) (a^g)$$
$$= \sum_{\substack{i=1\\ i \notin K_s}}^{r_1} |w_i| \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\} \colon x_l^* \in \begin{cases} \{1\} & \text{für } a_l^g - a_l^i > 0 \\ \{-1\} & \text{für } a_l^g - a_l^i < 0 \\ [-1, 1] & \text{für } a_l^g - a_l^i = 0 \end{cases} \right\}$$
$$+ \sum_{\substack{i=1\\ i \in K_s}}^{r_1} |w_i| \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\} \colon x_l^* \in [-1, 1] \right\}$$
$$+ \sum_{\substack{i=r_1+1\\ i \notin K_s}}^{r_2} |w_i| \left\{ \left(\frac{\operatorname{sgn}(a_1^g - a_1^i)|a_1^g - a_1^i|^{p_i-1}}{\|a^g - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a_n^g - a_n^i)|a_n^g - a_n^i|^{p_i-1}}{\|a^g - a^i\|_{p_i}^{p_i-1}} \right) \right\}$$
$$+ \sum_{\substack{i=r_1+1\\ i \in K_s}}^{r_2} |w_i| \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_i}{p_i-1}} \leq 1 \right\}$$
$$+ \sum_{\substack{i=r_2+1\\ i \notin K_s}}^{r} |w_i| \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\} \colon x_l^*(a_l^g - a_l^i) - |x_l^*| \max_{l \in \{1, \ldots, n\}} |a_l^g - a_l^i| = 0, \right.$$
$$\left. \sum_{l=1}^{n} |x_l^*| = 1 \right\} + \sum_{\substack{i=r_2+1\\ i \in K_s}}^{r} |w_i| \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_1 \leq 1 \right\}$$
$$+ \sum_{\substack{j=r+1\\ j \notin K_s}}^{r_3} |w_j| \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\} \colon x_l^* \in \begin{cases} \{-1\} & \text{für } a_l^g - a_l^j > 0 \\ \{1\} & \text{für } a_l^g - a_l^j < 0 \\ \{-1, 1\} & \text{für } a_l^g - a_l^j = 0 \end{cases} \right\}$$
$$+ \sum_{\substack{j=r+1\\ j \in K_s}}^{r_3} |w_j| \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1, \ldots, n\} \colon x_l^* \in \{-1, 1\} \right\}$$

$$+\sum_{\substack{j=r_3+1\\ j\notin K_s}}^{r_4}|w_j|\left\{\left(-\frac{\operatorname{sgn}(a_1^g-a_1^j)|a_1^g-a_1^j|^{p_j-1}}{\|a^g-a^j\|_{p_j}^{p_j-1}},\ldots,-\frac{\operatorname{sgn}(a_n^g-a_n^j)|a_n^g-a_n^j|^{p_j-1}}{\|a^g-a^j\|_{p_j}^{p_j-1}}\right)\right\}$$
$$+\sum_{\substack{j=r_3+1\\ j\in K_s}}^{r_4}|w_j|\left\{x^*\in\mathbb{R}^n \mid \|x^*\|_{\frac{p_j}{p_j-1}}=1\right\}$$
$$+\sum_{\substack{j=r_4+1\\ j\notin K_s}}^{m}|w_j|\left\{x^*\in\mathbb{R}^n \mid \forall\, l\in\{1,\ldots,n\}\colon x_l^*(a_l^g-a_l^j)+|x_l^*|\max_{l\in\{1,\ldots,n\}}|a_l^g-a_l^j|=0,\right.$$
$$\left.\sum_{l=1}^n|x_l^*|=1,\|x^*\|_\infty=1\right\}+\sum_{\substack{j=r_4+1\\ j\in K_s}}^{m}|w_j|\left\{x^*\in\mathbb{R}^n \mid \|x^*\|_1=1,\|x^*\|_\infty=1\right\}.$$

Das ist äquivalent zur Existenz eines $y^*\in\mathbb{R}^n$ mit

$$y^*\in\sum_{\substack{i=1\\ i\notin K_s}}^{r_1}|w_i|\left\{x^*\in\mathbb{R}^n \mid \forall\, l\in\{1,\ldots,n\}\colon x_l^*\in\begin{Bmatrix}\{1\} & \text{für } a_l^g-a_l^i>0\\ \{-1\} & \text{für } a_l^g-a_l^i<0\\ [-1,1] & \text{für } a_l^g-a_l^i=0\end{Bmatrix}\right\}$$
$$+\sum_{\substack{i=r_1+1\\ i\notin K_s}}^{r_2}|w_i|\left\{\left(\frac{\operatorname{sgn}(a_1^g-a_1^i)|a_1^g-a_1^i|^{p_i-1}}{\|a^g-a^i\|_{p_i}^{p_i-1}},\ldots,\frac{\operatorname{sgn}(a_n^g-a_n^i)|a_n^g-a_n^i|^{p_i-1}}{\|a^g-a^i\|_{p_i}^{p_i-1}}\right)\right\}$$
$$+\sum_{\substack{i=r_2+1\\ i\notin K_s}}^{r}|w_i|\left\{x^*\in\mathbb{R}^n \mid \forall\, l\in\{1,\ldots,n\}\colon x_l^*(a_l^g-a_l^i)-|x_l^*|\max_{l\in\{1,\ldots,n\}}|a_l^g-a_l^i|=0,\right.$$
$$\left.\sum_{l=1}^n|x_l^*|=1\right\}$$
$$+\sum_{\substack{j=r+1\\ j\notin K_s}}^{r_3}|w_j|\left\{x^*\in\mathbb{R}^n \mid \forall\, l\in\{1,\ldots,n\}\colon x_l^*\in\begin{Bmatrix}\{-1\} & \text{für } a_l^g-a_l^j>0\\ \{1\} & \text{für } a_l^g-a_l^j<0\\ \{-1,1\} & \text{für } a_l^g-a_l^j=0\end{Bmatrix}\right\}$$
$$-\sum_{\substack{j=r_3+1\\ j\notin K_s}}^{r_4}|w_j|\left\{\left(\frac{\operatorname{sgn}(a_1^g-a_1^j)|a_1^g-a_1^j|^{p_j-1}}{\|a^g-a^j\|_{p_j}^{p_j-1}},\ldots,\frac{\operatorname{sgn}(a_n^g-a_n^j)|a_n^g-a_n^j|^{p_j-1}}{\|a^g-a^j\|_{p_j}^{p_j-1}}\right)\right\}$$
$$+\sum_{\substack{j=r_4+1\\ j\notin K_s}}^{m}|w_j|\left\{x^*\in\mathbb{R}^n \mid \forall\, l\in\{1,\ldots,n\}\colon x_l^*(a_l^g-a_l^j)+|x_l^*|\max_{l\in\{1,\ldots,n\}}|a_l^g-a_l^j|=0,\right.$$
$$\left.\sum_{l=1}^n|x_l^*|=1,\|x^*\|_\infty=1\right\};$$
$$-y^*\in\sum_{\substack{i=1\\ i\in K_s}}^{r_1}\{x^*\in\mathbb{R}^n \mid \|x^*\|_\infty\le|w_i|\}+\sum_{\substack{i=r_1+1\\ i\in K_s}}^{r_2}\left\{x^*\in\mathbb{R}^n \mid \|x^*\|_{\frac{p_i}{p_i-1}}\le|w_i|\right\}$$
$$+\sum_{\substack{i=r_2+1\\ i\in K_s}}^{r}\{x^*\in\mathbb{R}^n \mid \|x^*\|_1\le|w_i|\}+\sum_{\substack{j=r_3+1\\ j\in K_s}}^{r_4}\left\{x^*\in\mathbb{R}^n \mid \|x^*\|_{\frac{p_j}{p_j-1}}=|w_j|\right\}$$
$$+\sum_{\substack{j=r+1\\ j\in K_s}}^{r_3}\{-w_j,w_j\}^n+\sum_{\substack{j=r_4+1\\ j\in K_s}}^{m}\{x^*\in\mathbb{R}^n \mid \|x^*\|_1=|w_j|,\|x^*\|_\infty=|w_j|\}$$

bzw. äquivalent dazu, dass für alle $h \in \{1, \ldots, r_1, r_2+1, \ldots, r_3, r_4+1, \ldots, m\} \setminus K_s$ Vektoren $x^{*_h} \in \mathbb{R}^n$ existieren, welche das System

$$
\begin{aligned}
(1) \quad & \sum_{\substack{i=1\\ i\notin K_s}}^{r_1} |w_i| x^{*_i} + \sum_{\substack{i=r_2+1\\ i\notin K_s}}^{r} |w_i| x^{*_i} + \sum_{\substack{j=r+1\\ j\notin K_s}}^{r_3} |w_j| x^{*_j} + \sum_{\substack{j=r_4+1\\ j\notin K_s}}^{m} |w_j| x^{*_j} \\
& + \sum_{\substack{i=r_1+1\\ i\notin K_s}}^{r_2} |w_i| \left(\frac{\operatorname{sgn}(a_1^g - a_1^i)|a_1^g - a_1^i|^{p_i-1}}{\|a^g - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a_n^g - a_n^i)|a_n^g - a_n^i|^{p_i-1}}{\|a^g - a^i\|_{p_i}^{p_i-1}} \right) \\
& - \sum_{\substack{j=r_3+1\\ j\notin K_s}}^{r_4} |w_j| \left(\frac{\operatorname{sgn}(a_1^g - a_1^j)|a_1^g - a_1^j|^{p_j-1}}{\|a^g - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(a_n^g - a_n^j)|a_n^g - a_n^j|^{p_j-1}}{\|a^g - a^j\|_{p_j}^{p_j-1}} \right) \\
& \in \sum_{\substack{i=1\\ i\in K_s}}^{r_1} \{x^* \in \mathbb{R}^n \mid \|x^*\|_\infty \le |w_i|\} + \sum_{\substack{i=r_1+1\\ i\in K_s}}^{r_2} \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_i}{p_i-1}} \le |w_i| \right\} \\
& + \sum_{\substack{i=r_2+1\\ i\in K_s}}^{r} \{x^* \in \mathbb{R}^n \mid \|x^*\|_1 \le |w_i|\} + \sum_{\substack{j=r_3+1\\ j\in K_s}}^{r_4} \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_j}{p_j-1}} = |w_j| \right\} \\
& + \sum_{\substack{j=r+1\\ j\in K_s}}^{r_3} \{-w_j, w_j\}^n + \sum_{\substack{j=r_4+1\\ j\in K_s}}^{m} \{x^* \in \mathbb{R}^n \mid \|x^*\|_1 = |w_j|, \|x^*\|_\infty = |w_j|\},
\end{aligned}
$$

(2) $\forall\, i \in \{1, \ldots, r_1\} \setminus K_s\ \forall\, l \in \{1, \ldots, n\}\colon x_l^{*_i} \in \begin{cases} \{1\} & \text{für } a_l^g - a_l^i > 0 \\ \{-1\} & \text{für } a_l^g - a_l^i < 0 \\ [-1, 1] & \text{für } a_l^g - a_l^i = 0 \end{cases},$

(3) $\forall\, j \in \{r+1, \ldots, r_3\} \setminus K_s\ \forall\, l \in \{1, \ldots, n\}\colon x_l^{*_j} \in \begin{cases} \{-1\} & \text{für } a_l^g - a_l^j > 0 \\ \{1\} & \text{für } a_l^g - a_l^j < 0 \\ \{-1, 1\} & \text{für } a_l^g - a_l^j = 0 \end{cases},$

(4) $\forall\, i \in \{r_2+1, \ldots, r\} \setminus K_s\ \forall\, l \in \{1, \ldots, n\}\colon x_l^{*_i}(a_l^g - a_l^i) - |x_l^{*_i}| \max_{l \in \{1,\ldots,n\}} |a_l^g - a_l^i| = 0,$

(5) $\forall\, j \in \{r_4+1, \ldots, m\} \setminus K_s\ \forall\, l \in \{1, \ldots, n\}\colon x_l^{*_j}(a_l^g - a_l^j) + |x_l^{*_j}| \max_{l \in \{1,\ldots,n\}} |a_l^g - a_l^j| = 0,$

(6) $\forall\, i \in \{r_2+1, \ldots, r\} \setminus K_s\colon \|x^{*_i}\|_1 = 1$ sowie

(7) $\forall\, j \in \{r_4+1, \ldots, m\} \setminus K_s\colon \|x^{*_j}\|_1 = 1$ und $\|x^{*_j}\|_\infty = 1$

erfüllen.

Ist hingegen $\overline{x} \notin \{a^1, \ldots, a^m\}$, dann liefert das Einsetzen der in den Unterkapiteln 3.1 und 3.2 berechneten Limiting-Subdifferentiale in die notwendige Optimalitätsbedingung

$$
\begin{aligned}
0 \in \Bigg(& \sum_{i=1}^{r_1} |w_i| \partial_L(\|x - a^i\|_1) + \sum_{i=r_1+1}^{r_2} |w_i| \partial_L(\|x - a^i\|_{p_i}) \\
& + \sum_{i=r_2+1}^{r} |w_i| \partial_L(\|x - a^i\|_\infty) + \sum_{j=r+1}^{r_3} |w_j| \partial_L(-\|x - a^j\|_1) \\
& + \sum_{j=r_3+1}^{r_4} |w_j| \partial_L(-\|x - a^j\|_{p_j}) + \sum_{j=r_4+1}^{m} |w_j| \partial_L(-\|x - a^j\|_\infty) \Bigg) (\overline{x})
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{i=1}^{r_1} |w_i| \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1,\dots,n\} \colon x_l^* \in \begin{cases} \{1\} & \text{für } \overline{x}_l - a_l^i > 0 \\ \{-1\} & \text{für } \overline{x}_l - a_l^i < 0 \\ [-1,1] & \text{für } \overline{x}_l - a_l^i = 0 \end{cases} \right\} \\
&\quad + \sum_{i=r_1+1}^{r_2} |w_i| \left\{ \left(\frac{\operatorname{sgn}(\overline{x}_1 - a_1^i)|\overline{x}_1 - a_1^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}}, \dots, \frac{\operatorname{sgn}(\overline{x}_n - a_n^i)|\overline{x}_n - a_n^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} \right) \right\} \\
&\quad + \sum_{i=r_2+1}^{r} |w_i| \Bigg\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1,\dots,n\} \colon x_l^*(\overline{x}_l - a_l^i) - |x_l^*| \max_{l \in \{1,\dots,n\}} |\overline{x}_l - a_l^i| = 0, \\
&\qquad \sum_{l=1}^{n} |x_l^*| = 1 \Bigg\} \\
&\quad + \sum_{j=r+1}^{r_3} |w_j| \left\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1,\dots,n\} \colon x_l^* \in \begin{cases} \{-1\} & \text{für } \overline{x}_l - a_l^j > 0 \\ \{1\} & \text{für } \overline{x}_l - a_l^j < 0 \\ \{-1,1\} & \text{für } \overline{x}_l - a_l^j = 0 \end{cases} \right\} \\
&\quad + \sum_{j=r_3+1}^{r_4} |w_j| \left\{ \left(-\frac{\operatorname{sgn}(\overline{x}_1 - a_1^j)|\overline{x}_1 - a_1^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}, \dots, -\frac{\operatorname{sgn}(\overline{x}_n - a_n^j)|\overline{x}_n - a_n^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}} \right) \right\} \\
&\quad + \sum_{j=r_4+1}^{m} |w_j| \Bigg\{ x^* \in \mathbb{R}^n \mid \forall\, l \in \{1,\dots,n\} \colon x_l^*(\overline{x}_l - a_l^j) + |x_l^*| \max_{l \in \{1,\dots,n\}} |\overline{x}_l - a_l^j| = 0, \\
&\qquad \sum_{l=1}^{n} |x_l^*| = 1, \|x^*\|_\infty = 1 \Bigg\}.
\end{aligned}
$$

Dies ist nun aber schon direkt äquivalent zur Existenz von Vektoren $x^{*_h} \in \mathbb{R}^n$ für alle $h \in \{1,\dots,r_1, r_2+1,\dots,r_3, r_4+1,\dots,m\}$, welche das System

$$
\begin{aligned}
(1) \qquad & \sum_{j=r_3+1}^{r_4} |w_j| \left(\frac{\operatorname{sgn}(\overline{x}_1 - a_1^j)|\overline{x}_1 - a_1^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}, \dots, \frac{\operatorname{sgn}(\overline{x}_n - a_n^j)|\overline{x}_n - a_n^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}} \right) \\
& - \sum_{i=r_1+1}^{r_2} |w_i| \left(\frac{\operatorname{sgn}(\overline{x}_1 - a_1^i)|\overline{x}_1 - a_1^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}}, \dots, \frac{\operatorname{sgn}(\overline{x}_n - a_n^i)|\overline{x}_n - a_n^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} \right) \\
& = \sum_{i=1}^{r_1} |w_i| x^{*_i} + \sum_{i=r_2+1}^{r} |w_i| x^{*_i} + \sum_{j=r+1}^{r_3} |w_j| x^{*_j} + \sum_{j=r_4+1}^{m} |w_j| x^{*_j},
\end{aligned}
$$

$$
(2)\ \forall\, i \in \{1,\dots,r_1\}\ \forall\, l \in \{1,\dots,n\} \colon x_l^{*_i} \in \begin{cases} \{1\} & \text{für } \overline{x}_l - a_l^i > 0 \\ \{-1\} & \text{für } \overline{x}_l - a_l^i < 0 \\ [-1,1] & \text{für } \overline{x}_l - a_l^i = 0 \end{cases},
$$

$$
(3)\ \forall\, j \in \{r+1,\dots,r_3\}\ \forall\, l \in \{1,\dots,n\} \colon x_l^{*_j} \in \begin{cases} \{-1\} & \text{für } \overline{x}_l - a_l^j > 0 \\ \{1\} & \text{für } \overline{x}_l - a_l^j < 0 \\ \{-1,1\} & \text{für } \overline{x}_l - a_l^j = 0 \end{cases},
$$

$$
(4)\ \forall\, i \in \{r_2+1,\dots,r\}\ \forall\, l \in \{1,\dots,n\} \colon x_l^{*_i}(\overline{x}_l - a_l^i) - |x_l^{*_i}| \max_{l \in \{1,\dots,n\}} |\overline{x}_l - a_l^i| = 0,
$$

$$
(5)\ \forall\, j \in \{r_4+1,\dots,m\}\ \forall\, l \in \{1,\dots,n\} \colon x_l^{*_j}(\overline{x}_l - a_l^j) + |x_l^{*_j}| \max_{l \in \{1,\dots,n\}} |\overline{x}_l - a_l^j| = 0,
$$

(6) $\forall\, i \in \{r_2+1,\dots,r\} \colon \|x^{*_i}\|_1 = 1$ sowie

(7) $\forall\, j \in \{r_4+1,\dots,m\} \colon \|x^{*_j}\|_1 = 1$ und $\|x^{*_j}\|_\infty = 1$

erfüllen. ■

Bemerkung 4.2.1

a) *Aufgrund unserer Betrachtungen zu Beginn dieses Kapitels kann in jeder Indexmenge gleicher existierender Standorte jede Abstandsfunktion höchstens einem dieser Standorte zugeordnet sein. Insbesondere kann also in jeder solchen Indexmenge auch jeweils höchstens ein (anziehender oder abstoßender) Standort enthalten sein, dessen Abstandsfunktion von der ℓ_1-Norm bzw. der ℓ_∞-Norm induziert wird.*

   Das heißt aber, dass in der notwendigen Optimalitätsbedingung für existierende Standorte mindestens zwei der rechtsseitigen Summen in Bedingung (1) immer leer sind. Diese Bedingung ist also stets leichter zu überprüfen, als dies in Satz 4.2.1 den Anschein hat. Dennoch ist die gewählte Formulierung in Satz 4.2.1 wesentlich kompakter und übersichtlicher als eine vollständige Unterscheidung aller durch diese Überlegung möglichen neun Fälle.

   Aus demselben Grund sind außerdem in jeder Indexmenge existierender Standorte die induzierenden ℓ_p-Normen der zugeordneten Abstandsfunktionen mit $p \in \mathbb{R}$ und $p > 1$ allesamt verschieden.

b) *In der Praxis tritt häufig der Spezialfall auf, dass die existierenden Standorte $\{a^1, \dots, a^m\}$ alle paarweise verschieden sind, also $|K_s| = 1$ für alle $s \in \{1, \dots, u\}$ gilt. In diesem Fall zerfällt die Testbedingung aus Satz 4.2.1, ob die notwendige Optimalitätsbedingung bereits von einem oder mehreren der existierenden Standorte a^k mit $k \in \{1, \dots, m\}$ erfüllt wird, in bis zu sechs einfacher zu prüfende Fälle für die (möglicherweise leeren) Mengen $k \in \{1, \dots, r_1\}$, $k \in \{r_1 + 1, \dots, r_2\}$, $k \in \{r_2 + 1, \dots, r\}$, $k \in \{r + 1, \dots, r_3\}$, $k \in \{r_3 + 1, \dots, r_4\}$ und $k \in \{r_4 + 1, \dots, m\}$. Dies entspricht den verschiedenen Möglichkeiten der induzierenden Norm für die dem Standort zugeordnete Abstandsfunktion, jeweils separat für die anziehenden und abstoßenden existierenden Standorte. Die notwendige Optimalitätsbedingung für Punkte, welche keine existierenden Standorte sind, bleibt dabei wie in Satz 4.2.1 erhalten.*

Wir haben mit Satz 4.2.1 eine allgemeine notwendige Optimalitätsbedingung für skalare Medianprobleme mit Anziehung und Abstoßung sowie beliebigen Minkowski-Metriken als Abstandsmaßen hergeleitet. In den folgenden Kapiteln 5, 6 und 7 werden wir sehen, dass sich zudem auch die grundlegenden Ergebnisse dieser Kapitel, nämlich notwendige Optimalitätsbedingungen für die jeweilige Aufgabenstellung mit Anziehung und Abstoßung sowie beliebigen Minkowski-Metriken (Satz 5.1.1, Satz 6.1.1 und Satz 7.1.2) unter gewissen Voraussetzungen stets auf die Aussagen aus Satz 4.2.1 zurückführen lassen. Daher bildet dieser Satz das Kernstück unserer Arbeit.

5 Notwendige Optimalitätsbedingungen für nichtkonvexe skalare Centerprobleme

Nach dem skalaren Medianproblem (SP) gehen wir im folgenden Kapitel zur nächsten Klasse von Standortoptimierungsproblemen über und beschäftigen uns nun mit dem skalaren Centerproblem (CP) mit Anziehung und Abstoßung aus Sektion 1.2.3, für dieses Problem wollen wir ebenfalls notwendige Optimalitätsbedingungen unter Verwendung unterschiedlicher Abstandsfunktionen herleiten.

Hierbei gehen wir jedoch genau andersherum vor als in Kapitel 4 und beweisen zunächst ein allgemeines Resultat für Centerprobleme mit beliebigen, von ℓ_p-Normen induzierten Abstandsfunktionen, aus welchem sich dann die weiteren Ergebnisse des zweiten Unterkapitels ableiten lassen. Sämtliche Aussagen im gesamten Kapitel 5 sind dabei neue Resultate dieser Arbeit.

5.1 Notwendige Optimalitätsbedingungen für Centerprobleme mit gemischten Minkowski-Metriken

Wir betrachten im Folgenden das Centerproblem (CP), wobei jedem der existierenden Standorte $a^1, \dots, a^m$ eine Abstandsfunktion zugeordnet ist, welche für alle $k \in \{1, \dots, m\}$ jeweils von einer ℓ_{p_k}-Norm mit $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ induziert wird. Ohne weitere Anforderungen an die Abstandsmaße hat eine solche Aufgabenstellung in der Klassifikation aus Sektion 1.2.1 das Schema $1/\mathbb{R}^n/w_k \gtrless 0/\ell_p^m/\max$.

Ein Standort mit mehreren verschiedenen zugeordneten Abstandsfunktionen wird dabei wie gehabt als mehrere gleiche existierende Standorte modelliert, welchen jeweils eine dieser Abstandsfunktionen zugeordnet ist, wobei diesen gleichen existierenden Standorten wiederum jede mögliche Abstandsfunktion höchstens einmal zugeordnet werden darf.

Die existierenden Standorte seien außerdem in gleicher Art und Weise geordnet und ihre Indizes in Indexmengen K_s mit $s \in \{1, \dots, u\}$ zusammengefasst, wie dies in Unterkapitel 4.2 für die entsprechenden Medianprobleme eingeführt wurde. Eine notwendige Optimalitätsbedingung für ein solches Centerproblem lässt sich dann folgendermaßen auf die in Satz 4.2.1 formulierte notwendige Optimalitätsbedingung eines zugeordneten Medianproblems zurückführen:

Satz 5.1.1 (*Notw. Optimalitätsbedingung für* $1/\mathbb{R}^n/w_k \gtrless 0/\ell_p^m/\max$)

Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CP), die existierenden Standorte in Indexmengen K_s mit $s \in \{1, \dots, u\}$ aufgeteilt, welche (a) - (c) genügen, sowie $d_k(x, a^k) = \|x - a^k\|_{p_k}$ mit $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ für alle $k \in \{1, \dots, m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit

$$\Lambda_{\overline{x}} = \Big\{ \lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1, \dots, r\} : \\ \lambda_i \left(|w_i| \|\overline{x} - a^i\|_{p_i} - \max_{i \in \{1,\dots,r\}} |w_i| \|\overline{x} - a^i\|_{p_i} \right) = 0 \Big\}$$

sowie

$$\Phi_{\overline{x}} = \Big\{ \varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1, \dots, m\} : \\ \varphi_j \left(|w_j|(-\|\overline{x} - a^j\|_{p_j}) - \max_{j \in \{r+1,\dots,m\}} |w_j|(-\|\overline{x} - a^j\|_{p_j}) \right) = 0 \Big\},$$

so dass in $\overline{x}$ die notwendige Optimalitätsbedingung aus Satz 4.2.1 für das Medianproblem

$$\sum_{i=1}^{r} \overline{\lambda}_i |w_i| \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j|(-\|x - a^j\|_{p_j}) \to \min_{x \in \mathbb{R}^n}$$

erfüllt ist.

Beweis: Ist $\overline{x}$ ein lokales Minimum des Problems (CP) mit $d_k(x, a^k) = \|x - a^k\|_{p_k}$, $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ für alle $k \in \{1, \dots, m\}$, so muss $\overline{x}$ nach (L4) aus Lemma 2.2.2 die notwendige Optimalitätsbedingung

$$\begin{aligned} 0 &\in \partial_L f_w(\overline{x}) \\ &= \partial_L \left(\max_{i \in \{1,\dots,r\}} |w_i| \|x - a^i\|_{p_i} + \max_{j \in \{r+1,\dots,m\}} |w_j|(-\|x - a^j\|_{p_j}) \right)(\overline{x}) \end{aligned}$$

erfüllen.

Diese Bedingung formen wir weiter um, indem wir mehrfach die Summenregel für Limiting-Subdifferentiale aus Satz 2.2.8, das Ergebnis zum Limiting-Subdifferential der Maximumsfunktion aus Satz 3.3.1 sowie Eigenschaft (L1) aus Lemma 2.2.2, die positive Homogenität des Limiting-Subdifferentials, nutzen. Es ergibt sich damit die notwendige Optimalitätsbedingung

$$\begin{aligned} 0 &\in \partial_L \left(\max_{i \in \{1,\dots,r\}} |w_i| \|x - a^i\|_{p_i} + \max_{j \in \{r+1,\dots,m\}} |w_j|(-\|x - a^j\|_{p_j}) \right)(\overline{x}) \\ &\subseteq \partial_L \left(\max_{i \in \{1,\dots,r\}} |w_i| \|x - a^i\|_{p_i} \right)(\overline{x}) + \partial_L \left(\max_{j \in \{r+1,\dots,m\}} |w_j|(-\|x - a^j\|_{p_j}) \right)(\overline{x}) \\ &\subseteq \bigcup_{\lambda \in \Lambda_{\overline{x}}} \partial_L \left(\sum_{i=1}^{r} \lambda_i |w_i| \|x - a^i\|_{p_i} \right)(\overline{x}) + \bigcup_{\varphi \in \Phi_{\overline{x}}} \partial_L \left(\sum_{j=r+1}^{m} \varphi_j |w_j|(-\|x - a^j\|_{p_j}) \right)(\overline{x}) \\ &\subseteq \bigcup_{\lambda \in \Lambda_{\overline{x}}} \sum_{i=1}^{r} \partial_L \left(\lambda_i |w_i| \|x - a^i\|_{p_i} \right)(\overline{x}) + \bigcup_{\varphi \in \Phi_{\overline{x}}} \sum_{j=r+1}^{m} \partial_L \left(\varphi_j |w_j|(-\|x - a^j\|_{p_j}) \right)(\overline{x}) \\ &= \bigcup_{\lambda \in \Lambda_{\overline{x}}} \sum_{i=1}^{r} \lambda_i |w_i| \partial_L \left(\|x - a^i\|_{p_i} \right)(\overline{x}) + \bigcup_{\varphi \in \Phi_{\overline{x}}} \sum_{j=r+1}^{m} \varphi_j |w_j| \partial_L(-\|x - a^j\|_{p_j})(\overline{x}), \end{aligned}$$

wobei

$$\Lambda_{\overline{x}} = \Big\{ \lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1,\ldots,r\}: \\ \lambda_i \left(|w_i| \|\overline{x} - a^i\|_{p_i} - \max_{i\in\{1,\ldots,r\}} |w_i| \|\overline{x} - a^i\|_{p_i} \right) = 0 \Big\}$$

und

$$\Phi_{\overline{x}} = \Big\{ \varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1,\ldots,m\}: \\ \varphi_j \left(|w_j| (-\|\overline{x} - a^j\|_{p_j}) - \max_{j\in\{r+1,\ldots,m\}} |w_j| (-\|\overline{x} - a^j\|_{p_j}) \right) = 0 \Big\}$$

sein sollen. Es müssen also $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ existieren, so dass

$$0 \in \sum_{i=1}^{r} \overline{\lambda}_i |w_i| \partial_L \left(\|x - a^i\|_{p_i} \right)(\overline{x}) + \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| \partial_L \left(-\|x - a^j\|_{p_j} \right)(\overline{x})$$

gilt.

Dies ist aber im Beweis von Satz 4.2.1 nach den ersten Umformungsschritten gerade die notwendige Optimalitätsbedingung für das Medianproblem

$$\sum_{i=1}^{r} \overline{\lambda}_i |w_i| \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| (-\|x - a^j\|_{p_j}) \to \min_{x\in\mathbb{R}^n},$$

aus welcher wir anschließend die Aussage dieses Satzes herleiten. Daher muss in $\overline{x}$ auch die notwendige Optimalitätsbedingung aus Satz 4.2.1 für das obige Medianproblem erfüllt sein. ■

Bemerkung 5.1.1

Fast immer werden viele der Einträge von $\overline{\lambda}$ und $\overline{\varphi}$ gleich Null sein (nämlich mindestens jene, für die die gewichtete Abstandsfunktion zum zugehörigen Standort nicht dem Maximum der gewichteten Abstände entspricht). Im Einklang mit unseren Festlegungen aus Sektion 1.2.2 werden diese Komponenten im zugeordneten Medianproblem dann auch nicht berücksichtigt und die Komplexität der Aufgabenstellung damit erheblich verringert.

Besonders bei der Testbedingung, ob bereits einer oder mehrere der existierenden Standorte lokale Optima der Problemstellung sein könnten, lassen sich die zu prüfenden Bedingungen dabei erheblich vereinfachen. Ist in einer Indexmenge K_s mit $s \in \{1,\ldots,u\}$ kein Index zu einem abstoßenden Standort enthalten, also $K_s \cap \{r+1,\ldots,m\} = \emptyset$, so fallen einfach die drei entsprechenden Mengen auf der rechten Seite von Bedingung (1) im ersten Teil von Satz 4.2.1 weg. Enthält K_s hingegen den Index mindestens eines abstoßenden Standortes, dann gilt $-d_j(\overline{x}, a^j) = 0$ für alle solchen Standorte mit $j \in K_s \cap \{r+1,\ldots,m\}$ und $-d_j(\overline{x}, a^j) < 0$ für alle a^j mit $j \in \{r+1,\ldots,m\} \setminus K_s$. Daraus folgt aufgrund der Definition von $\Phi_{\overline{x}}$ aber sofort $\overline{\varphi}_j = 0$ für alle $j \notin K_s \cap \{r+1,\ldots,m\}$, in Satz 4.2.1 entfallen damit die drei zugehörigen Summen auf der linken Seite von Bedingung (1) sowie die kompletten Bedingungen (3), (5) und (7).

Ähnliche Betrachtungen lassen sich für die anziehenden Standorte durchführen, auch hier entfallen einfach die drei entsprechenden Mengen auf der rechten Seite von Bedingung (1) des ersten Teils von Satz 4.2.1, wenn K_s mit $s \in \{1, \ldots, u\}$ keinen Index eines anziehenden Standortes enthält, also $K_s \cap \{1, \ldots, r\} = \emptyset$ gilt. Enthält K_s mindestens einen, aber nicht alle der anziehenden Standorte, dann gilt $d_i(\overline{x}, a^i) = 0$ für alle Standorte mit $i \in K_s \cap \{1, \ldots, r\}$ und $d_i(\overline{x}, a^i) > 0$ für alle übrigen anziehenden Standorte mit $i \in \{1, \ldots, r\} \setminus K_s$. Aus der Definition von $\Lambda_{\overline{x}}$ folgt dann $\overline{\lambda}_i = 0$ für alle $i \in K_s \cap \{1, \ldots, r\}$, daher verschwinden auch in diesem Fall dieselben drei Mengen auf der rechten Seite von Bedingung (1). Sind in K_s hingegen alle anziehenden Standorte enthalten, also $\{1, \ldots, r\} \subseteq K_s$, dann fallen in Satz 4.2.1 direkt die drei zu den anziehenden Standorten gehörenden Summen in Bedingung (1) sowie die Bedingungen (2), (4) und (6) weg.

Wir werden auf diese Überlegungen in Unterkapitel 5.2 nochmals zurückgreifen, um die dort gewonnenen notwendigen Optimalitätsbedingungen weiter zu vereinfachen.

Bemerkung 5.1.2

Auch wenn $\overline{x} \in \mathbb{R}^n$ in Satz 5.1.1 die notwendige Optimalitätsbedingung aus Satz 4.2.1 für das Medianproblem

$$\sum_{i=1}^{r} \overline{\lambda}_i |w_i| \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| (-\|x - a^j\|_{p_j}) \to \min_{x \in \mathbb{R}^n}$$

erfüllt, so folgt daraus nicht zwangsläufig die stärkere Bedingung

$$0 \in \partial_L \left(\sum_{i=1}^{r} \overline{\lambda}_i |w_i| \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| (-\|x - a^j\|_{p_j}) \right) (\overline{x})$$

und $\overline{x}$ ist daher im Allgemeinen auch kein lokales Minimum dieses Problems.

Wir betrachten als einfaches Gegenbeispiel dazu ein Centerproblem mit vier existierenden Standorten $a^1 = (-2, 0)$, $a^2 = (2, 0)$, $a^3 = (0, -2)$ und $a^4 = (0, 2)$, wobei a^1 und a^2 anziehende Standorte mit Gewichten $w_1 = w_2 = 1$ und a^3 sowie a^4 abstoßende Standorte mit Gewichten $w_3 = w_4 = -1$ sein sollen. Allen Standorten gleichermaßen ist als Abstandsfunktion die von der ℓ_1-Norm induzierte Metrik zugeordnet. Für $x \in [-2, 2] \times [-2, 2]$ erhält man als Zielfunktionswert des zugehörigen Problems (CP)

$$\begin{aligned} f(x) &= \max \{\|(-2, 0) - x\|_1, \|(2, 0) - x\|_1\} \\ &\quad + \max \{-\|(0, -2) - x\|_1, -\|(0, 2) - x\|_1\} \\ &= \max \{|-2 - x_1| + |-x_2|, |2 - x_1| + |-x_2|\} \\ &\quad + \max \{-|-x_1| - |-2 - x_2|, -|-x_1| - |2 - x_2|\} \\ &= 2 + |x_1| + |x_2| + (-2 - |x_1| + |x_2|) \\ &= 2 \cdot |x_2|, \end{aligned}$$

für $\overline{x} = (0, 0)$ ist also $f(\overline{x}) = 0$ und daher $\overline{x}$ nach Definition 2.1.19 ein lokales (und sogar globales Minimum) dieses Problems. Abbildung 5.1 visualisiert das Beispiel nochmals.

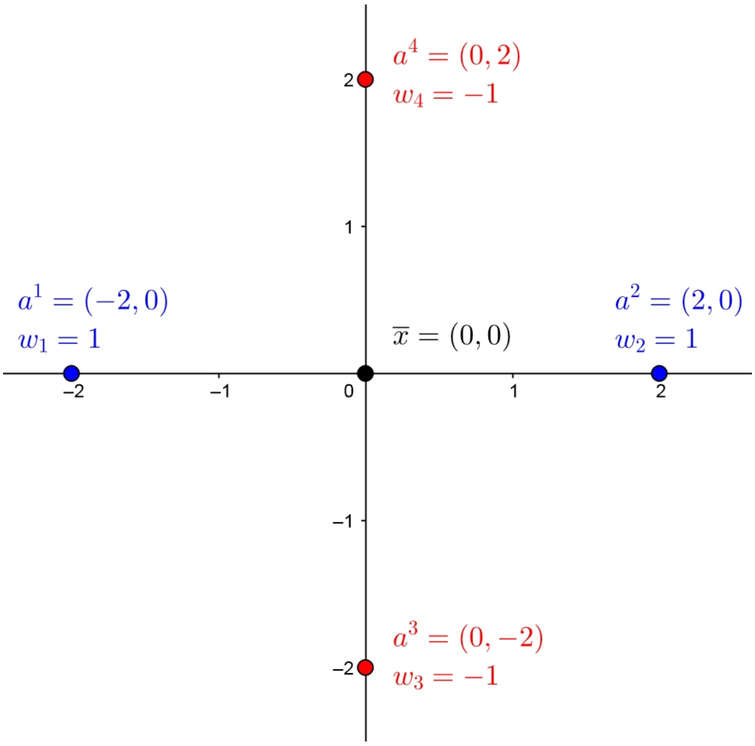

Abbildung 5.1: Ein Beispiel zu notwendigen Optimalitätsbedingungen von (CP) (eigene Darstellung mit GeoGebra)

Nach Satz 5.1.1 existieren dann

$$\overline{\lambda} \in \Lambda_{\overline{x}} = \{\lambda \in \mathbb{R}^2_+ \mid \lambda_1 + \lambda_2 = 1\} \text{ und } \overline{\varphi} \in \Phi_{\overline{x}} = \{\varphi \in \mathbb{R}^2_+ \mid \varphi_3 + \varphi_4 = 1\},$$

so dass die beiden notwendigen Optimalitätsbedingungen

$$\begin{aligned}
&0 \in \{\overline{\lambda}_1|w_1|\} + \{-\overline{\lambda}_2|w_2|\} + \{-\overline{\varphi}_3|w_3|, \overline{\varphi}_3|w_3|\} + \{-\overline{\varphi}_4|w_4|, \overline{\varphi}_4|w_4|\} \text{ und}\\
&0 \in [-\overline{\lambda}_1|w_1|, \overline{\lambda}_1|w_1|] + [-\overline{\lambda}_2|w_2|, \overline{\lambda}_2|w_2|] + \{-\overline{\varphi}_3|w_3|\} + \{\overline{\varphi}_4|w_4|\}
\end{aligned}$$

erfüllt sind. Dies ist beispielsweise für $\overline{\lambda}_1 = \overline{\lambda}_2 = \overline{\varphi}_3 = \overline{\varphi}_4 = \frac{1}{2}$ *der Fall, in diesem Beispiel gibt es aber sogar unendlich viele Möglichkeiten für die Wahl von* $\overline{\lambda} \in \Lambda_{\overline{x}}$ *und* $\overline{\varphi} \in \Phi_{\overline{x}}$. *Die Aussage von Satz 5.1.1 ist also korrekt.*

Der Zielfunktionswert des zugeordneten Medianproblems kann für jede der möglichen Gewichtungen $\overline{\lambda} \in \Lambda_{\overline{x}}$, $\overline{\varphi} \in \Phi_{\overline{x}}$ *mittels*

$$\begin{aligned}
f_M(x) &= \overline{\lambda}_1|w_1|\|x - a^1\|_1 + \overline{\lambda}_2|w_2|\|x - a^2\|_1\\
&\quad + \overline{\varphi}_3|w_3|(-\|x - a^3\|_1) + \overline{\varphi}_4|w_4|(-\|x - a^4\|_1)\\
&= \overline{\lambda}_1(|x_1 + 2| + |x_2|) + \overline{\lambda}_2(|x_1 - 2| + |x_2|)\\
&\quad - \overline{\varphi}_3(|x_1| + |x_2 + 2|) - \overline{\varphi}_4(|x_1| + |x_2 - 2|) \to \min_{x \in \mathbb{R}^n}
\end{aligned}$$

berechnet werden und beträgt für die drei Punkte $\overline{x} = (0, 0)$, $\widetilde{x} = (\varepsilon, 0)$ *und* $\widehat{x} = (-\varepsilon, 0)$ *mit beliebigem* $0 < \varepsilon < 2$ *immer*

$$\begin{aligned}
f_M(\overline{x}) &= 2\overline{\lambda}_1 + 2\overline{\lambda}_2 - 2\overline{\varphi}_3 - 2\overline{\varphi}_4\\
&= 2(\overline{\lambda}_1 + \overline{\lambda}_2) - 2(\overline{\varphi}_3 + \overline{\varphi}_4)\\
&= 0,
\end{aligned}$$

$$\begin{aligned}
f_M(\widetilde{x}) &= (2+\varepsilon)\overline{\lambda}_1 + (2-\varepsilon)\overline{\lambda}_2 - (2+\varepsilon)\overline{\varphi}_3 - (2+\varepsilon)\overline{\varphi}_4 \\
&= 2(\overline{\lambda}_1 + \overline{\lambda}_2) + \varepsilon(\overline{\lambda}_1 - \overline{\lambda}_2) - 2(\overline{\varphi}_3 + \overline{\varphi}_4) - \varepsilon(\overline{\varphi}_3 + \overline{\varphi}_4) \\
&= \varepsilon(\overline{\lambda}_1 - \overline{\lambda}_2 - 1) \\
&= -2\varepsilon\overline{\lambda}_2 \quad \textit{und} \\
f_M(\widehat{x}) &= (2-\varepsilon)\overline{\lambda}_1 + (2+\varepsilon)\overline{\lambda}_2 - (2+\varepsilon)\overline{\varphi}_3 - (2+\varepsilon)\overline{\varphi}_4 \\
&= 2(\overline{\lambda}_1 + \overline{\lambda}_2) + \varepsilon(\overline{\lambda}_2 - \overline{\lambda}_1) - 2(\overline{\varphi}_3 + \overline{\varphi}_4) - \varepsilon(\overline{\varphi}_3 + \overline{\varphi}_4) \\
&= \varepsilon(\overline{\lambda}_2 - \overline{\lambda}_1 - 1) \\
&= -2\varepsilon\overline{\lambda}_1.
\end{aligned}$$

Unabhängig von $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ existiert damit in jeder Umgebung von $\overline{x}$ mindestens ein Punkt x_0 mit $f_M(x_0) < f_M(\overline{x})$. Es gibt also unendlich viele mögliche Gewichtungen und entsprechende Medianprobleme, für welche $\overline{x}$ die notwendige Optimalitätsbedingung aus Satz 4.2.1 erfüllt, nach Definition 2.1.19 ist $\overline{x} = (0,0)$ aber für kein einziges dieser Medianprobleme ein lokales Minimum.

Eine allgemeine notwendige Optimalitätsbedingung für das nichtkonvexe Centerproblem mit Schema $1/\mathbb{R}^n/w_k \gtrless 0/\ell_p^m/\max$ kann daher nicht schärfer formuliert werden als in Satz 5.1.1.

5.2 Notwendige Optimalitätsbedingungen für Centerprobleme mit speziellen Metriken

Die soeben in Satz 5.1.1 hergeleitete notwendige Optimalitätsbedingung für das allgemeine Centerproblem mit Minkowski-Metriken wollen wir nun im zweiten Teil dieses Kapitels nutzen, um für einige speziellere Problemstellungen mit Maximumszielfunktion noch stärkere Resultate auszuarbeiten.

Dabei betrachten wir zuerst Aufgabenstellungen des Schemas $1/\mathbb{R}^n/w_k \gtrless 0/\ell_1/\max$, bei welchen allen der existierenden Standorte a^k mit $k \in \{1,\ldots,m\}$ die von der ℓ_1-Norm induzierte Abstandsfunktion $d_k(x,a^k) = d(x,a^k) = \|x-a^k\|_1$ zugeordnet wird.

Korollar 5.2.1 (*Notw. Optimalitätsbedingung für $1/\mathbb{R}^n/w_k \gtrless 0/\ell_1/\max$*)
Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CP) mit $d_k(x,a^k) = \|x-a^k\|_1$ für alle $k \in \{1,\ldots,m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit

$$\Lambda_{\overline{x}} = \Bigg\{\lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1,\ldots,r\}\colon \\ \lambda_i\left(|w_i|\|\overline{x}-a^i\|_1 - \max_{i\in\{1,\ldots,r\}} |w_i|\|\overline{x}-a^i\|_1\right) = 0\Bigg\}$$

sowie

$$\Phi_{\overline{x}} = \Bigg\{\varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1,\ldots,m\}\colon \\ \varphi_j\left(|w_j|(-\|\overline{x}-a^j\|_1) - \max_{j\in\{r+1,\ldots,m\}} |w_j|(-\|\overline{x}-a^j\|_1)\right) = 0\Bigg\},$$

so dass

$$0 \in \sum_{i=1}^{r} \begin{cases} \{\overline{\lambda}_i|w_i|\} & \text{für } \overline{x}_l > a_l^i \\ \{-\overline{\lambda}_i|w_i|\} & \text{für } \overline{x}_l < a_l^i \\ [-\overline{\lambda}_i|w_i|, \overline{\lambda}_i|w_i|] & \text{für } \overline{x}_l = a_l^i \end{cases} + \sum_{j=r+1}^{m} \begin{cases} \{-\overline{\varphi}_j|w_j|\} & \text{für } \overline{x}_l > a_l^j \\ \{\overline{\varphi}_j|w_j|\} & \text{für } \overline{x}_l < a_l^j \\ \{-\overline{\varphi}_j|w_j|, \overline{\varphi}_j|w_j|\} & \text{für } \overline{x}_l = a_l^j \end{cases}$$

für alle $l \in \{1, \ldots, n\}$ erfüllt ist.

Beweis: Die Aussage folgt sofort durch Einsetzen von $p_k = 1$ für alle $k \in \{1, \ldots, m\}$ in Satz 5.1.1. ■

Als Nächstes wollen wir einige notwendige Optimalitätsbedingungen für verschiedene Aufgabenstellungen herleiten, in deren Zielfunktionen keine von der ℓ_1-Norm oder von der ℓ_∞-Norm induzierten Metriken vorkommen, die zu den existierenden Standorten a^k zugehörigen Abstandsfunktionen werden demzufolge für alle $k \in \{1, \ldots, m\}$ von ℓ_{p_k}-Normen mit $p_k \in \mathbb{R}$ und $p_k > 1$ erzeugt. Das folgende Korollar gibt ein allgemeines Resultat für solche Problemstellungen mit dem Klassifikationsschema $1/\mathbb{R}^n/w_k \gtrless 0/\ell^m_{p,1<p<+\infty}/\max$ an.

Korollar 5.2.2 (*Notw. Optimalitätsbed. für* $1/\mathbb{R}^n/w_k \gtrless 0/\ell^m_{p,1<p<+\infty}/\max$)
Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CP), die existierenden Standorte in Indexmengen K_s mit $s \in \{1, \ldots, u\}$ aufgeteilt, welche (a) - (c) genügen, sowie $d_k(x, a^k) = \|x - a^k\|_{p_k}$ mit $p_k \in \mathbb{R}$ und $p_k > 1$ für alle $k \in \{1, \ldots, m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit

$$\Lambda_{\overline{x}} = \Bigg\{\lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1, \ldots, r\}: \\ \lambda_i \left(|w_i|\|\overline{x} - a^i\|_{p_i} - \max_{i \in \{1,\ldots,r\}} |w_i|\|\overline{x} - a^i\|_{p_i} \right) = 0 \Bigg\}$$

sowie

$$\Phi_{\overline{x}} = \Bigg\{\varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1, \ldots, m\}: \\ \varphi_j \left(|w_j|(-\|\overline{x} - a^j\|_{p_j}) - \max_{j \in \{r+1,\ldots,m\}} |w_j|(-\|\overline{x} - a^j\|_{p_j}) \right) = 0 \Bigg\},$$

und es ist entweder $\overline{x} = a^g$ für ein $g \in K_s$ mit $s \in \{1, \ldots, u\}$, so dass

$$\sum_{\substack{j=r+1 \\ j \notin K_s}}^{m} \overline{\varphi}_j|w_j| \left(\frac{\operatorname{sgn}(a^g_1 - a^j_1)|a^g_1 - a^j_1|^{p_j-1}}{\|a^g - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(a^g_n - a^j_n)|a^g_n - a^j_n|^{p_j-1}}{\|a^g - a^j\|_{p_j}^{p_j-1}} \right)$$
$$- \sum_{\substack{i=1 \\ i \notin K_s}}^{r} \overline{\lambda}_i|w_i| \left(\frac{\operatorname{sgn}(a^g_1 - a^i_1)|a^g_1 - a^i_1|^{p_i-1}}{\|a^g - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a^g_n - a^i_n)|a^g_n - a^i_n|^{p_i-1}}{\|a^g - a^i\|_{p_i}^{p_i-1}} \right)$$
$$\in \sum_{\substack{i=1 \\ i \in K_s}}^{r} \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_i}{p_i-1}} \leq \overline{\lambda}_i|w_i| \right\} + \sum_{\substack{j=r+1 \\ j \in K_s}}^{m} \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_j}{p_j-1}} = \overline{\varphi}_j|w_j| \right\}$$

gilt oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ und $\overline{x}$ löst das System nichtlinearer Gleichungen

$$0 = \sum_{i=1}^{r} \overline{\lambda}_i |w_i| \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis: Die Aussage folgt mit $p_k \neq 1$ und $p_k \neq \infty$ für alle $k \in \{1, \ldots, m\}$ direkt aus Satz 5.1.1. ■

Sind alle der existierenden Standorte $a^1, \ldots, a^m$ paarweise verschieden, dann lässt sich die notwendige Optimalitätsbedingung aus Korollar 5.2.2 nochmals stark vereinfachen:

Korollar 5.2.3

Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CP) mit paarweise verschiedenen existierenden Standorten $a^1, \ldots, a^m$ sowie $d_k(x, a^k) = \|x - a^k\|_{p_k}$ mit $p_k \in \mathbb{R}$ und $p_k > 1$ für alle $k \in \{1, \ldots, m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit

$$\Lambda_{\overline{x}} = \Big\{ \lambda \in \mathbb{R}_+^r \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1, \ldots, r\} : \\ \lambda_i \left(|w_i| \|\overline{x} - a^i\|_{p_i} - \max_{i \in \{1,\ldots,r\}} |w_i| \|\overline{x} - a^i\|_{p_i} \right) = 0 \Big\}$$

sowie

$$\Phi_{\overline{x}} = \Big\{ \varphi \in \mathbb{R}_+^{m-r} \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1, \ldots, m\} : \\ \varphi_j \left(|w_j|(-\|\overline{x} - a^j\|_{p_j}) - \max_{j \in \{r+1,\ldots,m\}} |w_j|(-\|\overline{x} - a^j\|_{p_j}) \right) = 0 \Big\},$$

und es ist entweder $\overline{x} = a^k$ für ein $k \in \{r+1, \ldots, m\}$, so dass

$$\left\| \sum_{i=1}^{r} \overline{\lambda}_i |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}} \right) \right\|_{\frac{p_k}{p_k-1}} = |w_k|$$

gilt oder es ist $\overline{x} = a^1$ der einzige existierende anziehende Standort und es gilt

$$\left\| \sum_{j=2}^{m} \overline{\varphi}_j |w_j| \left(\frac{\operatorname{sgn}(a_1^1 - a_1^j)|a_1^1 - a_1^j|^{p_j-1}}{\|a^1 - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(a_n^1 - a_n^j)|a_n^1 - a_n^j|^{p_j-1}}{\|a^1 - a^j\|_{p_j}^{p_j-1}} \right) \right\|_{\frac{p_1}{p_1-1}} \leq |w_1|$$

oder es ist $\overline{x} = a^k$ für ein $k \in \{1, \ldots, r\}$ mit $r \geq 2$ und es gilt

$$0 = \sum_{\substack{i=1 \\ i \neq k}}^{r} \overline{\lambda}_i |w_i| \frac{\operatorname{sgn}(a_l^k - a_l^i)|a_l^k - a_l^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| \frac{\operatorname{sgn}(a_l^k - a_l^j)|a_l^k - a_l^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}}$$

für alle $l \in \{1, \ldots, n\}$ oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ und $\overline{x}$ löst das System nichtlinearer Gleichungen

$$0 = \sum_{i=1}^{r} \overline{\lambda}_i |w_i| \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p_i - 1}}{\|\overline{x} - a^i\|_{p_i}^{p_i - 1}} - \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p_j - 1}}{\|\overline{x} - a^j\|_{p_j}^{p_j - 1}}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis: Die Aussage folgt mit $|K_s| = 1$ für alle $s \in \{1, \ldots, u\}$ aus Korollar 5.2.2. Zur Vereinfachung der notwendigen Optimalitätsbedingungen in den existierenden Standorten $\overline{x} = a^k$ für alle $k \in \{1, \ldots, m\}$ nutzen wir außerdem die Vorbetrachtungen aus Bemerkung 5.1.1. ■

In Korollar 5.2.3 ist auch der Spezialfall inbegriffen, dass alle Abstandsfunktionen zu den existierenden Standorten von derselben ℓ_p-Norm mit $p \in \mathbb{R}$ und $p > 1$ induziert werden, dass also $d_k(x, a^k) = d(x, a^k) = \|x - a^k\|_p$ für alle $k \in \{1, \ldots, m\}$ gilt. Wir wollen diese Folgerung aufgrund ihrer Wichtigkeit aber noch einmal separat festhalten.

Korollar 5.2.4 (*Notw. Optimalitätsbed. für* $\mathbf{1/\mathbb{R}^n/w_k \gtrless 0/\ell_{p,1<p<+\infty}/\max}$)
Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CP) mit $d_k(x, a^k) = \|x - a^k\|_p$ für ein $p \in \mathbb{R}$ mit $p > 1$ und für alle $k \in \{1, \ldots, m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit

$$\Lambda_{\overline{x}} = \Bigg\{ \lambda \in \mathbb{R}_+^r \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1, \ldots, r\}: \\ \lambda_i \left(|w_i| \|\overline{x} - a^i\|_p - \max_{i \in \{1, \ldots, r\}} |w_i| \|\overline{x} - a^i\|_p \right) = 0 \Bigg\}$$

sowie

$$\Phi_{\overline{x}} = \Bigg\{ \varphi \in \mathbb{R}_+^{m-r} \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1, \ldots, m\}: \\ \varphi_j \left(|w_j|(-\|\overline{x} - a^j\|_p) - \max_{j \in \{r+1, \ldots, m\}} |w_j|(-\|\overline{x} - a^j\|_p) \right) = 0 \Bigg\},$$

und es ist entweder $\overline{x} = a^k$ für ein $k \in \{r+1, \ldots, m\}$, so dass

$$\left\| \sum_{i=1}^{r} \overline{\lambda}_i |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} \right) \right\|_{\frac{p}{p-1}} = |w_k|$$

gilt oder es ist $\overline{x} = a^1$ der einzige existierende anziehende Standort und es gilt

$$\left\| \sum_{j=2}^{m} \overline{\varphi}_j |w_j| \left(\frac{\operatorname{sgn}(a_1^1 - a_1^j)|a_1^1 - a_1^j|^{p-1}}{\|a^1 - a^j\|_p^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^1 - a_n^j)|a_n^1 - a_n^j|^{p-1}}{\|a^1 - a^j\|_p^{p-1}} \right) \right\|_{\frac{p}{p-1}} \leq |w_1|$$

oder es ist $\overline{x} = a^k$ für ein $k \in \{1, \ldots, r\}$ mit $r \geq 2$ und es gilt

$$0 = \sum_{\substack{i=1 \\ i \neq k}}^{r} \overline{\lambda}_i |w_i| \frac{\operatorname{sgn}(a_l^k - a_l^i)|a_l^k - a_l^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} - \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| \frac{\operatorname{sgn}(a_l^k - a_l^j)|a_l^k - a_l^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}$$

für alle $l \in \{1, \ldots, n\}$ oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ und $\overline{x}$ löst das System nichtlinearer Gleichungen

$$0 = \sum_{i=1}^{r} \overline{\lambda}_i |w_i| \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p-1}}{\|\overline{x} - a^i\|_p^{p-1}} - \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p-1}}{\|\overline{x} - a^j\|_p^{p-1}}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis: Dieses Resultat ist Korollar 5.2.3 mit $p_k = p$ für alle $k \in \{1, \ldots, m\}$. ∎

Für praktische Anwendungen wie die Standortoptimierung ist die euklidische Norm ℓ_2 die wichtigste der ℓ_p-Normen, da sie den intuitiven Abstandsbegriff widerspiegelt. Deswegen geben wir noch eine eigene notwendige Optimalitätsbedingung für Aufgabenstellungen mit Maximumszielfunktion an, bei denen die Abstandsfunktionen $d_k(x, a^k) = \|x - a^k\|_2$ für alle $k \in \{1, \ldots, m\}$ von der ℓ_2-Norm induziert sind.

Korollar 5.2.5 (*Notw. Optimalitätsbedingung für* $\mathbf{1/\mathbb{R}^n/w_k \gtrless 0/\ell_2/\max}$)

Ist $\overline{x} \in \mathbb{R}^n$ *ein lokales Minimum des Problems (CP) mit* $d_k(x, a^k) = \|x - a^k\|_2$ *für alle* $k \in \{1, \ldots, m\}$, *dann existieren* $\overline{\lambda} \in \Lambda_{\overline{x}}$ *und* $\overline{\varphi} \in \Phi_{\overline{x}}$ *mit*

$$\Lambda_{\overline{x}} = \Bigg\{ \lambda \in \mathbb{R}_+^r \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1, \ldots, r\} : \\ \lambda_i \left(|w_i| \|\overline{x} - a^i\|_2 - \max_{i \in \{1,\ldots,r\}} |w_i| \|\overline{x} - a^i\|_2 \right) = 0 \Bigg\}$$

sowie

$$\Phi_{\overline{x}} = \Bigg\{ \varphi \in \mathbb{R}_+^{m-r} \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1, \ldots, m\} : \\ \varphi_j \left(|w_j|(-\|\overline{x} - a^j\|_2) - \max_{j \in \{r+1,\ldots,m\}} |w_j|(-\|\overline{x} - a^j\|_2) \right) = 0 \Bigg\},$$

und es ist entweder $\overline{x} = a^k$ *für ein* $k \in \{r+1, \ldots, m\}$, *so dass*

$$\left\| \sum_{i=1}^{r} \overline{\lambda}_i |w_i| \frac{a^k - a^i}{\|a^k - a^i\|_2} \right\|_2 = |w_k|$$

gilt oder es ist $\overline{x} = a^1$ *der einzige existierende anziehende Standort und es gilt*

$$\left\| \sum_{j=2}^{m} \overline{\varphi}_j |w_j| \frac{a^1 - a^j}{\|a^1 - a^j\|_2} \right\|_2 \leq |w_1|$$

oder es ist $\overline{x} = a^k$ *für ein* $k \in \{1, \ldots, r\}$ *mit* $r \geq 2$ *und es gilt*

$$0 = \sum_{\substack{i=1 \\ i \neq k}}^{r} \overline{\lambda}_i |w_i| \frac{a_l^k - a_l^i}{\|a^k - a^i\|_2} - \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| \frac{a_l^k - a_l^j}{\|a^k - a^j\|_2}$$

für alle $l \in \{1, \ldots, n\}$ *oder es ist* $\overline{x} \notin \{a^1, \ldots, a^m\}$ *und* $\overline{x}$ *löst das System nichtlinearer Gleichungen*

$$0 = \sum_{i=1}^{r} \overline{\lambda}_i |w_i| \frac{\overline{x}_l - a_l^i}{\|\overline{x} - a^i\|_2} - \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| \frac{\overline{x}_l - a_l^j}{\|\overline{x} - a^j\|_2}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis: Die Aussage folgt mit $p = 2$ direkt aus Korollar 5.2.4. ■

Zum Abschluss dieses Unterkapitels präsentieren wir nun noch eine notwendige Optimalitätsbedingung für Problemstellungen mit ausschließlich von der ℓ_∞-Norm induzierten Metriken, wenn also $d_k(x, a^k) = \|x - a^k\|_\infty$ für alle $k \in \{1, \ldots, m\}$ gilt.

Korollar 5.2.6 (*Notw. Optimalitätsbedingung für* $\mathbf{1/\mathbb{R}^n/w_k \gtrless 0/\ell_\infty/\max}$)
Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CP) mit $d_k(x, a^k) = \|x - a^k\|_\infty$ für alle $k \in \{1, \ldots, m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit

$$\Lambda_{\overline{x}} = \Bigg\{ \lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1, \ldots, r\} : \\ \lambda_i \left(|w_i| \|\overline{x} - a^i\|_\infty - \max_{i \in \{1,\ldots,r\}} |w_i| \|\overline{x} - a^i\|_\infty \right) = 0 \Bigg\}$$

sowie

$$\Phi_{\overline{x}} = \Bigg\{ \varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1, \ldots, m\} : \\ \varphi_j \left(|w_j| (-\|\overline{x} - a^j\|_\infty) - \max_{j \in \{r+1,\ldots,m\}} |w_j| (-\|\overline{x} - a^j\|_\infty) \right) = 0 \Bigg\},$$

und es ist entweder $\overline{x} = a^k$ für ein $k \in \{r+1, \ldots, m\}$ und es existieren $x^{*_i} \in \mathbb{R}^n$ für alle $i \in \{1, \ldots, r\}$, welche

$$\text{(1)} \quad \left\| \sum_{i=1}^{r} \overline{\lambda}_i |w_i| x^{*_i} \right\|_1 = |w_k|, \quad \left\| \sum_{i=1}^{r} \overline{\lambda}_i |w_i| x^{*_i} \right\|_\infty = |w_k|,$$

$$\text{(2)} \quad x^{*_i}_l (a^k_l - a^i_l) - |x^{*_i}_l| \max_{l \in \{1,\ldots,n\}} |a^k_l - a^i_l| = 0 \text{ und } \|x^{*_i}\|_1 = 1$$

für alle $i \in \{1, \ldots, r\}$ und für alle $l \in \{1, \ldots, n\}$ erfüllen oder es ist $\overline{x} = a^1$ der einzige existierende anziehende Standort und es existieren $x^{*_j} \in \mathbb{R}^n$ für alle $j \in \{2, \ldots, m\}$, welche

$$\text{(1)} \quad \left\| \sum_{j=2}^{m} \overline{\varphi}_j |w_j| x^{*_j} \right\|_1 \leq |w_1|,$$

$$\text{(2)} \quad x^{*_j}_l (a^1_l - a^j_l) + |x^{*_j}_l| \max_{l \in \{1,\ldots,n\}} |a^1_l - a^j_l| = 0,$$

$$\text{(3)} \quad \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1$$

für alle $j \in \{2, \ldots, m\}$ und für alle $l \in \{1, \ldots, n\}$ erfüllen oder es ist $\overline{x} = a^k$ für ein $k \in \{1, \ldots, r\}$ mit $r \geq 2$ und es existieren $x^{*_i} \in \mathbb{R}^n$ für alle $i \in \{1, \ldots, r\} \backslash \{k\}$ sowie $x^{*_j} \in \mathbb{R}^n$ für alle $j \in \{r+1, \ldots, m\}$, welche

$$\text{(1)} \quad \sum_{\substack{i=1 \\ i \neq k}}^{r} \overline{\lambda}_i |w_i| x^{*_i} + \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| x^{*_j} = 0,$$

$$\text{(2)} \quad x^{*_i}_l (a^k_l - a^i_l) - |x^{*_i}_l| \max_{l \in \{1,\ldots,n\}} |a^k_l - a^i_l| = 0,$$

$$\text{(3)} \quad x^{*_j}_l (a^k_l - a^j_l) + |x^{*_j}_l| \max_{l \in \{1,\ldots,n\}} |a^k_l - a^j_l| = 0,$$

$$\text{(4)} \quad \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1$$

für alle $i \in \{1, \ldots, r\} \setminus \{k\}$, *für alle* $j \in \{r+1, \ldots, m\}$ *und für alle* $l \in \{1, \ldots, n\}$ *erfüllen oder es ist schließlich* $\overline{x} \notin \{a^1, \ldots, a^m\}$ *und es existieren* $x^{*_i} \in \mathbb{R}^n$ *für alle* $i \in \{1, \ldots, r\}$ *sowie* $x^{*_j} \in \mathbb{R}^n$ *für alle* $j \in \{r+1, \ldots, m\}$, *welche*

$$\text{(1)} \quad \sum_{i=1}^{r} \overline{\lambda}_i |w_i| x^{*_i} + \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| x^{*_j} = 0,$$

$$\text{(2)} \quad x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^i| = 0,$$

$$\text{(3)} \quad x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^j| = 0,$$

$$\text{(4)} \quad \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1$$

für alle $i \in \{1, \ldots, r\}$, *für alle* $j \in \{r+1, \ldots, m\}$ *und für alle* $l \in \{1, \ldots, n\}$ *erfüllen.*

Beweis: Die Aussage folgt mit $p_k = \infty$ für alle $k \in \{1, \ldots, m\}$ und unter Nutzung der vorherigen Überlegungen in Bemerkung 5.1.1 direkt aus Satz 5.1.1. ■

6 Notwendige Optimalitätsbedingungen für nichtkonvexe skalare Centdianprobleme

Im folgenden Kapitel wollen wir mit dem skalaren Centdianproblem die letzte der drei skalaren Problemstellungen untersuchen, welche wir am Anfang der Arbeit in Unterkapitel 1.2 formuliert hatten. Bei Aufgabenstellungen mit Centdian-Ansatz wird die Zielfunktion durch eine Konvexkombination aus den Zielfunktionen je eines Median- und eines Centerproblems gebildet. Es liegt daher nahe, auch zur Gewinnung von notwendigen Optimalitätsbedingungen für das Centdianproblem diese zunächst auf notwendige Optimalitätsbedingungen des Median- und des Centerproblems zurückzuführen und dann auf die für diese beiden Probleme bereits in den Kapiteln 4 und 5 hergeleiteten Aussagen zurückzugreifen. Erneut sind alle der in diesem Kapitel präsentierten Ergebnisse bisher unveröffentlichte und im Rahmen dieser Promotion neu entwickelte Resultate.

6.1 Notwendige Optimalitätsbedingungen für Centdianprobleme mit gemischten Minkowski-Metriken

Wie bereits im vorhergehenden Kapitel wollen wir auch für die Centdianprobleme zuerst eine notwendige Optimalitätsbedingung für den allgemeinen Fall herleiten, in welchem die zu den existierenden Standorten a^k gehörenden Abstandsfunktionen für alle $k \in \{1, \dots, m\}$ von beliebigen und gegebenenfalls komplett verschiedenen ℓ_{p_k}-Normen mit $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ induziert werden. In der Klassifikation der Standortoptimierungsprobleme hat diese Aufgabenstellung das Schema $1/\mathbb{R}^n/w_k \gtreqless 0/\ell_p^m/\,\mathrm{CD}$. Das erzielte Ergebnis für den allgemeinen Fall nutzen wir dann im nächsten Unterkapitel 6.2, um durch den Übergang zu konkreten Normen notwendige Optimalitätsbedingungen für einige wichtige speziellere Problemstellungen zu folgern.

Für die Modellierung fordern wir auch hier, dass jedem existierenden Standort jede Abstandsfunktion höchstens einmal zugeordnet ist, anderenfalls lassen sich die Gewichte der betreffenden Standorte einfach aufaddieren. Mit gleichen existierenden Standorten, denen jedoch jeweils verschiedene Abstandsfunktionen zugeordnet sind, bilden wir in unserem Modell wie zuvor den möglicherweise auftretenden Fall eines Standortes mit mehreren zugeordneten Abstandsfunktionen ab.

Zudem ordnen wir die existierenden Standorte wieder genauso nach ihren zugeordneten Abstandsfunktionen um und fassen Indizes gleicher Standorte in Indexmengen K_s mit $s \in \{1, \dots, u\}$ zusammen, wie dies am Anfang von Unterkapitel 4.2 eingeführt wurde. Damit lässt sich nun unter Nutzung von Satz 4.2.1 und Satz 5.1.1 eine allgemeine notwendige Optimalitätsbedingung für das Problem (CDP) formulieren.

Satz 6.1.1 (***Notw. Optimalitätsbedingung für*** $\mathbf{1/\mathbb{R}^n/w_k \gtrless 0/\ell_p^m/\,CD}$)

Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CDP) mit $\alpha \in [0,1]$, die existierenden Standorte in Indexmengen K_s mit $s \in \{1,\ldots,u\}$ aufgeteilt, welche (a) - (c) genügen, sowie $d_k(x,a^k) = \|x-a^k\|_{p_k}$ mit $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ für alle $k \in \{1,\ldots,m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit

$$\Lambda_{\overline{x}} = \Bigg\{\lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^r \lambda_i = 1, \forall\, i \in \{1,\ldots,r\}: \\ \lambda_i \left(|w_i|\|\overline{x}-a^i\|_{p_i} - \max_{i\in\{1,\ldots,r\}} |w_i|\|\overline{x}-a^i\|_{p_i}\right) = 0\Bigg\}$$

sowie

$$\Phi_{\overline{x}} = \Bigg\{\varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^m \varphi_j = 1, \forall\, j \in \{r+1,\ldots,m\}: \\ \varphi_j \left(|w_j|(-\|\overline{x}-a^j\|_{p_j}) - \max_{j\in\{r+1,\ldots,m\}} |w_j|(-\|\overline{x}-a^j\|_{p_j})\right) = 0\Bigg\},$$

so dass in $\overline{x}$ die notwendige Optimalitätsbedingung aus Satz 4.2.1 für das Medianproblem

$$\sum_{i=1}^r(\alpha+(1-\alpha)\overline{\lambda}_i)|w_i|\|x-a^i\|_{p_i} + \sum_{j=r+1}^m (\alpha+(1-\alpha)\overline{\varphi}_j)|w_j|(-\|x-a^j\|_{p_j}) \to \min_{x\in\mathbb{R}^n}$$

erfüllt ist.

Beweis: Ist $\overline{x}$ ein lokales Minimum des Problems (CDP) mit $d_k(x,a^k) = \|x-a^k\|_{p_k}$, $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ für alle $k \in \{1,\ldots,m\}$ sowie $\alpha \in [0,1]$, dann folgt aufgrund von Eigenschaft (L4) aus Lemma 2.2.2, dass $\overline{x}$ stets die notwendige Optimalitätsbedingung

$$\begin{aligned} 0 &\in \partial_L f_{w,\alpha}(\overline{x}) \\ &= \partial_L \Bigg(\alpha\left(\sum_{i=1}^r |w_i|\|x-a^i\|_{p_i} + \sum_{j=r+1}^m |w_j|(-\|x-a^j\|_{p_j})\right) \\ &\qquad + (1-\alpha)\left(\max_{i\in\{1,\ldots,r\}} |w_i|\|x-a^i\|_{p_i} + \max_{j\in\{r+1,\ldots,m\}} |w_j|(-\|x-a^j\|_{p_j})\right)\Bigg)(\overline{x}) \end{aligned}$$

erfüllen muss. Wir formen diese Bedingung um, indem wir wiederum Satz 2.2.8 (die Summenregel für das Limiting-Subdifferential) und (L1) aus Lemma 2.2.2 (die positive Homogenität des Limiting-Subdifferentials) anwenden. Damit muss für jedes lokale Minimum $\overline{x}$ des Problems (CDP)

$$\begin{aligned} 0 \in \partial_L &\Bigg(\alpha\left(\sum_{i=1}^r |w_i|\|x-a^i\|_{p_i} + \sum_{j=r+1}^m |w_j|(-\|x-a^j\|_{p_j})\right) \\ &\qquad + (1-\alpha)\left(\max_{i\in\{1,\ldots,r\}} |w_i|\|x-a^i\|_{p_i} + \max_{j\in\{r+1,\ldots,m\}} |w_j|(-\|x-a^j\|_{p_j})\right)\Bigg)(\overline{x}) \end{aligned}$$

$$\begin{aligned}
&\subseteq \partial_L \left(\alpha \left(\sum_{i=1}^{r} |w_i| \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} |w_j| (-\|x - a^j\|_{p_j}) \right) \right) (\overline{x}) \\
&\quad + \partial_L \left((1-\alpha) \left(\max_{i \in \{1,\dots,r\}} |w_i| \|x - a^i\|_{p_i} + \max_{j \in \{r+1,\dots,m\}} |w_j| (-\|x - a^j\|_{p_j}) \right) \right) (\overline{x}) \\
&= \alpha \left(\partial_L \left(\sum_{i=1}^{r} |w_i| \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} |w_j| (-\|x - a^j\|_{p_j}) \right) \right) (\overline{x}) \\
&\quad + (1-\alpha) \left(\partial_L \left(\max_{i \in \{1,\dots,r\}} |w_i| \|x - a^i\|_{p_i} + \max_{j \in \{r+1,\dots,m\}} |w_j| (-\|x - a^j\|_{p_j}) \right) \right) (\overline{x})
\end{aligned}$$

gelten.

Die weiteren Umformungen für diese beiden Limiting-Subdifferentiale haben wir in den Beweisen von Satz 4.2.1 und von Satz 5.1.1 bereits vorgenommen. Setzen wir die dabei erhaltenen Resultate hier ein, so ergibt sich für das Centdianproblem (CDP) die notwendige Optimalitätsbedingung

$$\begin{aligned}
0 &\in \alpha \left(\partial_L \left(\sum_{i=1}^{r} |w_i| \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} |w_j| (-\|x - a^j\|_{p_j}) \right) \right) (\overline{x}) \\
&\quad + (1-\alpha) \left(\partial_L \left(\max_{i \in \{1,\dots,r\}} |w_i| \|x - a^i\|_{p_i} + \max_{j \in \{r+1,\dots,m\}} |w_j| (-\|x - a^j\|_{p_j}) \right) \right) (\overline{x}) \\
&\subseteq \alpha \left(\sum_{i=1}^{r} |w_i| \partial_L (\|x - a^i\|_{p_i}) + \sum_{j=r+1}^{m} |w_j| \partial_L (-\|x - a^j\|_{p_j}) \right) (\overline{x}) \\
&\quad + (1-\alpha) \left(\bigcup_{\lambda \in \Lambda_{\overline{x}}} \sum_{i=1}^{r} \lambda_i |w_i| \partial_L \left(\|x - a^i\|_{p_i} \right) + \bigcup_{\varphi \in \Phi_{\overline{x}}} \sum_{j=r+1}^{m} \varphi_j |w_j| \partial_L (-\|x - a^j\|_{p_j}) \right) (\overline{x}),
\end{aligned}$$

wobei wieder

$$\begin{aligned}
\Lambda_{\overline{x}} = \Bigg\{ \lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1,\dots,r\} : \\
\lambda_i \left(|w_i| \|\overline{x} - a^i\|_{p_i} - \max_{i \in \{1,\dots,r\}} |w_i| \|\overline{x} - a^i\|_{p_i} \right) = 0 \Bigg\}
\end{aligned}$$

und

$$\begin{aligned}
\Phi_{\overline{x}} = \Bigg\{ \varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1,\dots,m\} : \\
\varphi_j \left(|w_j| (-\|\overline{x} - a^j\|_{p_j}) - \max_{j \in \{r+1,\dots,m\}} |w_j| (-\|\overline{x} - a^j\|_{p_j}) \right) = 0 \Bigg\}
\end{aligned}$$

sein sollen. Somit müssen $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ existieren, so dass

$$\begin{aligned}
0 &\in \alpha \left(\sum_{i=1}^{r} |w_i| \partial_L (\|x - a^i\|_{p_i}) + \sum_{j=r+1}^{m} |w_j| \partial_L (-\|x - a^j\|_{p_j}) \right) (\overline{x}) \\
&\quad + (1-\alpha) \left(\sum_{i=1}^{r} \overline{\lambda}_i |w_i| \partial_L \left(\|x - a^i\|_{p_i} \right) + \sum_{j=r+1}^{m} \overline{\varphi}_j |w_j| \partial_L (-\|x - a^j\|_{p_j}) \right) (\overline{x})
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{r} \alpha |w_i| \partial_L (\|x - a^i\|_{p_i})(\overline{x}) + \sum_{j=r+1}^{m} \alpha |w_j| \partial_L (-\|x - a^j\|_{p_j})(\overline{x}) \\
&\quad + \sum_{i=1}^{r} (1-\alpha)\overline{\lambda}_i |w_i| \partial_L \left(\|x - a^i\|_{p_i}\right)(\overline{x}) + \sum_{j=r+1}^{m} (1-\alpha)\overline{\varphi}_j |w_j| \partial_L (-\|x - a^j\|_{p_j})(\overline{x}) \\
&= \sum_{i=1}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i) |w_i| \partial_L (\|x - a^i\|_{p_i})(\overline{x}) \\
&\quad + \sum_{j=r+1}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j) |w_j| \partial_L (-\|x - a^j\|_{p_j})(\overline{x})
\end{aligned}
$$

gilt.

Das entspricht aber genau der notwendigen Optimalitätsbedingung im Punkt $\overline{x}$ für das Medianproblem

$$
\sum_{i=1}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| (-\|x - a^j\|_{p_j}) \to \min_{x \in \mathbb{R}^n},
$$

welche am Anfang des Beweises von Satz 4.2.1 durch die ersten Umformungen entstanden ist.

Dementsprechend muss $\overline{x}$ auch die im weiteren Verlauf des Beweises daraus gefolgerten Bedingungen erfüllen, also muss in $\overline{x}$ die notwendige Optimalitätsbedingung aus Satz 4.2.1 für ebendieses Medianproblem erfüllt sein. ■

Bemerkung 6.1.1

Die Einschränkungen, welche wir in Bemerkung 5.1.2 hinsichtlich der notwendigen Optimalitätsbedingung des Centerproblems (CP) aus Satz 5.1.1 festgestellt haben, gelten in analoger Weise auch für das Centdianproblem (für $\alpha \in [0,1)$ kann beispielsweise sogar das gleiche Gegenbeispiel wie in Bemerkung 5.1.2 betrachtet werden).

Wenn $\overline{x} \in \mathbb{R}^n$ die notwendige Optimalitätsbedingung aus Satz 6.1.1 erfüllt und somit der notwendigen Optimalitätsbedingung aus Satz 4.2.1 für das Medianproblem

$$
\sum_{i=1}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| (-\|x - a^j\|_{p_j}) \to \min_{x \in \mathbb{R}^n}
$$

mit $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ genügt, dann folgt daraus im Allgemeinen nicht die stärkere Aussage

$$
\begin{aligned}
0 \in \partial_L \Bigg(&\sum_{i=1}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \|x - a^i\|_{p_i} \\
&+ \sum_{j=r+1}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| (-\|x - a^j\|_{p_j}) \Bigg)(\overline{x}).
\end{aligned}
$$

Es lässt sich also aus Satz 6.1.1 nicht ableiten, ob $\overline{x}$ auch ein lokales Minimum eines solchen verknüpften Medianproblems ist.

6.2 Notwendige Optimalitätsbedingungen für Centdianprobleme mit speziellen Metriken

Wir wollen die soeben in Unterkapitel 6.1 hergeleitete notwendige Optimalitätsbedingung für den allgemeinen Fall aus Satz 6.1.1 nun nutzen, um für einige Problemstellungen mit Centdian-Zielfunktion und gleichen oder gleichartigen Abstandsfunktionen zu allen existierenden Standorten besser handhabbare Resultate abzuleiten.

Zuerst betrachten wir dabei wieder den Fall, dass allen existierenden Standorten die von der ℓ_1-Norm induzierte Metrik als Abstandsfunktion zugeordnet ist, dass also $d_k(x,a^k) = d(x,a^k) = \|x - a^k\|_1$ für alle $k \in \{1,\dots,m\}$ gilt. Diese Aufgabe hat in der Klassifikation der Standortoptimierungsprobleme das Schema $1/\mathbb{R}^n/w_k \gtrless 0/\ell_1/$ CD.

Korollar 6.2.1 (*Notw. Optimalitätsbedingung für* $\boldsymbol{1/\mathbb{R}^n/w_k \gtrless 0/\ell_1/}$ CD)
Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CDP) mit $\alpha \in [0,1]$ und $d_k(x,a^k) = \|x - a^k\|_1$ für alle $k \in \{1,\dots,m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit

$$\Lambda_{\overline{x}} = \Bigg\{ \lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1,\dots,r\}: \\ \lambda_i \left(|w_i| \|\overline{x} - a^i\|_1 - \max_{i \in \{1,\dots,r\}} |w_i| \|\overline{x} - a^i\|_1 \right) = 0 \Bigg\}$$

sowie

$$\Phi_{\overline{x}} = \Bigg\{ \varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1,\dots,m\}: \\ \varphi_j \left(|w_j|(-\|\overline{x} - a^j\|_1) - \max_{j \in \{r+1,\dots,m\}} |w_j|(-\|\overline{x} - a^j\|_1) \right) = 0 \Bigg\},$$

so dass

$$0 \in \sum_{i=1}^{r} \begin{cases} \{(\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|\} & \text{für } \overline{x}_l > a^i_l \\ \{-(\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|\} & \text{für } \overline{x}_l < a^i_l \\ [-(\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|, (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|] & \text{für } \overline{x}_l = a^i_l \end{cases} \\ + \sum_{j=r+1}^{m} \begin{cases} \{-(\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|\} & \text{für } \overline{x}_l > a^j_l \\ \{(\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|\} & \text{für } \overline{x}_l < a^j_l \\ \{-(\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|, (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|\} & \text{für } \overline{x}_l = a^j_l \end{cases}$$

für alle $l \in \{1,\dots,n\}$ erfüllt ist.

Beweis: Die Aussage folgt mit $p_k = 1$ für alle $k \in \{1,\dots,m\}$ direkt aus Satz 6.1.1. ∎

Die nächsten Korollare geben notwendige Optimalitätsbedingungen für Problemstellungen an, bei denen die Abstandsfunktionen zu den existierenden Standorten a^k für alle $k \in \{1,\dots,m\}$ von ℓ_{p_k}-Normen mit $p_k \in \mathbb{R}$ und $p_k > 1$ erzeugt werden. Für den allgemeinen Fall mit dem Schema $1/\mathbb{R}^n/w_k \gtrless 0/\ell^m_{p,1<p<+\infty}/$ CD erhalten wir dabei in Abhängigkeit davon, ob den existierenden Standorten auch mehrere Abstandsfunktionen zugeordnet sein dürfen oder nicht, die folgenden beiden Resultate.

Korollar 6.2.2 (*Notw. Optimalitätsbed. für* $\mathbf{1/\mathbb{R}^n/w_k \gtrless 0/\ell^m_{p,1<p<+\infty}/}$ CD) *Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CDP) mit $\alpha \in [0,1]$, die existierenden Standorte in Indexmengen K_s mit $s \in \{1,\dots,u\}$ aufgeteilt, welche (a) - (c) genügen, sowie $d_k(x,a^k) = \|x-a^k\|_{p_k}$ mit $p_k \in \mathbb{R}$ und $p_k > 1$ für alle $k \in \{1,\dots,m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit*

$$\Lambda_{\overline{x}} = \Bigg\{\lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^r \lambda_i = 1, \forall\, i \in \{1,\dots,r\}: \\ \lambda_i\left(|w_i|\|\overline{x}-a^i\|_{p_i} - \max_{i\in\{1,\dots,r\}} |w_i|\|\overline{x}-a^i\|_{p_i}\right) = 0\Bigg\}$$

sowie

$$\Phi_{\overline{x}} = \Bigg\{\varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^m \varphi_j = 1, \forall\, j \in \{r+1,\dots,m\}: \\ \varphi_j\left(|w_j|(-\|\overline{x}-a^j\|_{p_j}) - \max_{j\in\{r+1,\dots,m\}} |w_j|(-\|\overline{x}-a^j\|_{p_j})\right) = 0\Bigg\}$$

und es ist entweder $\overline{x} = a^g$ für ein $g \in K_s$ mit $s \in \{1,\dots,u\}$, so dass

$$\sum_{\substack{j=r+1\\ j\notin K_s}}^m \frac{(\alpha+(1-\alpha)\overline{\varphi}_j)|w_j|}{\|a^g-a^j\|_{p_j}^{p_j-1}} \left((a_1^g-a_1^j)|a_1^g-a_1^j|^{p_j-2},\dots,(a_n^g-a_n^j)|a_n^g-a_n^j|^{p_j-2}\right)$$
$$- \sum_{\substack{i=1\\ i\notin K_s}}^r \frac{(\alpha+(1-\alpha)\overline{\lambda}_i)|w_i|}{\|a^g-a^i\|_{p_i}^{p_i-1}} \left((a_1^g-a_1^i)|a_1^g-a_1^i|^{p_i-2},\dots,(a_n^g-a_n^i)|a_n^g-a_n^i|^{p_i-2}\right)$$
$$\in \sum_{\substack{i=1\\ i\in K_s}}^r \left\{x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_i}{p_i-1}} \le (\alpha+(1-\alpha)\overline{\lambda}_i)|w_i|\right\}$$
$$+ \sum_{\substack{j=r+1\\ j\in K_s}}^m \left\{x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_j}{p_j-1}} = (\alpha+(1-\alpha)\overline{\varphi}_j)|w_j|\right\}$$

gilt oder es ist $\overline{x} \notin \{a^1,\dots,a^m\}$ und $\overline{x}$ löst das System nichtlinearer Gleichungen

$$0 = \sum_{i=1}^r (\alpha+(1-\alpha)\overline{\lambda}_i)|w_i| \frac{\operatorname{sgn}(\overline{x}_l-a_l^i)|\overline{x}_l-a_l^i|^{p_i-1}}{\|\overline{x}-a^i\|_{p_i}^{p_i-1}} \\ - \sum_{j=r+1}^m (\alpha+(1-\alpha)\overline{\varphi}_j)|w_j| \frac{\operatorname{sgn}(\overline{x}_l-a_l^j)|\overline{x}_l-a_l^j|^{p_j-1}}{\|\overline{x}-a^j\|_{p_j}^{p_j-1}}$$

für alle $l \in \{1,\dots,n\}$.

Beweis: Dieses Ergebnis ist Satz 6.1.1 mit $p_k \neq 1$ und $p_k \neq \infty$ für alle $k \in \{1,\dots,m\}$. ■

Korollar 6.2.3

Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CDP) mit $\alpha \in [0,1]$, paarweise verschiedenen existierenden Standorten $a^1, \ldots, a^m$ sowie $d_k(x, a^k) = \|x - a^k\|_{p_k}$ mit $p_k \in \mathbb{R}$ und $p_k > 1$ für alle $k \in \{1, \ldots, m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit

$$\Lambda_{\overline{x}} = \Bigg\{ \lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1, \ldots, r\} : \\ \lambda_i \left(|w_i| \|\overline{x} - a^i\|_{p_i} - \max_{i \in \{1,\ldots,r\}} |w_i| \|\overline{x} - a^i\|_{p_i} \right) = 0 \Bigg\}$$

sowie

$$\Phi_{\overline{x}} = \Bigg\{ \varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1, \ldots, m\} : \\ \varphi_j \left(|w_j|(-\|\overline{x} - a^j\|_{p_j}) - \max_{j \in \{r+1,\ldots,m\}} |w_j|(-\|\overline{x} - a^j\|_{p_j}) \right) = 0 \Bigg\}$$

und es ist entweder $\overline{x} = a^k$ für ein $k \in \{1, \ldots, r\}$, so dass

$$\Bigg\| \sum_{j=r+1}^{m} \left((\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| \frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p_j - 1}}{\|a^k - a^j\|_{p_j}^{p_j - 1}}, \ldots, \right. \\ \left. (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p_j - 1}}{\|a^k - a^j\|_{p_j}^{p_j - 1}} \right) \\ - \sum_{\substack{i=1 \\ i \neq k}}^{r} \left((\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p_i - 1}}{\|a^k - a^i\|_{p_i}^{p_i - 1}}, \ldots, \right. \\ \left. (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p_i - 1}}{\|a^k - a^i\|_{p_i}^{p_i - 1}} \right) \Bigg\|_{\frac{p_k}{p_k - 1}} \leq (\alpha + (1-\alpha)\overline{\lambda}_k)|w_k|$$

gilt oder es ist $\overline{x} = a^k$ für ein $k \in \{r+1, \ldots, m\}$ und es gilt

$$\Bigg\| \sum_{\substack{j=r+1 \\ j \neq k}}^{m} \left((\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| \frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p_j - 1}}{\|a^k - a^j\|_{p_j}^{p_j - 1}}, \ldots, \right. \\ \left. (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p_j - 1}}{\|a^k - a^j\|_{p_j}^{p_j - 1}} \right) \\ - \sum_{i=1}^{r} \left((\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p_i - 1}}{\|a^k - a^i\|_{p_i}^{p_i - 1}}, \ldots, \right. \\ \left. (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p_i - 1}}{\|a^k - a^i\|_{p_i}^{p_i - 1}} \right) \Bigg\|_{\frac{p_k}{p_k - 1}} = (\alpha + (1-\alpha)\overline{\varphi}_k)|w_k|$$

oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ *und* $\overline{x}$ *löst das System nichtlinearer Gleichungen*

$$\begin{aligned}0 = &\sum_{i=1}^{r}(\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|\frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}}\\ &- \sum_{j=r+1}^{m}(\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|\frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}\end{aligned}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis: Die Aussage erhält man mit $|K_s| = 1$ für alle $s \in \{1, \ldots, u\}$ direkt aus Korollar 6.2.2. ■

Für praktische Anwendungen ist der Fall einheitlicher Abstandsfunktionen einmal mehr von besonderem Interesse. Werden diese Abstandsfunktionen alle von derselben ℓ_p-Norm mit $p \in \mathbb{R}$ und $p > 1$ induziert, so erhalten wir die folgende notwendige Optimalitätsbedingung:

Korollar 6.2.4 (*Notw. Optimalitätsbed. für* $\mathbf{1/\mathbb{R}^n/w_k \gtrless 0/\ell_{p,1<p<+\infty}/}$ CD) *Ist* $\overline{x} \in \mathbb{R}^n$ *ein lokales Minimum des Problems (CDP) mit* $\alpha \in [0,1]$ *sowie* $d_k(x, a^k) = \|x - a^k\|_p$ *für ein* $p \in \mathbb{R}$ *mit* $p > 1$ *und für alle* $k \in \{1, \ldots, m\}$, *dann existieren* $\overline{\lambda} \in \Lambda_{\overline{x}}$ *und* $\overline{\varphi} \in \Phi_{\overline{x}}$ *mit*

$$\begin{aligned}\Lambda_{\overline{x}} = \Bigg\{&\lambda \in \mathbb{R}_+^r \mid \sum_{i=1}^{r}\lambda_i = 1, \forall\, i \in \{1, \ldots, r\}:\\ &\lambda_i\left(|w_i|\|\overline{x} - a^i\|_p - \max_{i\in\{1,\ldots,r\}}|w_i|\|\overline{x} - a^i\|_p\right) = 0\Bigg\}\end{aligned}$$

sowie

$$\begin{aligned}\Phi_{\overline{x}} = \Bigg\{&\varphi \in \mathbb{R}_+^{m-r} \mid \sum_{j=r+1}^{m}\varphi_j = 1, \forall\, j \in \{r+1, \ldots, m\}:\\ &\varphi_j\left(|w_j|(-\|\overline{x} - a^j\|_p) - \max_{j\in\{r+1,\ldots,m\}}|w_j|(-\|\overline{x} - a^j\|_p)\right) = 0\Bigg\}\end{aligned}$$

und es ist entweder $\overline{x} = a^k$ *für ein* $k \in \{1, \ldots, r\}$, *so dass*

$$\begin{aligned}\Bigg\|&\sum_{j=r+1}^{m}\Bigg((\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}, \ldots,\\ &\qquad(\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|\frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}\Bigg)\\ &- \sum_{\substack{i=1\\ i\neq k}}^{r}\Bigg((\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \ldots,\\ &\qquad(\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|\frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}\Bigg)\Bigg\|_{\frac{p}{p-1}} \le (\alpha + (1-\alpha)\overline{\lambda}_k)|w_k|\end{aligned}$$

gilt oder es ist $\overline{x} = a^k$ *für ein* $k \in \{r+1,\dots,m\}$ *und es gilt*

$$\left\| \sum_{\substack{j=r+1\\ j\neq k}}^{m} \left((\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| \frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}}, \dots, (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_p^{p-1}} \right) - \sum_{i=1}^{r} \left((\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}}, \dots, (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_p^{p-1}} \right) \right\|_{\frac{p}{p-1}} = (\alpha + (1-\alpha)\overline{\varphi}_k)|w_k|$$

oder es ist $\overline{x} \notin \{a^1,\dots,a^m\}$ *und* $\overline{x}$ *löst das System nichtlinearer Gleichungen*

$$0 = \sum_{i=1}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p-1}}{\|\overline{x} - a^i\|_p^{p-1}} - \sum_{j=r+1}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p-1}}{\|\overline{x} - a^j\|_p^{p-1}}$$

für alle $l \in \{1,\dots,n\}$.

Beweis: Durch Setzen von $p_k = p$ für alle $k \in \{1,\dots,m\}$ in Korollar 6.2.3 erhält man sofort die Aussage. ■

Unter den ℓ_p-Normen mit $p \in \mathbb{R}$ und $p > 1$ ist die ℓ_2-Norm die mit Abstand wichtigste Norm zur Abstandsmessung. Wir wollen daher als nächstes Resultat eine notwendige Optimalitätsbedingung für das Centdianproblem angeben, wenn bei diesem alle Metriken von der ℓ_2-Norm erzeugt werden.

Korollar 6.2.5 (*Notw. Optimalitätsbedingung für* $1/\mathbb{R}^n/w_k \gtrless 0/\ell_2/$ CD)

Ist $\overline{x} \in \mathbb{R}^n$ *ein lokales Minimum des Problems (CDP) mit* $\alpha \in [0,1]$ *sowie* $d_k(x,a^k) = \|x - a^k\|_2$ *für alle* $k \in \{1,\dots,m\}$*, dann existieren* $\overline{\lambda} \in \Lambda_{\overline{x}}$ *und* $\overline{\varphi} \in \Phi_{\overline{x}}$ *mit*

$$\Lambda_{\overline{x}} = \left\{ \lambda \in \mathbb{R}_+^r \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1,\dots,r\} : \lambda_i \left(|w_i|\|\overline{x} - a^i\|_2 - \max_{i\in\{1,\dots,r\}} |w_i|\|\overline{x} - a^i\|_2 \right) = 0 \right\}$$

sowie

$$\Phi_{\overline{x}} = \left\{ \varphi \in \mathbb{R}_+^{m-r} \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1,\dots,m\} : \varphi_j \left(|w_j|(-\|\overline{x} - a^j\|_2) - \max_{j\in\{r+1,\dots,m\}} |w_j|(-\|\overline{x} - a^j\|_2) \right) = 0 \right\}$$

und es ist entweder $\overline{x} = a^k$ für ein $k \in \{1,\dots,r\}$, so dass

$$\left\| \sum_{j=r+1}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| \frac{a^k - a^j}{\|a^k - a^j\|_2} - \sum_{\substack{i=1\\ i\neq k}}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \frac{a^k - a^i}{\|a^k - a^i\|_2} \right\|_2$$
$$\leq (\alpha + (1-\alpha)\overline{\lambda}_k)|w_k|$$

gilt oder es ist $\overline{x} = a^k$ für ein $k \in \{r+1,\dots,m\}$ und es gilt

$$\left\| \sum_{\substack{j=r+1\\ j\neq k}}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| \frac{a^k - a^j}{\|a^k - a^j\|_2} - \sum_{i=1}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \frac{a^k - a^i}{\|a^k - a^i\|_2} \right\|_2$$
$$= (\alpha + (1-\alpha)\overline{\varphi}_k)|w_k|$$

oder es ist $\overline{x} \notin \{a^1,\dots,a^m\}$ und $\overline{x}$ löst das System nichtlinearer Gleichungen

$$0 = \sum_{i=1}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i| \frac{\overline{x}_l - a^i_l}{\|\overline{x} - a^i\|_2} - \sum_{j=r+1}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j| \frac{\overline{x}_l - a^j_l}{\|\overline{x} - a^j\|_2}$$

für alle $l \in \{1,\dots,n\}$.

Beweis: Dieses Ergebnis ist Korollar 6.2.4 mit $p = 2$. ■

Als letztes Korollar in diesem Unterkapitel formulieren wir nachfolgend noch eine notwendige Optimalitätsbedingung für Centdianprobleme mit dem Klassifikationsschema $1/\mathbb{R}^n/w_k \gtrless 0/\ell_\infty/$ CD. Bei diesen Problemstellungen werden alle Abstandsfunktionen von der ℓ_∞-Norm induziert, es gilt also $d_k(x,a^k) = d(x,a^k) = \|x - a^k\|_\infty$ für alle $k \in \{1,\dots,m\}$.

Korollar 6.2.6 (*Notw. Optimalitätsbedingung für* $1/\mathbb{R}^n/w_k \gtrless 0/\ell_\infty/$ CD)
Ist $\overline{x} \in \mathbb{R}^n$ ein lokales Minimum des Problems (CDP) mit $\alpha \in [0,1]$ sowie $d_k(x,a^k) = \|x - a^k\|_\infty$ für alle $k \in \{1,\dots,m\}$, dann existieren $\overline{\lambda} \in \Lambda_{\overline{x}}$ und $\overline{\varphi} \in \Phi_{\overline{x}}$ mit

$$\Lambda_{\overline{x}} = \Bigg\{ \lambda \in \mathbb{R}^r_+ \mid \sum_{i=1}^{r} \lambda_i = 1, \forall\, i \in \{1,\dots,r\}:$$
$$\lambda_i \left(|w_i|\|\overline{x} - a^i\|_\infty - \max_{i\in\{1,\dots,r\}} |w_i|\|\overline{x} - a^i\|_\infty \right) = 0 \Bigg\}$$

sowie

$$\Phi_{\overline{x}} = \Bigg\{ \varphi \in \mathbb{R}^{m-r}_+ \mid \sum_{j=r+1}^{m} \varphi_j = 1, \forall\, j \in \{r+1,\dots,m\}:$$
$$\varphi_j \left(|w_j|(-\|\overline{x} - a^j\|_\infty) - \max_{j\in\{r+1,\dots,m\}} |w_j|(-\|\overline{x} - a^j\|_\infty) \right) = 0 \Bigg\}$$

und es ist entweder $\overline{x} = a^k$ für ein $k \in \{1,\dots,r\}$ und es existieren $x^{*_i} \in \mathbb{R}^n$ für *alle $i \in \{1, \ldots, r\} \setminus \{k\}$ sowie $x^{*_j} \in \mathbb{R}^n$ für alle $j \in \{r+1, \ldots, m\}$, welche*

$$\begin{aligned}
&(1) \quad \left\| \sum_{\substack{i=1\\ i\neq k}}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|x^{*_i} + \sum_{j=r+1}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|x^{*_j} \right\|_1 \\
&\qquad \leq (\alpha + (1-\alpha)\overline{\lambda}_k)|w_k|, \\
&(2) \quad x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^i| = 0, \\
&(3) \quad x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^j| = 0, \\
&(4) \quad \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1
\end{aligned}$$

für alle $i \in \{1, \ldots, r\} \setminus \{k\}$, für alle $j \in \{r+1, \ldots, m\}$ und für alle $l \in \{1, \ldots, n\}$ erfüllen oder es ist $\overline{x} = a^k$ für ein $k \in \{r+1, \ldots, m\}$ und es existieren $x^{*_i} \in \mathbb{R}^n$ für alle $i \in \{1, \ldots, r\}$ sowie $x^{*_j} \in \mathbb{R}^n$ für alle $j \in \{r+1, \ldots, m\} \setminus \{k\}$, welche

$$\begin{aligned}
&(1) \quad \left\| \sum_{i=1}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|x^{*_i} + \sum_{\substack{j=r+1\\ j\neq k}}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|x^{*_j} \right\|_1 \\
&\qquad = \left\| \sum_{i=1}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|x^{*_i} + \sum_{\substack{j=r+1\\ j\neq k}}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|x^{*_j} \right\|_\infty \\
&\qquad = (\alpha + (1-\alpha)\overline{\varphi}_k)|w_k|, \\
&(2) \quad x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^i| = 0, \\
&(3) \quad x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^j| = 0, \\
&(4) \quad \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1
\end{aligned}$$

für alle $i \in \{1, \ldots, r\}$, für alle $j \in \{r+1, \ldots, m\} \setminus \{k\}$ und für alle $l \in \{1, \ldots, n\}$ erfüllen oder es ist schließlich $\overline{x} \notin \{a^1, \ldots, a^m\}$ und es existieren $x^{*_i} \in \mathbb{R}^n$ für alle $i \in \{1, \ldots, r\}$ sowie $x^{*_j} \in \mathbb{R}^n$ für alle $j \in \{r+1, \ldots, m\}$, welche

$$\begin{aligned}
&(1) \quad \sum_{i=1}^{r} (\alpha + (1-\alpha)\overline{\lambda}_i)|w_i|x^{*_i} + \sum_{j=r+1}^{m} (\alpha + (1-\alpha)\overline{\varphi}_j)|w_j|x^{*_j} = 0, \\
&(2) \quad x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^i| = 0, \\
&(3) \quad x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max_{l \in \{1,\ldots,n\}} |a_l^k - a_l^j| = 0, \\
&(4) \quad \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1
\end{aligned}$$

für alle $i \in \{1, \ldots, r\}$, für alle $j \in \{r+1, \ldots, m\}$ und für alle $l \in \{1, \ldots, n\}$ erfüllen.

Beweis: Die Aussage folgt mit $p_k = \infty$ für alle $k \in \{1, \ldots, m\}$ direkt aus Satz 6.1.1. ■

7 Notwendige Optimalitätsbedingungen für nichtkonvexe mehrkriterielle Standortoptimierungsprobleme

Neben den skalaren Problemstellungen, für welche wir in den Kapiteln 4, 5 und 6 bereits eine Vielzahl an notwendigen Optimalitätsbedingungen unter verschiedenen Voraussetzungen hergeleitet haben, hatten wir in Unterkapitel 1.2 mit der mehrkriteriellen Formulierung (MP) noch ein weiteres Modell eingeführt, welches sich erheblich von den zuvor betrachteten Modellierungen unterscheidet. Bei der mehrkriteriellen Formulierung studieren wir im Folgenden eine vektorwertige Zielfunktion und suchen nach neuen Standorten, welche für diese Zielfunktion bezüglich einer gegebenen Ordnungsmenge A das Kriterium der schwachen Minimalität aus Definition 2.1.21 erfüllen.

Auch in diesem Kapitel wollen wir zunächst eine notwendige Optimalitätsbedingung für die allgemein formulierte Aufgabenstellung beweisen und aus dieser dann stärkere Resultate für Aufgabenstellungen mit spezifischen Abstandsfunktionen ableiten. Gerade für mehrkriterielle bzw. vektorielle Optimierungsprobleme stellen solche notwendigen Optimalitätsbedingungen oder Multiplikatorenregeln einen besonderen Schwerpunkt der aktuellen Forschung dar, wie beispielsweise bei Khan und Raciti (2003) oder Khan und Sama (2012). Alle Aussagen in diesem Kapitel sind wieder neue Ergebnisse der Promotion.

7.1 Notwendige Optimalitätsbedingungen für mehrkriterielle Standortoptimierungsprobleme mit gemischten Minkowski-Metriken

Bevor wir eine allgemeine notwendige Optimalitätsbedingung für das Problem (MP) präsentieren können, müssen wir zuerst auf einen wichtigen Aspekt bei der Modellierung der Aufgabenstellung eingehen. In Sektion 1.2.5 hatten wir bereits erwähnt, dass im Falle von mehrkriteriellen Standortoptimierungsproblemen die Gewichtung der existierenden Standorte abseits von positiv und negativ nicht notwendig ist, ganz im Gegenteil wird bei der Lösung des mehrkriteriellen Standortproblems mittels Skalarisierung eine entsprechende Gewichtung sogar mitbestimmt. Gerade diese Eigenschaft macht die mehrkriterielle Formulierung für praktische Anwendungen sehr nützlich, weil vorab kein Gewichtungsvektor bekannt sein muss.

Da in der Standortoptimierung die Problemstellungen auch fast immer durch praktische Anwendungen motiviert sind, wird für Standortoptimierungsprobleme zumeist der gewöhnliche Ordnungskegel $\mathbb{R}^n_+$ als Ordnungsmenge verwendet. Bezüglich dieses Kegels haben Gewichte überhaupt keinen Einfluss auf die Lösungsmenge und mit dem folgenden Lemma zeigen wir, warum mehrkriteriellen Standortoptimierungsproblemen diese Eigenschaft innewohnt.

Ehrgott (2005) beweist eine etwas stärkere Aussage mittels der Betrachtung von Niveaulinien, wir werden den Beweis hier allerdings elementar und anschaulich über die Verknüpfung des gewöhnlichen Ordnungskegels mit den gewohnten Konzepten „besser“ und „schlechter“ führen. Die verwendeten Benennungen entsprechen dabei alle den in Sektion 1.2.5 eingeführten Bezeichnungen.

Lemma 7.1.1

Für jeden Gewichtsvektor $(w_1, \dots, w_m) \in \mathbb{R}^m$ mit $w_i > 0$ für alle $i \in \{1, \dots, r\}$ und $w_j < 0$ für alle $j \in \{r+1, \dots, m\}$ ist $\overline{x} \in \mathbb{R}^n$ über $\mathbb{R}^n$ genau dann ein schwaches Minimum der gewichteten Funktion $f_w \colon \mathbb{R}^n \to \mathbb{R}^m$ mit

$$f_w(x) = \begin{pmatrix} w_1 d_1(x, a^1) \\ \vdots \\ w_r d_r(x, a^r) \\ w_{r+1} d_{r+1}(x, a^{r+1}) \\ \vdots \\ w_m d_m(x, a^m) \end{pmatrix}$$

bezüglich des gewöhnlichen Ordnungskegels $\mathbb{R}^m_+$, wenn $\overline{x}$ über $\mathbb{R}^n$ ein schwaches Minimum von $f \colon \mathbb{R}^n \to \mathbb{R}^m$ mit

$$f(x) = \begin{pmatrix} d_1(x, a^1) \\ \vdots \\ d_r(x, a^r) \\ -d_{r+1}(x, a^{r+1}) \\ \vdots \\ -d_m(x, a^m) \end{pmatrix}$$

bezüglich $\mathbb{R}^m_+$ ist.

Beweis: Ist $\overline{x} \in \mathbb{R}^n$ über $\mathbb{R}^n$ ein schwaches Minimum von f bezüglich $\mathbb{R}^m_+$, so ist $\overline{y} = f(\overline{x})$ ein schwach minimales Element von $f(\mathbb{R}^n)$ bezüglich $\mathbb{R}^m_+$. Das heißt, es existiert kein Punkt $y_0 \in f(\mathbb{R}^n)$ mit $\overline{y} \in y_0 + \operatorname{int} \mathbb{R}^m_+$, also gibt es für jedes $y \in f(\mathbb{R}^n)$ immer mindestens ein $k \in \{1, \dots, m\}$, so dass $\overline{y}_k \leq y_k$ gilt.

Damit gilt auch $|w_k|\overline{y}_k \leq |w_k| y_k$ für alle $y \in f(\mathbb{R}^n)$, mindestens ein $k \in \{1, \dots, m\}$ und jeden beliebigen Gewichtsvektor $w = (w_1, \dots, w_m) \in \mathbb{R}^m$. Demzufolge kann kein Element $(|w_1|, \dots, |w_m|) y_0 \in (|w_1|, \dots, |w_m|) f(\mathbb{R}^n)$ existieren, für welches $(|w_1|, \dots, |w_m|)\overline{y} \in (|w_1|, \dots, |w_m|) y_0 + \operatorname{int} \mathbb{R}^m_+$ gilt.

Wegen $f_w(x) = (|w_1|, \dots, |w_m|) f(x)$ für alle $x \in \mathbb{R}^n$ liefert das aber gerade, dass $(|w_1|, \dots, |w_m|)\overline{y} = f_w(\overline{x})$ ein schwach minimales Element von $f_w(\mathbb{R}^n)$ bezüglich $\mathbb{R}^m_+$ ist bzw. dass $\overline{x}$ über $\mathbb{R}^n$ ein schwaches Minimum von f_w bezüglich $\mathbb{R}^m_+$ ist.

Die Rückrichtung lässt sich völlig analog beweisen, indem man den Gewichtsvektor $\left(\frac{1}{|w_1|}, \dots, \frac{1}{|w_m|}\right)$ betrachtet. ■

Wir wollen nun zwei Resultate zu notwendigen Optimalitätsbedingungen des allgemeinen mehrkriteriellen Problems (MP) angeben, bei welchem die Abstandsfunktionen für alle $k \in \{1, \dots, m\}$ Minkowski-Metriken zu beliebigen ℓ_{p_k}-Normen mit $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ sein sollen. Diese Aufgabenstellung hat in der Klassifikation der Standortoptimierungsprobleme das Schema $1/\mathbb{R}^n/w_k = \pm 1/\ell_p^m/m - \sum_{wpar}$. Wir unterscheiden für die folgenden Ergebnisse aber noch, bezüglich welcher Ordnungsmenge wir die schwachen Minima des Problems bestimmen.

Zuerst nutzen wir dafür eine Ordnungsmenge A, an die wir möglichst wenige Voraussetzungen stellen, und betrachten eine Skalarisierung des mehrkriteriellen Problems mithilfe des in (2.1) definierten nichtlinearen Skalarisierungsfunktionals $\varphi_{A,k}$ aus Sektion 2.1.4, um schwach minimale Elemente von $f(\mathbb{R}^n)$ bezüglich A zu bestimmen. Dieses Resultat ist eine Verallgemeinerung von Theorem 5.6 aus Bao, Hillmann und Tammer (2017).

Satz 7.1.1 (*Notw. Optimalitätsbed. für* $1/\mathbb{R}^n/w_k = \pm 1/\ell_p^m/m - \sum_{wpar}$)

Es sei $A \subset \mathbb{R}^m$ eine abgeschlossene Menge mit nichtleerem algebraischen Inneren, $0 \in \operatorname{bd} A$ und $k \in \mathbb{R}^m \setminus \{0\}$ ein Richtungsvektor, so dass (2.3) erfüllt ist, also $A + (0, +\infty) \cdot k \subseteq \operatorname{int} A$ gilt. Weiter sei $f(\overline{x}) \in \operatorname{WMin}(f(\mathbb{R}^n), A)$, die Funktion $\Phi_{A,k}$ sei für alle $y \in \mathbb{R}^m$ definiert mittels $\Phi_{A,k}(y) := \varphi_{A,k}(y - f(\overline{x}))$ und lokal Lipschitz-stetig in $f(\overline{x})$.

Dann existiert ein $y^* \in \mathcal{A} := \{y^* \in \mathbb{R}^m \mid y^*(k) = 1, -y^* \in N_L(\varphi_{A,k}(0)k, \operatorname{bd} A)\}$ mit

$$0 \in \partial_L \langle y^*, f \rangle (\overline{x}).$$

Beweis: Theorem 6.2.31 in Tammer und Weidner (2020) besagt, dass genau dann $f(\overline{x}) \in \operatorname{WMin}(f(\mathbb{R}^n), A)$ gilt, wenn $\overline{x}$ ein (lokales) Minimum des skalarisierten Problems

$$\Phi_{A,k} \circ f(x) \to \min_{x \in \mathbb{R}^n}$$

ist. Eigenschaft (L4) aus Lemma 2.2.2 liefert dann die notwendige Optimalitätsbedingung $0 \in \partial_L (\Phi_{A,k} \circ f)(\overline{x})$.

f ist eine Funktion, die komponentenweise nur aus Normfunktionen besteht, und damit lokal Lipschitz-stetig in $\overline{x}$ und $\Phi_{A,k}$ ist nach Voraussetzung lokal Lipschitz-stetig in $f(\overline{x})$, also können wir die Kettenregel für Limiting-Subdifferentiale (Satz 2.2.9) anwenden und erhalten

$$0 \in \partial_L (\Phi_{A,k} \circ f)(\overline{x}) \subseteq \bigcup_{y^* \in \partial_L \Phi_{A,k}(f(\overline{x}))} \partial_L \langle y^*, f \rangle (\overline{x}).$$

Das Limiting-Subdifferential des nichtlinearen Skalarisierungsfunktionals $\varphi_{A,k}$ haben wir bereits in Satz 3.3.2 bestimmt, damit ist $\partial_L \Phi_{A,k}(f(\overline{x})) = \partial_L \varphi_{A,k}(0) = \mathcal{A}$ und die Aussage des Satzes gezeigt. ■

Bemerkung 7.1.1

Um im Beweis von Satz 7.1.1 die Kettenregel für Limiting-Subdifferentiale aus Satz 2.2.9 anwenden zu können, müssen wir die lokale Lipschitz-Stetigkeit von $\Phi_{A,k}$ in $f(\overline{x})$ fordern. Diese Eigenschaft ist z. B. stets gegeben, wenn A zusätzlich zu den Voraussetzungen von Satz 7.1.1 noch konvex ist.

Wählen wir als Ordnungsmenge $A \subset \mathbb{R}^m$ den gewöhnlichen Ordnungskegel $\mathbb{R}^m_+$, so lässt sich auch die notwendige Optimalitätsbedingung für das allgemeine mehrkriterielle Standortoptimierungsproblem mit beliebigen, für alle $k \in \{1, \dots, m\}$ von ℓ_{p_k}-Normen mit $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ induzierten Abstandsfunktionen auf die notwendige Optimalitätsbedingung eines zugehörigen Medianproblems aus Satz 4.2.1 zurückführen.

Mithilfe unseres Ansatzes können wir also nicht nur die notwendigen Optimalitätsbedingungen für die in den Kapiteln 4, 5 und 6 betrachteten skalaren Zielfunktionen vereinheitlichen, sondern auch der vektorielle Ansatz fügt sich nahtlos in die entwickelte Theorie ein.

Um die notwendige Optimalitätsbedingung aus Satz 4.2.1 direkt anwenden zu können, ohne die existierenden Standorte nochmals umsortieren zu müssen, fordern wir auch hier, dass die existierenden Standorte vorab nach ihren zugeordneten Abstandsfunktionen sortiert und deren Indizes in Indexmengen K_s mit $s \in \{1, \dots, u\}$ zusammengefasst seien, wie wir dies in Unterkapitel 4.2 eingeführt haben.

Satz 7.1.2 (*Notw. Optimalitätsbed. für* $\boldsymbol{1/\mathbb{R}^n/w_k = \pm 1/\ell_p^m/m - \sum_{wpar}}$)

Ist $f(\overline{x}) \in \mathrm{WMin}(f(\mathbb{R}^n), \mathbb{R}^m_+)$ ein schwach minimales Element des Problems (MP) bezüglich $\mathbb{R}^m_+$, $k \in \operatorname{int} \mathbb{R}^m_+$ ein Richtungsvektor, die existierenden Standorte in Indexmengen K_s mit $s \in \{1, \dots, u\}$ aufgeteilt, welche (a) - (c) genügen, sowie $d_k(x, a^k) = \|x - a^k\|_{p_k}$ mit $p_k \in \overline{\mathbb{R}}$ und $p_k \geq 1$ für alle $k \in \{1, \dots, m\}$, dann existiert ein $y^* \in \{y^* \in \mathbb{R}^m_+ \mid y^*(k) = 1\}$, so dass in $\overline{x}$ die notwendige Optimalitätsbedingung aus Satz 4.2.1 für das Medianproblem

$$\sum_{i=1}^{r} y_i^* \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} y_j^* (-\|x - a^j\|_{p_j}) \to \min_{x \in \mathbb{R}^n}$$

erfüllt ist.

Beweis: Wir wenden Satz 7.1.1 unter Nutzung der Ordnungsmenge $A = \mathbb{R}^m_+$ an. Ist $k \in \operatorname{int} \mathbb{R}^m_+$, so sind alle Voraussetzungen von Satz 7.1.1 bereits erfüllt. Die lokale Lipschitz-Stetigkeit der Funktion $\Phi_{\mathbb{R}^m_+, k}$ folgt dabei, wie in Bemerkung 7.1.1 beschrieben, aus der Konvexität des Ordnungskegels $\mathbb{R}^m_+$.

Es existiert also ein $y^* \in \{y^* \in \mathbb{R}^m \mid y^*(k) = 1, -y^* \in N_L(\varphi_{\mathbb{R}^m_+, k}(0)k, \operatorname{bd} \mathbb{R}^m_+)\}$ mit $0 \in \partial_L \langle y^*, f \rangle (\overline{x})$. Wegen $\varphi_{\mathbb{R}^m_+, k}(0) = 0$ und $N_L(0, \operatorname{bd} \mathbb{R}^m_+) = -\mathbb{R}^m_+$ gilt weiter

$$\begin{aligned}
&\{y^* \in \mathbb{R}^m \mid y^*(k) = 1, -y^* \in N_L(\varphi_{\mathbb{R}^m_+, k}(0)k, \operatorname{bd} \mathbb{R}^m_+)\} \\
= &\{y^* \in \mathbb{R}^m \mid y^*(k) = 1, -y^* \in N_L(0, \operatorname{bd} \mathbb{R}^m_+)\} \\
= &\{y^* \in \mathbb{R}^m \mid y^*(k) = 1, -y^* \in -\mathbb{R}^m_+\} \\
= &\{y^* \in \mathbb{R}^m_+ \mid y^*(k) = 1\}.
\end{aligned}$$

Die notwendige Optimalitätsbedingung

$$\begin{aligned}
0 &\in \partial_L \langle y^*, f \rangle (\overline{x}) \\
&= \partial_L \left(\sum_{i=1}^{r} y_i^* \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} y_j^* (-\|x - a^j\|_{p_j}) \right) (\overline{x})
\end{aligned}$$

entspricht genau der notwendigen Optimalitätsbedingung, die Eigenschaft (L4) aus Lemma 2.2.2 für das Medianproblem

$$\sum_{i=1}^{r} y_i^* \|x - a^i\|_{p_i} + \sum_{j=r+1}^{m} y_j^* (-\|x - a^j\|_{p_j}) \to \min_{x \in \mathbb{R}^n}$$

liefert. Damit muss dieses Medianproblem in $\overline{x}$ auch der notwendigen Optimalitätsbedingung aus Satz 4.2.1 genügen. ■

7.2 Notwendige Optimalitätsbedingungen für mehrkriterielle Standortoptimierungsprobleme mit speziellen Metriken

Auch in diesem Kapitel wollen wir das zweite Unterkapitel dafür verwenden, aus der allgemeinen notwendigen Optimalitätsbedingung des Satzes 7.1.2 weitere Resultate für speziellere Problemstellungen herzuleiten. Wir betrachten dazu im Folgenden mehrkriterielle Standortoptimierungsprobleme, bei denen die Abstandsfunktionen zu allen existierenden Standorten homogen sind.

Zuerst untersuchen wir wieder Aufgabenstellungen mit ausschließlich von der ℓ_1-Norm induzierten Metriken, bei Verwendung einer mehrkriteriellen Zielfunktion haben diese das Schema $1/\mathbb{R}^n/w_k = \pm 1/\ell_1/m - \sum_{wpar}$.

Korollar 7.2.1 (*Notw. Opt.bed. für* $1/\mathbb{R}^n/w_k = \pm 1/\ell_1/m - \sum_{wpar}$)
Ist $f(\overline{x}) \in \mathrm{WMin}(f(\mathbb{R}^n), \mathbb{R}^m_+)$ ein schwach minimales Element des Problems (MP) bezüglich $\mathbb{R}^m_+$, $k \in \operatorname{int} \mathbb{R}^m_+$ ein Richtungsvektor und $d_k(x, a^k) = \|x - a^k\|_1$ für alle $k \in \{1, \ldots, m\}$, dann existiert ein $y^* \in \{y^* \in \mathbb{R}^m_+ \mid y^*(k) = 1\}$, so dass

$$0 \in \sum_{i=1}^{r} \begin{cases} \{y_i^*\} & \text{für } \overline{x}_l > a_l^i \\ \{-y_i^*\} & \text{für } \overline{x}_l < a_l^i \\ [-y_i^*, y_i^*] & \text{für } \overline{x}_l = a_l^i \end{cases} + \sum_{j=r+1}^{m} \begin{cases} \{-y_j^*\} & \text{für } \overline{x}_l > a_l^j \\ \{y_j^*\} & \text{für } \overline{x}_l < a_l^j \\ \{-y_j^*, y_j^*\} & \text{für } \overline{x}_l = a_l^j \end{cases}$$

für alle $l \in \{1, \ldots, n\}$ erfüllt ist.

Beweis: Diese Aussage ist Satz 7.1.2 mit $p_k = 1$ für alle $k \in \{1, \ldots, m\}$. ■

Nun wollen wir zur Betrachtung solcher Aufgabenstellungen übergehen, bei denen die Abstandsfunktionen zu den existierenden Standorten für alle $k \in \{1, \ldots, m\}$ von ℓ_{p_k}-Normen mit $p_k \in \mathbb{R}$ und $p_k > 1$ induziert werden. Dabei beginnen wir mit dem allgemeineren Fall, in welchem die Abstandsfunktionen alle voneinander verschieden sein können.

Dieses Problem entspricht dem Schema $1/\mathbb{R}^n/w_k = \pm 1/\ell^m_{p,1<p<+\infty}/m - \sum_{wpar}$ in der Klassifikation der Standortoptimierungsprobleme. Die folgenden beiden Korollare unterscheiden sich jedoch nochmals darin, ob die existierenden Standorte gleich sein bzw. mehrere Zielfunktionen zugeordnet bekommen dürfen oder ob diese alle paarweise verschieden sein sollen.

Korollar 7.2.2 (*Notw. Opt.bed. für* $\mathbf{1/\mathbb{R}^n/w_k = \pm 1/\ell^m_{p,1<p<+\infty}/m - \sum_{wpar}}$)

Ist $f(\overline{x}) \in \mathrm{WMin}(f(\mathbb{R}^n), \mathbb{R}^m_+)$ ein schwach minimales Element des Problems (MP) bezüglich $\mathbb{R}^m_+$, $k \in \operatorname{int} \mathbb{R}^m_+$ ein Richtungsvektor, die existierenden Standorte in Indexmengen K_s mit $s \in \{1, \ldots, u\}$ aufgeteilt, welche (a) - (c) genügen, sowie $d_k(x, a^k) = \|x - a^k\|_{p_k}$ mit $p_k \in \mathbb{R}$ und $p_k > 1$ für alle $k \in \{1, \ldots, m\}$, dann existiert ein $y^* \in \{y^* \in \mathbb{R}^m_+ \mid y^*(k) = 1\}$ und es ist entweder $\overline{x} = a^g$ für ein $g \in K_s$ mit $s \in \{1, \ldots, u\}$, so dass

$$\sum_{\substack{j=r+1 \\ j \notin K_s}}^{m} y_j^* \left(\frac{\operatorname{sgn}(a_1^g - a_1^j)|a_1^g - a_1^j|^{p_j - 1}}{\|a^g - a^j\|_{p_j}^{p_j - 1}}, \ldots, \frac{\operatorname{sgn}(a_n^g - a_n^j)|a_n^g - a_n^j|^{p_j - 1}}{\|a^g - a^j\|_{p_j}^{p_j - 1}} \right)$$
$$- \sum_{\substack{i=1 \\ i \notin K_s}}^{r} y_i^* \left(\frac{\operatorname{sgn}(a_1^g - a_1^i)|a_1^g - a_1^i|^{p_i - 1}}{\|a^g - a^i\|_{p_i}^{p_i - 1}}, \ldots, \frac{\operatorname{sgn}(a_n^g - a_n^i)|a_n^g - a_n^i|^{p_i - 1}}{\|a^g - a^i\|_{p_i}^{p_i - 1}} \right)$$
$$\in \sum_{\substack{i=1 \\ i \in K_s}}^{r} \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_i}{p_i - 1}} \leq y_i^* \right\} + \sum_{\substack{j=r+1 \\ j \in K_s}}^{m} \left\{ x^* \in \mathbb{R}^n \mid \|x^*\|_{\frac{p_j}{p_j - 1}} = y_j^* \right\}$$

gilt oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ und $\overline{x}$ löst das System nichtlinearer Gleichungen

$$0 = \sum_{i=1}^{r} y_i^* \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p_i - 1}}{\|\overline{x} - a^i\|_{p_i}^{p_i - 1}} - \sum_{j=r+1}^{m} y_j^* \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p_j - 1}}{\|\overline{x} - a^j\|_{p_j}^{p_j - 1}}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis: Die Aussage folgt sofort, indem man in Satz 7.1.2 sowohl $p_k \neq 1$ als auch $p_k \neq \infty$ für alle $k \in \{1, \ldots, m\}$ setzt. ■

Falls alle der existierenden Standorte $a^1, \ldots, a^m$ paarweise verschieden sind, dann erhält man aus Korollar 7.2.2 die vereinfachte Formulierung:

Korollar 7.2.3

Ist $f(\overline{x}) \in \mathrm{WMin}(f(\mathbb{R}^n), \mathbb{R}^m_+)$ ein schwach minimales Element des Problems (MP) bezüglich $\mathbb{R}^m_+$, $k \in \operatorname{int} \mathbb{R}^m_+$ ein Richtungsvektor und sind die existierenden Standorte $a^1, \ldots, a^m$ paarweise verschieden mit $d_k(x, a^k) = \|x - a^k\|_{p_k}$ mit $p_k \in \mathbb{R}$ und $p_k > 1$ für alle $k \in \{1, \ldots, m\}$, dann existiert ein $y^* \in \{y^* \in \mathbb{R}^m_+ \mid y^*(k) = 1\}$ und es ist entweder $\overline{x} = a^k$ für ein $k \in \{1, \ldots, r\}$ und es gilt

$$\Bigg\| \sum_{j=r+1}^{m} y_j^* \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p_j - 1}}{\|a^k - a^j\|_{p_j}^{p_j - 1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p_j - 1}}{\|a^k - a^j\|_{p_j}^{p_j - 1}} \right)$$
$$- \sum_{\substack{i=1 \\ i \neq k}}^{r} y_i^* \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p_i - 1}}{\|a^k - a^i\|_{p_i}^{p_i - 1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p_i - 1}}{\|a^k - a^i\|_{p_i}^{p_i - 1}} \right) \Bigg\|_{\frac{p_k}{p_k - 1}} \leq y_k^*$$

oder es ist $\overline{x} = a^k$ für ein $k \in \{r + 1, \ldots, m\}$ und es gilt

$$\left\| \sum_{\substack{j=r+1\\j\neq k}}^{m} y_j^* \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}} \right) \right.$$
$$\left. - \sum_{i=1}^{r} y_i^* \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}} \right) \right\|_{\frac{p_k}{p_k-1}} = y_k^*$$

oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ *und* $\overline{x}$ *löst das System nichtlinearer Gleichungen*

$$0 = \sum_{i=1}^{r} y_i^* \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} y_j^* \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis: Die Aussage folgt mit $|K_s| = 1$ für alle $s \in \{1, \ldots, u\}$ direkt aus Korollar 7.2.2. ∎

Werden alle der Abstandsfunktionen im Problem (MP) von derselben ℓ_p-Norm mit $p \in \mathbb{R}$ und $p > 1$ induziert, so ergibt sich das folgende Resultat als wichtiger Spezialfall von Korollar 7.2.3.

Korollar 7.2.4 (*Notw. Opt.bed. für* $\mathbf{1/\mathbb{R}^n/w_k = \pm 1/\ell_{p,1<p<+\infty}/m - \sum_{wpar}}$)
Ist $f(\overline{x}) \in \operatorname{WMin}(f(\mathbb{R}^n), \mathbb{R}_+^m)$ *ein schwach minimales Element des Problems (MP) bezüglich* $\mathbb{R}_+^m$, $k \in \operatorname{int} \mathbb{R}_+^m$ *ein Richtungsvektor und gilt* $d_k(x, a^k) = \|x - a^k\|_p$ *für ein* $p \in \mathbb{R}$ *mit* $p > 1$ *sowie für alle* $k \in \{1, \ldots, m\}$, *dann existiert ein* $y^* \in \{y^* \in \mathbb{R}_+^m \mid y^*(k) = 1\}$ *und es ist entweder* $\overline{x} = a^k$ *für ein* $k \in \{1, \ldots, r\}$ *und es gilt*

$$\left\| \sum_{j=r+1}^{m} y_j^* \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_{p}^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_{p}^{p-1}} \right) \right.$$
$$\left. - \sum_{\substack{i=1\\i\neq k}}^{r} y_i^* \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_{p}^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_{p}^{p-1}} \right) \right\|_{\frac{p}{p-1}} \leq y_k^*$$

oder es ist $\overline{x} = a^k$ *für ein* $k \in \{r+1, \ldots, m\}$ *und es gilt*

$$\left\| \sum_{\substack{j=r+1\\j\neq k}}^{m} y_j^* \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p-1}}{\|a^k - a^j\|_{p}^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p-1}}{\|a^k - a^j\|_{p}^{p-1}} \right) \right.$$
$$\left. - \sum_{i=1}^{r} y_i^* \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p-1}}{\|a^k - a^i\|_{p}^{p-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p-1}}{\|a^k - a^i\|_{p}^{p-1}} \right) \right\|_{\frac{p}{p-1}} = y_k^*$$

oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ *und* $\overline{x}$ *löst das System nichtlinearer Gleichungen*

$$0 = \sum_{i=1}^{r} y_i^* \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p-1}}{\|\overline{x} - a^i\|_p^{p-1}} - \sum_{j=r+1}^{m} y_j^* \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p-1}}{\|\overline{x} - a^j\|_p^{p-1}}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis: Dieses Resultat ist Korollar 7.2.3 mit $p_k = p$ für alle $k \in \{1, \ldots, m\}$. ■

Ein weiterer wichtiger Spezialfall von Korollar 7.2.3 bzw. von Korollar 7.2.4 sind Aufgabenstellungen, die nur die ℓ_2-Norm als induzierende Funktion für alle Metriken besitzen. Diese Problemstellungen sind für den praktischen Einsatz von so großer Bedeutung, dass wir auch für diesen Fall ein separates Ergebnis angeben wollen.

Korollar 7.2.5 (*Notw. Opt.bed. für* $1/\mathbb{R}^n/w_k = \pm 1/\ell_2/m - \sum_{wpar}$)

Ist $f(\overline{x}) \in \operatorname{WMin}(f(\mathbb{R}^n), \mathbb{R}^m_+)$ *ein schwach minimales Element des Problems (MP) bezüglich* $\mathbb{R}^m_+$, $k \in \operatorname{int} \mathbb{R}^m_+$ *ein Richtungsvektor und* $d_k(x, a^k) = \|x - a^k\|_2$ *für alle* $k \in \{1, \ldots, m\}$, *dann existiert ein* $y^* \in \{y^* \in \mathbb{R}^m_+ \mid y^*(k) = 1\}$ *und es ist entweder* $\overline{x} = a^k$ *für ein* $k \in \{1, \ldots, r\}$ *und es gilt*

$$\left\| \sum_{j=r+1}^{m} y_j^* \frac{a^k - a^j}{\|a^k - a^j\|_2} - \sum_{\substack{i=1 \\ i \neq k}}^{r} y_i^* \frac{a^k - a^i}{\|a^k - a^i\|_2} \right\|_2 \leq y_k^*$$

oder es ist $\overline{x} = a^k$ *für ein* $k \in \{r+1, \ldots, m\}$ *und es gilt*

$$\left\| \sum_{\substack{j=r+1 \\ j \neq k}}^{m} y_j^* \frac{a^k - a^j}{\|a^k - a^j\|_2} - \sum_{i=1}^{r} y_i^* \frac{a^k - a^i}{\|a^k - a^i\|_2} \right\|_2 = y_k^*$$

oder es ist $\overline{x} \notin \{a^1, \ldots, a^m\}$ *und* $\overline{x}$ *löst das System nichtlinearer Gleichungen*

$$0 = \sum_{i=1}^{r} y_i^* \frac{\overline{x}_l - a_l^i}{\|\overline{x} - a^i\|_2} - \sum_{j=r+1}^{m} y_j^* \frac{\overline{x}_l - a_l^j}{\|\overline{x} - a^j\|_2}$$

für alle $l \in \{1, \ldots, n\}$.

Beweis: Die Aussage des Korollars folgt mit $p = 2$ sofort aus Korollar 7.2.4. ■

Zum Abschluss dieses Kapitels wollen wir noch ein Ergebnis für mehrkriterielle Standortprobleme mit ausschließlich von der ℓ_∞-Norm erzeugten Abstandsfunktionen, wenn also $d_k(x, a^k) = d(x, a^k) = \|x - a^k\|_\infty$ für alle $k \in \{1, \ldots, m\}$ gilt, aus Satz 7.1.2 ableiten.

Dies ist zugleich die letzte notwendige Optimalitätsbedingung in dieser Arbeit, nachdem wir nun für alle der vier Modellierungen aus Unterkapitel 1.2 eine Vielzahl solcher Optimalitätsbedingungen für die unterschiedlichsten Problemstellungen herleiten konnten.

Korollar 7.2.6 (*Notw. Opt.bed. für* $1/\mathbb{R}^n/w_k = \pm 1/\ell_\infty/m - \sum_{wpar}$)

Ist $f(\overline{x}) \in \operatorname{WMin}(f(\mathbb{R}^n), \mathbb{R}^m_+)$ ein schwach minimales Element des Problems (MP) bezüglich $\mathbb{R}^m_+$, $k \in \operatorname{int} \mathbb{R}^m_+$ ein Richtungsvektor und $d_k(x, a^k) = \|x - a^k\|_\infty$ für alle $k \in \{1, \ldots, m\}$, dann existiert ein $y^* \in \{y^* \in \mathbb{R}^m_+ \mid y^*(k) = 1\}$ und es ist entweder $\overline{x} = a^k$ für ein $k \in \{1, \ldots, r\}$ und es existieren $x^{*_h} \in \mathbb{R}^n$ für alle $h \in \{1, \ldots, m\} \setminus \{k\}$, welche

$$\begin{array}{ll}
(1) & \left\| \sum_{\substack{h=1 \\ h \neq k}}^m y_h^* x^{*_h} \right\|_1 \leq y_k^*, \\
(2) & x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max\limits_{l \in \{1,\ldots,n\}} |a_l^k - a_l^i| = 0, \\
(3) & x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max\limits_{l \in \{1,\ldots,n\}} |a_l^k - a_l^j| = 0, \\
(4) & \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1
\end{array}$$

für alle $i \in \{1, \ldots, r\} \setminus \{k\}$, für alle $j \in \{r+1, \ldots, m\}$ und für alle $l \in \{1, \ldots, n\}$ erfüllen oder es ist $\overline{x} = a^k$ für ein $k \in \{r+1, \ldots, m\}$ und es existieren $x^{*_h} \in \mathbb{R}^n$ für alle $h \in \{1, \ldots, m\} \setminus \{k\}$, welche

$$\begin{array}{ll}
(1) & \left\| \sum_{\substack{h=1 \\ h \neq k}}^m y_h^* x^{*_h} \right\|_1 = y_k^*, \quad \left\| \sum_{\substack{h=1 \\ h \neq k}}^m y_h^* x^{*_h} \right\|_\infty = y_k^*, \\
(2) & x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max\limits_{l \in \{1,\ldots,n\}} |a_l^k - a_l^i| = 0, \\
(3) & x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max\limits_{l \in \{1,\ldots,n\}} |a_l^k - a_l^j| = 0, \\
(4) & \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1
\end{array}$$

für alle $i \in \{1, \ldots, r\}$, für alle $j \in \{r+1, \ldots, m\} \setminus \{k\}$ und für alle $l \in \{1, \ldots, n\}$ erfüllen oder es ist schließlich $\overline{x} \notin \{a^1, \ldots, a^m\}$ und es existieren $x^{*_h} \in \mathbb{R}^n$ für alle $h \in \{1, \ldots, m\}$, welche

$$\begin{array}{ll}
(1) & \sum_{h=1}^m y_h^* x^{*_h} = 0, \\
(2) & x_l^{*_i}(a_l^k - a_l^i) - |x_l^{*_i}| \max\limits_{l \in \{1,\ldots,n\}} |a_l^k - a_l^i| = 0, \\
(3) & x_l^{*_j}(a_l^k - a_l^j) + |x_l^{*_j}| \max\limits_{l \in \{1,\ldots,n\}} |a_l^k - a_l^j| = 0, \\
(4) & \|x^{*_i}\|_1 = 1, \|x^{*_j}\|_1 = 1 \text{ und } \|x^{*_j}\|_\infty = 1
\end{array}$$

für alle $i \in \{1, \ldots, r\}$, für alle $j \in \{r+1, \ldots, m\}$ und für alle $l \in \{1, \ldots, n\}$ erfüllen.

Beweis: Dieses Ergebnis erhalten wir direkt aus Satz 7.1.2, indem wir $p_k = \infty$ für alle $k \in \{1, \ldots, m\}$ einsetzen. ■

8 Algorithmen

Wir wollen diese Arbeit mit einem kurzen Einblick in die Möglichkeiten zur praktischen Anwendung der erzielten theoretischen Resultate für die Herleitung von Algorithmen abschließen. Aus den verschiedenen notwendigen Optimalitätsbedingungen lassen sich Algorithmen entwickeln, welche einzelne oder sogar alle Elemente in $\mathbb{R}^n$ liefern, die die jeweilige notwendige Optimalitätsbedingung erfüllen. Auch ohne hinreichendes Kriterium schränken solche Algorithmen die Menge der als Lösung infrage kommenden Elemente erheblich ein oder führen sogar direkt zu Lösungen.

In seiner Masterarbeit hat Hillmann (2013) bereits einen Algorithmus zur Lösung des teil-abstoßenden Medianproblems mit der ℓ_1-Norm als induzierender Norm aller Abstandsfunktionen entwickelt. Dieser Algorithmus basiert auf der notwendigen Optimalitätsbedingung für das Problem $1/\mathbb{R}^n/w_k \gtrless 0/\ell_1/\Sigma$, die wir in Satz 4.1.1 vorgestellt haben.

Da in $\mathbb{R}^2$ die ℓ_1-Norm und die ℓ_∞-Norm äquivalent sind und für alle $x = (x_1, x_2) \in \mathbb{R}^2$ mittels der linearen Transformation $\|(x_1, x_2)\|_\infty = \left\|\frac{1}{2}(x_1 + x_2, -x_1 + x_2)\right\|_1$ ineinander überführt werden können, löst der Algorithmus in $\mathbb{R}^2$ außerdem auch solche teil-abstoßenden Medianprobleme, bei denen alle Abstandsfunktionen von der ℓ_∞-Norm induziert werden.

Wir wollen nun einen Algorithmus für Aufgabenstellungen mit dem Klassifikationsschema $1/\mathbb{R}^n/w_k \gtrless 0/\ell^m_{p,1<p<+\infty}/\Sigma$ entwickeln, bei denen für alle $k \in \{1, \ldots, m\}$ der Abstand zwischen dem existierenden Standort a^k und dem neuen Standort x durch die Minkowski-Metrik zu einer ℓ_{p_k}-Norm mit $p_k \in \mathbb{R}$ und $p_k > 1$ gemessen wird. Damit schließen wir die bestehende Lücke zwischen den beiden bereits von Hillmann (2013) hergeleiteten Algorithmen. Die Minkowski-Metriken können dabei für alle existierenden Standorte unterschiedlich sein, wir wollen allerdings voraussetzen, dass auch alle der existierenden Standorte paarweise verschieden sind.

Der Algorithmus wird die notwendige Optimalitätsbedingung aus Satz 4.1.2 bzw. aus der anschließenden Bemerkung 4.1.2 nutzen, um iterativ Elemente zu generieren, welche diese notwendige Optimalitätsbedingung näherungsweise immer besser erfüllen. Zur Berechnung dieser Elemente formen wir die notwendige Optimalitätsbedingung für den Fall $\overline{x} \notin \{a^1, \ldots, a^m\}$ aus Bemerkung 4.1.2 um. Die Formel lässt sich zwar nicht geschlossen nach x auflösen, allerdings erhalten wir so trotzdem eine geeignete Iterationsvorschrift.

Für alle $l \in \{1, \ldots, n\}$ muss dabei

$$\begin{aligned} 0 &= \sum_{i=1}^{r} |w_i| \frac{\operatorname{sgn}(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p_i-1}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} |w_j| \frac{\operatorname{sgn}(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p_j-1}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}} \\ &= \sum_{i=1}^{r} |w_i| \frac{(\overline{x}_l - a_l^i)|\overline{x}_l - a_l^i|^{p_i-2}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} |w_j| \frac{(\overline{x}_l - a_l^j)|\overline{x}_l - a_l^j|^{p_j-2}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{r} \overline{x}_l \frac{|w_i||\overline{x}_l - a_l^i|^{p_i-2}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{i=1}^{r} a_l^i \frac{|w_i||\overline{x}_l - a_l^i|^{p_i-2}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} \\
&- \sum_{j=r+1}^{m} \overline{x}_l \frac{|w_j||\overline{x}_l - a_l^j|^{p_j-2}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}} + \sum_{j=r+1}^{m} a_l^j \frac{|w_j||\overline{x}_l - a_l^j|^{p_j-2}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}
\end{aligned}
$$

gelten, dies ist äquivalent zu

$$
\begin{aligned}
&\overline{x}_l \left(\sum_{i=1}^{r} \frac{|w_i||\overline{x}_l - a_l^i|^{p_i-2}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} \frac{|w_j||\overline{x}_l - a_l^j|^{p_j-2}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}} \right) \\
&= \sum_{i=1}^{r} a_l^i \frac{|w_i||\overline{x}_l - a_l^i|^{p_i-2}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} a_l^j \frac{|w_j||\overline{x}_l - a_l^j|^{p_j-2}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}
\end{aligned}
$$

bzw. zu

$$
\overline{x}_l = \frac{\sum_{i=1}^{r} a_l^i \frac{|w_i||\overline{x}_l - a_l^i|^{p_i-2}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} a_l^j \frac{|w_j||\overline{x}_l - a_l^j|^{p_j-2}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}}{\sum_{i=1}^{r} \frac{|w_i||\overline{x}_l - a_l^i|^{p_i-2}}{\|\overline{x} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} \frac{|w_j||\overline{x}_l - a_l^j|^{p_j-2}}{\|\overline{x} - a^j\|_{p_j}^{p_j-1}}}.
$$

Wenn wir auf der rechten Seite dieser Gleichung das zuletzt berechnete Element einsetzen, so können wir eine bessere Näherungslösung erhalten und diese dann wiederum im nächsten Iterationsschritt verwenden. Als Startpunkt bietet sich hierfür zum Beispiel ein Erfahrungswert oder die Lösung desselben Problems mit quadratischer ℓ_2-Norm als induzierender Metrik an, ein entsprechender Lösungsalgorithmus findet sich ebenfalls bei Hillmann (2013). Durch die Vorgabe unterschiedlicher Startpunkte lassen sich gegebenfalls auch mehrere verschiedene Elemente generieren, welche die notwendige Optimalitätsbedingung näherungsweise erfüllen.

Vor dem Beginn der Iterationen prüfen wir außerdem mit den beiden anderen Gleichungen aus Satz 4.1.2 bzw. aus Bemerkung 4.1.2, ob bereits einer oder mehrere der existierenden Standorte optimal sein könnten, da für diese nicht die der Iterationsvorschrift zugrunde liegende notwendige Optimalitätsbedingung gilt. Falls es existierende Standorte gibt, die die jeweils notwendige Optimalitätsbedingung erfüllen, so geben wir diese Menge zusätzlich zu dem iterativ berechneten Element aus.

In Pseudocode ergibt sich damit folgender Algorithmus:

Algorithmus 8.0.1 (*Iterativer Alg. für* $\mathbf{1/\mathbb{R}^n/w_k \gtrless 0/\ell^m_{p,1<p<+\infty}/\Sigma}$)

Eingabe: Existierende Standorte $a^1, \dots, a^m \in \mathbb{R}^n$*, Vektor* $w = (w_1, \dots, w_m) \in \mathbb{R}^m$ *der zugehörigen Gewichte, Vektor* $(p_1, \dots, p_m) \in \mathbb{R}^m$ *der Parameter für die zugehörigen induzierenden Normen, Startpunkt* x^{neu}

Ausgabe: Approximation eines Punktes x^{neu}*, welcher die notwendige Optimalitätsbedingung für das Problem* $1/\mathbb{R}^n/w_k \gtrless 0/\ell^m_{p,1<p<+\infty}/\Sigma$ *erfüllt, sowie alle existierenden Standorte, welche die entsprechende notwendige Optimalitätsbedingung erfüllen*

Start;
Initialisiere $N = \emptyset$*;*

Für $k = 1$ *bis* r

Falls

$$\left\| \sum_{\substack{j=r+1}}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}} \right) - \sum_{\substack{i=1\\ i\neq k}}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}} \right) \right\|_{\frac{p_k}{p_k-1}} \leq |w_k|$$

Füge a^k *zu* N *hinzu;*

Ende;

Ende;

Für $k = r + 1$ *bis* m

Falls

$$\left\| \sum_{\substack{j=r+1\\ j\neq k}}^{m} |w_j| \left(\frac{\operatorname{sgn}(a_1^k - a_1^j)|a_1^k - a_1^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^j)|a_n^k - a_n^j|^{p_j-1}}{\|a^k - a^j\|_{p_j}^{p_j-1}} \right) - \sum_{i=1}^{r} |w_i| \left(\frac{\operatorname{sgn}(a_1^k - a_1^i)|a_1^k - a_1^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}}, \ldots, \frac{\operatorname{sgn}(a_n^k - a_n^i)|a_n^k - a_n^i|^{p_i-1}}{\|a^k - a^i\|_{p_i}^{p_i-1}} \right) \right\|_{\frac{p_k}{p_k-1}} = |w_k|$$

Füge a^k *zu* N *hinzu;*

Ende;

Ende;

Solange eine Abbruchbedingung nicht erfüllt ist

Setze $x^{alt} = x^{neu}$;

Für $l = 1$ *bis* n

Setze

$$x_l^{neu} = \frac{\sum_{i=1}^{r} a_l^i \frac{|w_i||x_l^{alt} - a_l^i|^{p_i-2}}{\|x^{alt} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} a_l^j \frac{|w_j||x_l^{alt} - a_l^j|^{p_j-2}}{\|x^{alt} - a^j\|_{p_j}^{p_j-1}}}{\sum_{i=1}^{r} \frac{|w_i||x_l^{alt} - a_l^i|^{p_i-2}}{\|x^{alt} - a^i\|_{p_i}^{p_i-1}} - \sum_{j=r+1}^{m} \frac{|w_j||x_l^{alt} - a_l^j|^{p_j-2}}{\|x^{alt} - a^j\|_{p_j}^{p_j-1}}};$$

Ende;

Ende;

Gib N *als Menge der Standorte aus, die bereits Lösungen sein könnten;*

Gib x^{neu} *als Approximation eines Punktes aus, welcher die notwendige Optimalitätsbedingung erfüllt;*

Ende;

Als Abbruchbedingung kann hier beispielsweise gewählt werden, dass der Abstand zwischen dem neu berechneten und dem im letzten Iterationsschritt bestimmten Element weniger als eine vorgegebene Genauigkeit $\varepsilon > 0$ beträgt, aber auch das

Erreichen einer festgelegten Anzahl an Iterationsschritten oder eines vorgegebenen Zielfunktionswertes sind mögliche Abbruchbedingungen.

Für den Fall ausschließlich anziehender existierender Standorte mit der ℓ_2-Norm als induzierender Norm für alle Abstandsfunktionen veröffentlichte bereits Weiszfeld (1937) einen Algorithmus zur Bestimmung von Näherungslösungen. Dieser fand zunächst wenig Beachtung, bevor Kuhn und Kuenne (1962) ihn erneut beschrieben. Plastria (2011) gibt einen umfassenden aktuellen Überblick über den Weiszfeld-Algorithmus sowie dessen Varianten und Weiterentwicklungen.

Wenn $p_k = 2$ für alle $k \in \{1, \ldots, m\}$ ist, dann entspricht Algorithmus 8.0.1 nun genau einer Verallgemeinerung des Weiszfeld-Algorithmus. Sowohl die Testbedingungen für die existierenden Standorte als auch die Iterationsvorschrift verallgemeinern dann die Funktionsweise des Weiszfeld-Algorithmus für teil-abstoßende Probleme und im Falle $r = m$, also bei Problemstellungen ohne abstoßende existierende Standorte, geht Algorithmus 8.0.1 genau in den Weiszfeld-Algorithmus über. Damit kann unser Algorithmus als natürliche Erweiterung eines der wichtigsten Lösungsalgorithmen aus der konvexen Standortoptimierung betrachtet werden.

9 Abschließende Betrachtungen und mögliche Anknüpfungspunkte

In der vorliegenden Promotion haben wir einen neuartigen und vereinheitlichten Ansatz verfolgt, um mithilfe der von Ioffe sowie von Kruger und Mordukhovich bereits in den 80er Jahren des letzten Jahrhunderts entwickelten nichtkonvexen Subdifferentiale eine Vielzahl von notwendigen Optimalitätsbedingungen für mehrere Klassen von Standortoptimierungsproblemen mit Anziehung und Abstoßung herzuleiten.

Dabei haben wir vier verschiedene Ansätze, welche klassischerweise für konvexe oder konkave Zielstellungen verwendet werden, für die von uns betrachteten Aufgaben angepasst und erweitert. Die hierdurch erhaltenen Modelle für nichtkonvexe Standortoptimierungsprobleme (SP) mit Median-Zielfunktion, (CP) mit Center-Zielfunktion, (CDP) mit Centdian-Zielfunktion sowie (MP) im mehrkriteriellen Fall wurden in Kapitel 1 vorgestellt und in das Klassifikationsschema nach Hamacher (1995) eingeordnet.

Als Abstandsfunktionen, welche die Distanz zwischen den existierenden Standorten und dem neu zu platzierenden Standort messen, verwendeten wir in allen Aufgabenstellungen die von den ℓ_p-Normen mit $p \in \overline{\mathbb{R}}$ und $p \geq 1$ induzierten Metriken. Nach einem ausführlichen Einblick in die Theorie nichtkonvexer Subdifferentiale und ihrer Kalküle in Kapitel 2 haben wir daher im darauf folgenden Kapitel 3 die Limiting-Subdifferentiale (und damit auch direkt die für unseren Fall äquivalenten approximierenden Subdifferentiale) für die sowohl mit positivem als auch mit negativem Vorzeichen behafteten ℓ_p-Normen exakt berechnet. Des Weiteren haben wir in diesem Kapitel auch Resultate zu den Limiting-Subdifferentialen der Maximumsfunktion und eines nichtlinearen Skalarisierungsfunktionals vorgestellt.

Mittels dieser Resultate ließen sich nun in den Kapiteln 4 bis 7 notwendige Optimalitätsbedingungen für alle der zuvor modellierten Standortoptimierungsprobleme herleiten. Dabei konnten wir weitreichende theoretische Aussagen zu notwendigen Optimalitätsbedingungen für die allgemeinste Form des jeweiligen Problemtyps mit beliebigen gemischten Abstandsfunktionen zu den existierenden Standorten gewinnen. Zudem ließen sich noch stärkere Resultate für praktisch relevante Aufgabenstellungen generieren, bei denen allen existierenden Standorten die gleichen oder zumindest gleichartige Metriken zugeordnet sind.

Abgeschlossen wurde die Arbeit mit einem kurzen Abriss zu möglichen Anwendungen der erzielten Ergebnisse in Form von Algorithmen, welche Elemente generieren, die die notwendige Optimalitätsbedingung der jeweiligen Problemstellung erfüllen.

Mit dieser Promotion wurde die von Hillmann (2013) im Rahmen der Masterarbeit begonnene Forschungstätigkeit fortgeführt. Erfreulicherweise konnten hier viele der in dieser Masterarbeit noch als Anknüpfungspunkte aufgeworfenen Fragestellungen umfassend beantwortet werden. Insbesondere die Ergebnisse zu den Limiting-Subdifferentialen der (negativen) ℓ_p-Normen für $p \in \mathbb{R}$ und $p > 1$ und die daraus abgeleiteten notwendigen Optimalitätsbedingungen, die notwendigen Optimalitätsbedingungen für Probleme mit ℓ_∞-Norm als abstandsinduzierende Funktion zu allen Standorten, die Betrachtung von Problemstellungen mit gemischten Abstandsfunktionen und die Herleitung von notwendigen Optimalitätsbedingungen für diese sowie die Modellierung und Erforschung von teil-abstoßenden Standortoptimierungsproblemen mit anderen Zielfunktionen sind Resultate, mit denen mehrere der wichtigsten von Hillmann (2013) vorgestellten Forschungsbereiche abgeschlossen werden konnten.

Für exakte Lösungsmethoden benötigt man neben den notwendigen Optimalitätsbedingungen jedoch zumeist auch hinreichende Optimalitätsbedingungen, um nicht alle der (gegebenenfalls unendlich vielen) möglicherweise minimalen Punkte auf ihre tatsächliche Minimalität prüfen zu müssen. Zur Herleitung solcher hinreichender Optimalitätsbedingungen hatte Hillmann (2013) die Verwendung des Limiting-Subdifferentials zweiter Ordnung und seines zugehörigen Kalküls vorgeschlagen. Im Rahmen der Promotion wurde zwar auch in dieser Richtung recherchiert, allerdings erweist sich das Limiting-Subdifferential zweiter Ordnung selbst für einfache praktische Aufgabenstellungen schnell als zu komplex und nicht mehr beherrschbar. Da hinreichende Optimalitätsbedingungen aber auch weiterhin von großem Interesse für die Lösung teil-abstoßender Standortoptimierungsprobleme sind, wären diesbezügliche Untersuchungen ein wichtiger Anknüpfungspunkt an diese Arbeit.

Ein weiterer wesentlicher Aspekt möglicher zukünftiger Forschungsaktivität liegt zudem in der Untersuchung von N-Standortproblemen. Während im Rahmen dieser Promotion stets Aufgabenstellungen betrachtet wurden, in denen nur ein einziger neuer Standort zu lokalisieren ist, kommt in praktischen Anwendungen häufig auch die gleichzeitige Platzierung mehrerer neuer Standorte vor. Solche Aufgabenstellungen sind vor allem dadurch motiviert, dass hierbei auch die Entfernungen zwischen allen neu zu errichtenden Standorten direkt in der Zielfunktion mitberücksichtigt werden und somit möglicherweise Synergieeffekte und Kostenersparnisse gegenüber der zeitlich versetzten Errichtung mehrerer einzelner neuer Standorte realisiert werden können.

Zuletzt sind durch die hergeleiteten notwendigen Optimalitätsbedingungen, auch ohne hinreichendes Kriterium, viele Möglichkeiten zur Entwicklung von weiteren Algorithmen gegeben. Insbesondere für generalisierte Proximal-Point-Algorithmen liegen den entsprechenden notwendigen Optimalitätsbedingungen bereits geeignete Strukturen zugrunde. Die Implementierung der Ergebnisse in Algorithmen, deren Erprobung und das Erbringen von Nachweisen über ihre Funktionalität und Eigenschaften bietet somit ein weiteres vielschichtiges Betätigungsfeld für anschließende Forschungen.

Literaturverzeichnis

[1] Truong Q. Bao, Marcus Hillmann und Christiane Tammer. Subdifferentials of nonlinear scalarization functions and applications. In: *Journal of Nonlinear and Convex Analysis* 4 (2017), S. 589–605.

[2] Ehrhard Behrends. *Analysis Band 1*. Wiesbaden: Vieweg, 2004.

[3] Ehrhard Behrends. *Analysis Band 2*. Wiesbaden: Vieweg, 2004.

[4] Jonathan M. Borwein und Alexander D. Ioffe. Proximal analysis in smooth spaces. In: *Set-Valued Analysis* 4 (1996), S. 1–24.

[5] Jonathan M. Borwein und Qiji J. Zhu. A survey of subdifferential calculus with applications. In: *Nonlinear Analysis: Theory, Methods & Applications* 38 (1999), S. 687–773.

[6] Jonathan M. Borwein und Adrian S. Lewis. *Convex analysis and nonlinear optimization*. New York: Springer, 2006.

[7] Radu Ioan Boţ, Sorin-Mihai Grad und Gert Wanka. *Duality in vector optimization*. Berlin, Heidelberg: Springer, 2009.

[8] Jack Brimberg und Zvi Drezner. A new heuristic for solving the p-median problem in the plane. In: *Computers & Operations Research* 40.1 (2013), S. 427–437.

[9] Bundesministerium für Wirtschaft und Energie. *Abschlussbericht der Kommission „Wachstum, Strukturwandel und Beschäftigung"*. `https://www.bmwi.de/Redaktion/DE/Downloads/A/abschlussbericht-kommission-wachstum-strukturwandel-und-beschaeftigung.pdf?__blob=publicationFile`. 2019.

[10] Emilio Carrizosa und Frank Plastria. Location of semi-obnoxious facilities. In: *Studies in Locational Analysis* 12 (1999), S. 1–27.

[11] Pey-Chun Chen, Pierre Hansen, Brigitte Jaumard und Hoang Tuy. Weber's problem with attraction and repulsion. In: *Journal of Regional Science* 32.4 (1992), S. 467–486.

[12] Frank H. Clarke. *Necessary conditions for nonsmooth problems in optimal control and the calculus of variations*. Ph.D. thesis. University of Washington, 1973.

[13] Frank H. Clarke. Generalized gradients and applications. In: *Transactions of the American Mathematical Society* 205 (1975), S. 247–262.

[14] Frank H. Clarke. *Optimization and nonsmooth analysis*. New York: Wiley, 1983.

[15] Leon Cooper. Location-allocation problems. In: *Operations Research* 11.3 (1963), S. 331–343.

[16] Zvi Drezner, Hrsg. *Facility location: A survey of applications and methods*. New York: Springer, 1995.

[17] Zvi Drezner und George O. Wesolowsky. The Weber problem on the plane with some negative weights. In: *Information Systems and Operational Research* 29.2 (1991), S. 87–99.

[18] Zvi Drezner und Horst W. Hamacher, Hrsg. *Facility location: Applications and theory*. Berlin, Heidelberg: Springer, 2002.

[19] Zvi Drezner, Jack Brimberg, Nenad Mladenović und Said Salhi. New heuristic algorithms for solving the planar p-median problem. In: *Computers & Operations Research* 62 (2015), S. 296–304.

[20] Matthias Ehrgott. *Multicriteria optimization*. Berlin, Heidelberg: Springer, 2005.

[21] Wilhelm Forst und Dieter Hoffmann. *Optimization - theory and practice*. New York: Springer, 2010.

[22] GeoGebra. `https://www.geogebra.org/geometry`. Zuletzt abgerufen am 03.06.2023.

[23] Christiane Gerstewitz (Tammer) und Emil Iwanow. Dualität für nichtkonvexe Vektoroptimierungsprobleme. In: *Wissenschaftliche Zeitschrift der TH Ilmenau* 31.2 (1985), S. 61–81.

[24] Christiane Gerstewitz (Tammer). *Beiträge zur Dualitätstheorie der nichtlinearen Vektoroptimierung (Contributions to duality theory in nonlinear vector optimization)*. Dissertation. Technische Hochschule Leuna-Merseburg, 1984.

[25] Christiane Gerth (Tammer) und Petra Weidner. Nonconvex separation theorems and some applications in vector optimization. In: *Journal of Optimization Theory and Applications* 67 (1990), S. 297–320.

[26] Alfred Göpfert und Thomas Riedrich. *Funktionalanalysis*. Leipzig: B. G. Teubner, 1994.

[27] Alfred Göpfert, Thomas Riedrich und Christiane Tammer. *Angewandte Funktionalanalysis*. Wiesbaden: Vieweg+Teubner, 2009.

[28] Alfred Göpfert, Thomas Riedrich und Christiane Tammer. *Approximation und nichtlineare Optimierung in Praxisaufgaben*. Wiesbaden: Springer Spektrum, 2017.

[29] Alfred Göpfert, Hassan Riahi, Christiane Tammer und Constantin Zălinescu. *Variational methods in partially ordered spaces*. New York: Springer, 2003.

[30] Christian Günther, Marcus Hillmann, Christiane Tammer und Brian Winkler. *Facility Location Optimizer (FLO) - A tool for solving location problems*. `http://www.project-flo.de`. Version 1.2.2. 2016.

[31] S. Louis Hakimi. Optimum locations of switching centers and the absolute centers and medians of a graph. In: *Operations Research* 12.3 (1964), S. 450–459.

[32] S. Louis Hakimi. Optimum distribution of switching centers in a communication network and some related graph theoretic problems. In: *Operations Research* 13.3 (1965), S. 462–475.

[33] Jonathan Halpern. The location of a center-median convex combination on an undirected tree. In: *Journal of Regional Science* 16.2 (1976), S. 237–245.

[34] Jonathan Halpern. Finding minimal center-median convex combination (cent-dian) of a graph. In: *Management Science* 24.5 (1978), S. 485–586.

[35] Horst W. Hamacher. *Mathematische Lösungsverfahren für planare Standortprobleme*. Vieweg-Lehrbuch Wirtschaftsmathematik. Braunschweig: Vieweg, 1995.

[36] Horst W. Hamacher, Stefan Nickel und Anja Schneider. *Classification of location problems*. Techn. Ber. Universität Kaiserslautern, 1996.

[37] Horst W. Hamacher und Stefan Nickel. Classification of location models. In: *Location Science* 6.1-4 (1998), S. 229–242.

[38] Horst W. Hamacher, Kathrin Klamroth und Christiane Tammer. Standortoptimierung. In: *Die Kunst des Modellierens: Mathematisch-ökonomische Modelle*. Hrsg. von Bernd Luderer. Wiesbaden: Vieweg+Teubner, 2008, S. 139–156.

[39] Horst W. Hamacher, Andrea Maier, Aleksandra Gross und Philipp Heßler. *Library of Location and Layout Algorithms (LoLoLa)*. `http://lolola.mathematik.uni-kl.de:8000/`. 2015.

[40] Marcus Hillmann. *Lagrange-Multiplikatorenregeln und Algorithmen für nichtkonvexe Standortprobleme*. Masterarbeit. Martin-Luther-Universität Halle-Wittenberg, 2013.

[41] Alexander D. Ioffe. Nonsmooth analysis: Differential calculus of nondifferentiable mappings. In: *Transactions of the American Mathematical Society* 266 (1981), S. 1–56.

[42] Alexander D. Ioffe. Sous-differentielles approaches de fonctions numériques. In: *Comptes rendus de l'Académie des sciences* 292 (1981), S. 675–678.

[43] Alexander D. Ioffe. *Calculus of Dini subdifferentials*. CEREMADE publication 8110. Université de Paris IX Dauphine, 1981.

[44] Alexander D. Ioffe. *Approximate subdifferentials of nonconvex functions*. CEREMADE publication 8120. Université de Paris IX Dauphine, 1981.

[45] Alexander D. Ioffe. Approximate subdifferentials and applications. I: The finite dimensional theory. In: *Transactions of the American Mathematical Society* 281 (1984), S. 389–416.

[46] Alexander D. Ioffe. Approximate subdifferentials and applications. II: Functions on locally convex spaces. In: *Mathematika* 33 (1986), S. 111–128.

[47] Alexander D. Ioffe. Approximate subdifferentials and applications. III: The metric theory. In: *Mathematika* 36 (1989), S. 1–38.

[48] Alexander D. Ioffe. Metric regularity and subdifferential calculus. In: *Russian Mathematical Surveys* 55:3 (2000), S. 501–558.

[49] Alexander D. Ioffe. On the theory of subdifferentials. In: *Advances in Nonlinear Analysis* 1 (2012), S. 47–120.

[50] Alexander D. Ioffe. *Variational analysis of regular mappings*. Cham: Springer, 2017.

[51] Johannes Jahn. *Vector optimization: Theory, applications, and extensions*. New York: Springer, 2004.

[52] Akhtar A. Khan und Fabio Raciti. A multiplier rule in set-valued optimization. In: *Bulletin of the Australian Mathematical Society* 68 (2003), S. 93–100.

[53] Akhtar A. Khan und Miguel Sama. A multiplier rule for stable problems in vector optimization. In: *Journal of Convex Analysis* 19.2 (2012), S. 525–539.

[54] Akhtar A. Khan, Christiane Tammer und Constantin Zălinescu. *Set-valued optimization*. Berlin, Heidelberg: Springer, 2015.

[55] Alexander Y. Kruger. *Generalized differentials of nonsmooth functions and necessary conditions for an extremum*. Ph.D. thesis. Belarus State University, Minsk, 1981.

[56] Alexander Y. Kruger. *Generalized differentials of nonsmooth functions*. Deposited in VINITI #1332-81. Moskau, 1981.

[57] Alexander Y. Kruger. Generalized differentials of nonsmooth functions and necessary conditions for an extremum. In: *Siberian Mathematical Journal* 26 (1985), S. 370–379.

[58] Alexander Y. Kruger und Boris S. Mordukhovich. Minimization of nonsmooth functionals in optimal control problems. In: *Engineering Cybernetics* 16 (1978), S. 126–133.

[59] Alexander Y. Kruger und Boris S. Mordukhovich. Extremal points and the Euler equation in nonsmooth optimization. In: *Doklady Akademii Nauk SSSR* 24 (1980), S. 684–687.

[60] Alexander Y. Kruger und Boris S. Mordukhovich. *Generalized normals and derivatives, and necessary optimality conditions in nondifferential programming*. Deposited in VINITI, #408-80(I), #494-80(II). Moskau, 1980.

[61] Harold W. Kuhn und Robert E. Kuenne. An efficient algorithm for the numerical solution of the generalized Weber problem in spatial economics. In: *Journal of Regional Science* 4.2 (1962), S. 21–33.

[62] Gilbert Laporte, Stefan Nickel und Francisco Saldanha da Gama, Hrsg. *Location Science*. Cham: Springer, 2015.

[63] Wilhelm Launhardt. *Theorie des Trassierens, Teil 1*. Hannover: Schmorl und von Sefeld, 1887.

[64] Wilhelm Launhardt. *The theory of the trace: Being a discussion of the principles of location*. Madras: Lawrence Asylum Press, 1900.

[65] Robert F. Love, James G. Morris und George O. Wesolowsky. *Facilities locations: Models and methods*. Amsterdam: North Holland, 1988.

[66] Marco Manetti. *Topology*. Cham: Springer, 2015.

[67] Costas D. Maranas und Christodoulos A. Floudas. A global optimization method for Weber's problem with attraction and repulsion. In: *Large Scale Optimization*. Hrsg. von Panos M. Pardalos, William W. Hager und Donald W. Hearn. Boston: Springer, 1994, S. 259–285.

[68] William Miehle. Link-length minimization in networks. In: *Operations Research* 6.2 (1958), S. 232–243.

[69] Boris S. Mordukhovich. Maximum principle in problems of time optimal control with nonsmooth constraints. In: *Journal of Applied Mathematics and Mechanics* 40 (1976), S. 960–969.

[70] Boris S. Mordukhovich. Approximation and maximum principle for nonsmooth problems of optimal control. In: *Russian Mathematical Surveys* 196 (1977), S. 263–264.

[71] Boris S. Mordukhovich. Metric approximations and necessary optimality conditions for general classes of extremal problems. In: *Soviet Mathematics Doklady* 22 (1980), S. 526–530.

[72] Boris S. Mordukhovich. Nonsmooth analysis with nonconvex generalized differentials and adjoint mappings. In: *Doklady Akademii Nauk SSSR* 28 (1984), S. 976–979.

[73] Boris S. Mordukhovich. *Approximation methods in problems of optimization and control*. Moskau: Nauka, 1988.

[74] Boris S. Mordukhovich. *Variational analysis and generalized differentation I. Basic theory*. Bd. 330. Grundlehren der mathematischen Wissenschaften - A series of comprehensive studies in mathematics. Heidelberg: Springer, 2006.

[75] Boris S. Mordukhovich. *Variational analysis and generalized differentation II. Applications*. Bd. 331. Grundlehren der mathematischen Wissenschaften - A series of comprehensive studies in mathematics. Heidelberg: Springer, 2006.

[76] Boris S. Mordukhovich. *Variational analysis and applications*. Cham: Springer, 2018.

[77] Boris S. Mordukhovich und Alexander Y. Kruger. Necessary optimality conditions for a terminal control problem with nonfunctional constraints. In: *Doklady Akademii Nauk SSSR* 20 (1976), S. 1064–1067.

[78] Boris S. Mordukhovich und Yongheng Shao. On nonconvex subdifferential calculus in Banach spaces. In: *Journal of Convex Analysis* 2 (1995), S. 211–227.

[79] Boris S. Mordukhovich und Yongheng Shao. Nonsmooth sequential analysis in Asplund spaces. In: *Transactions of the American Mathematical Society* 348 (1996), S. 1235–1280.

[80] Jean-Jacques Moreau. Fonctions duales et points proximaux dans un espace hilbertien. In: *Comptes rendus de l'Académie des sciences* 255 (1962), S. 2897–2899.

[81] Jean-Jacques Moreau. Fonctionelles sous-differentiables. In: *Comptes rendus de l'Académie des sciences* 257 (1963), S. 4117–4119.

[82] Stefan Nickel und Eva-Maria Dudenhöffer. Weber's problem with attraction and repulsion under polyhedral gauges. In: *Journal of Global Optimization* 11.4 (1997), S. 409–432.

[83] Stefan Nickel und Justo Puerto. *Location theory: A unified approach*. Berlin, Heidelberg: Springer, 2005.

[84] OpenStreetMap. `https://www.openstreetmap.de/index.html`. Zuletzt abgerufen am 24.07.2019.

[85] Adriano Pascoletti und Paolo Serafini. Scalarizing vector optimization problems. In: *Journal of Optimization Theory and Applications* 42 (1984), S. 499–524.

[86] Dionisio Pérez-Brito und José A. Moreno Pérez. The generalized p-centdian on network. In: *TOP* 8 (2000), S. 265–285.

[87] Robert R. Phelps. *Convex functions, monotone operators and differentiability*. Bd. 1364. Lecture notes in mathematics. Berlin, Heidelberg: Springer, 1993.

[88] James V. Pinto. Launhardt and location theory: Rediscovery of a neglected book. In: *Journal of Regional Science* 17.1 (1977), S. 17–29.

[89] Frank Plastria. The Weiszfeld algorithm: Proof, amendments, and extensions. In: *Foundations of location analysis*. Hrsg. von Horst A. Eiselt und Vladimir Marianov. New York: Springer, 2011, S. 357–389.

[90] R. Tyrrell Rockafellar. *Convex functions and dual extremum problems*. Ph.D. thesis. Harvard University, 1963.

[91] R. Tyrrell Rockafellar. Duality theorems for convex functions. In: *Bulletin of the American Mathematical Society* 70 (1964), S. 189–192.

[92] R. Tyrrell Rockafellar. *Convex analysis*. Princeton: Princeton University Press, 1970.

[93] R. Tyrrell Rockafellar und Roger J.-B. Wets. *Variational analysis*. New York: Springer, 1998.

[94] Thomas Simpson. *The doctrine and application of fluxions*. London, 1750.

[95] Ivan Singer. A Fenchel-Rockafellar type duality theorem for maximization. In: *Bulletin of the Australian Mathematical Society* 20.2 (1979), S. 193–198.

[96] Christiane Tammer und Petra Weidner. *Scalarization and separation by translation invariant functions*. Cham: Springer, 2020.

[97] Luc-Normand Tellier. *Économie spatiale: Rationalité économique de l'espace habité*. Chicoutinmi: Gaëtan Morin, 1985.

[98] John F. Toland. Duality in nonconvex optimization. In: *Journal of Mathematical Analysis and Applications* 66.2 (1978), S. 399–415.

[99] Andrea Wagner. *A new duality based approach for the problem of locating a semi-obnoxious facility*. Dissertation. Martin-Luther-Universität Halle-Wittenberg, 2015.

[100] Andrea Wagner, Juan-Enrique Martínez-Legaz und Christiane Tammer. Locating a semi-obnoxious facility - a Toland-Singer duality based approach. In: *Journal of Convex Analysis* 23.4 (2016), S. 1185–1204.

[101] Alfred Weber. *Über den Standort der Industrien*. Tübingen: J.C.B. Mohr, 1909.

[102] Endre Weiszfeld. Sur le point pour lequel la somme des distances de n points donnés est minimum. In: *Tohoku Mathematical Journal* (1937), S. 355–386.

[103] Dirk Werner. *Funktionalanalysis*. Berlin: Springer Spektrum, 2018.

[104] George O. Wesolowsky. The Weber problem: History and perspectives. In: *Location Science* 1 (1993), S. 5–23.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift zum Thema „Notwendige Optimalitätsbedingungen für nichtkonvexe Standortoptimierungsprobleme“ eigenständig und nur unter Zuhilfenahme der angegebenen Quellen angefertigt zu haben. Alle Inhalte, die ich diesen Werken wörtlich oder inhaltlich entnommen habe, wurden als solche kenntlich gemacht.

Insbesondere wurden alle der vorgestellten Resultate und Algorithmen, welche den wissenschaftlichen Erkenntnisgewinn durch diese Arbeit widerspiegeln, von mir selbst hergeleitet bzw. entwickelt und sind nach meinem besten Wissen und Gewissen in dieser Form bisher unveröffentlicht.

Die Dissertationsschrift wurde zuvor weder an der Martin-Luther-Universität Halle-Wittenberg als Prüfungsleistung für eine Qualifikation eingereicht noch einer anderen Hochschule im In- oder Ausland in gleicher oder ähnlicher Art und Weise vorgelegt.

Weimar, den 8. März 2024

Lebenslauf

Persönliche Angaben

Name:	Marcus Hillmann
Geburtsdatum:	15. April 1988
Geburtsort:	Halle (Saale)

Berufliche Tätigkeit

08/2023 - 10/2023	Wissenschaftlicher Mitarbeiter am Institut für Bauphysikalische Qualitätssicherung, Fachhochschule Erfurt • Mitarbeit im Projekt ThÖMie
08/2022 - 08/2023	Wissenschaftlicher Mitarbeiter im Team des Vizepräsidenten für Studium und Lehre, Fachhochschule Erfurt • Mitarbeit im Projekt ProMINT
04/2021 - 01/2023	Wissenschaftlicher Mitarbeiter an der Fakultät für Landschaftsarchitektur, Gartenbau und Forst, Fachhochschule Erfurt • Forschungstätigkeit im Projekt „Optische Bio-Masseermittlung“ (OBioM)
03/2018 - 10/2018	Lehrkraft für besondere Aufgaben am Institut für Mathematik, Martin-Luther-Universität (MLU) Halle-Wittenberg

Publikationen

05/2017	T. Q. Bao, M. Hillmann and Chr. Tammer *Subdifferentials of nonlinear scalarization functions and applications* In: Journal of Nonlinear and Convex Analysis 4/2017, pp. 589-605.
04/2015	C. Günther, M. Hillmann, Chr. Tammer and B. Winkler Facility Location Optimizer (FLO) - A tool for solving location problems, http://www.project-flo.de

Ausbildung

10/2013 - 11/2023	Doktorand am Institut für Mathematik, MLU Halle-Wittenberg Thema der Promotion: „Notwendige Optimalitätsbedingungen für nichtkonvexe Standortoptimierungsprobleme“ Betreuerin: Frau Prof. Dr. Christiane Tammer 04/2014 - 03/2017 Graduiertenstipendium des Landes Sachsen-Anhalt 10/2013 - 03/2014 Reisestipendium der Stiftung Theoretische Physik und Mathematik 11/2013 - 01/2014 Forschungsaufenthalt am Rochester Institute of Technology (RIT) in Rochester, NY, USA
10/2012 - 03/2014	M.Sc. in Mathematik, MLU Halle-Wittenberg Thema der Masterarbeit: „Lösungsmethoden für Maximum-Entropy-Probleme“ Betreuerin: Frau Prof. Dr. Christiane Tammer
10/2010 - 07/2013	M.Sc. in Wirtschaftsmathematik, MLU Halle-Wittenberg Thema der Masterarbeit: „Lagrange-Multiplikatorenregeln und Algorithmen für nichtkonvexe Standortprobleme“ Betreuerin: Frau Prof. Dr. Christiane Tammer 10/2011 - 09/2012 Deutschlandstipendium
10/2007 - 07/2010	B.Sc. in Wirtschaftsmathematik, MLU Halle-Wittenberg Thema der Bachelorarbeit: „Der Benson-Algorithmus“ Betreuer: Herr Prof. Dr. Andreas Löhne
08/1998 - 03/2007	Abitur, Elisabeth-Gymnasium Halle

Weimar, den 8. März 2024