

Warmstart-Strategien in der parametrisierten Optimierung mit Anwendung in der multikriteriellen Optimierung

DISSERTATION

zur Erlangung des Grades eines
Doktors der Naturwissenschaften

dem Fachbereich Mathematik der Universität Dortmund

vorgelegt am 4. Mai 2005

von

Andree Heseler

Eingereicht am 4. Mai 2005

Tag der mündlichen Prüfung am 16. Dezember 2005

Vorsitzender: Prof. Dr. W. Kaballo

Erstgutachter: Prof. Dr. J. Fliege

Zweitgutachter: Prof. Dr. J. Jahn

Drittgutachter: Prof. Dr. W. Achtziger

Symbolverzeichnis

I	Einheitsmatrix	
e	Einsvektor $(1, \dots, 1)^\top$	
Ξ	Menge der Parameter für (P_ξ)	(2.1)
$\mathcal{F}(\xi)$	Restriktionsmenge für (P_ξ)	(2.2)
$h : \mathbb{R}^n \times \Xi \rightarrow \mathbb{R}^{m_h}$	Gleichungsrestriktionen	(2.2)
$g : \mathbb{R}^n \times \Xi \rightarrow \mathbb{R}^{m_g}$	Ungleichungsrestriktionen	(2.2)
$f : \mathbb{R}^n \times \Xi \rightarrow \mathbb{R}$	Zielfunktion	
S, Π, \dots	Diagonalmatrizen $\text{diag}(s)$, $\text{diag}(\pi)$	
\tilde{n}	$\tilde{n} = n + m_g + m_h$	(2.6)
\mathcal{L}	Lagrangefunktion	(2.3), (2.7)
F, F_τ	implizite (τ -pert.) KKT-Abb.	(2.8), (2.22), (2.26)
$r : \mathbb{R}^{\tilde{n}} \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}$	Residuenabbildung zu F, F_τ	(2.10)
τ	Perturbation in Kompl.bed.	
$\mathcal{N}_F(\Xi, \tau)$	τ -Urbildmenge zu (P_ξ)	Def. 2.2.7
L	Kurzschreibweise für $D_{xx}\mathcal{L}(w; \xi)$	Bew. Lem. 2.2.8
A	Kurzschreibweise für $D_x h(x; \xi)$	Bew. Lem. 2.2.8
C	Kurzschreibweise für $D_x g(x; \xi)$	Bew. Lem. 2.2.8
x'	Kurzschreibweise für $D_\xi x(\xi)$	Bew. Lem. 2.2.8
λ'	Kurzschreibweise für $D_\xi \lambda(\xi)$	Bew. Lem. 2.2.8
s'	Kurzschreibweise für $D_\xi s(\xi)$	Bew. Lem. 2.2.8
π'	Kurzschreibweise für $D_\xi \pi(\xi)$	Bew. Lem. 2.2.8
r_L	Kurzschreibweise für $D_{\xi x}\mathcal{L}(w; \xi)$	Bew. Lem. 2.2.8
r_A	Kurzschreibweise für $D_\xi h(x; \xi)$	Bew. Lem. 2.2.8
r_C	Kurzschreibweise für $D_\xi g(x; \xi)$	Bew. Lem. 2.2.8
$r_{S\Pi}$	Kurzschreibweise für $D_\xi(S_\xi \Pi_\xi e - \tau)$	Bew. Lem. 2.2.8

ρ, ρ_w, ρ_F	Warmstartabbildungen	Def. 2.3.1, 2.32, 2.34
$\Delta w, \Delta w_w, \Delta w_F$	Differenzenabbildungen	Def. 2.3.1, 2.33, 2.35
ΔF	Perturbationsabb. in (2.56)	(2.59)
$M(w; \xi)$	Entspr. $D_w F_\tau(w; \xi)$ in (2.56)	(2.60)
$\Delta M(w)$	Perturbationsmatrix in (2.56)	(2.62)
$\mathcal{T}(\tau, \theta)$	θ -Umgebung um τ -Urbilder	Def. 2.4.7
$\mathcal{E}(Z)$	Menge der effizienten Punkte z aus Z	Def. 4.1.1
$\mathcal{P}(f, X)$	Pareto­menge bez. $f : X \rightarrow \mathbb{R}^n$	Def. 4.1.1
$\mathcal{K}(f, X)$	KKT-Punktmenge bez. $\min_{x \in X} f$	Def. 4.2.3
Ω	Menge d. Skalarisierungsparameter	(4.15)
$N_i(\alpha)$	Niveaumenge d. Fkt. f_i zu $f_i(x) = \alpha$	(4.31)
P_ξ	Parametrisiertes Problem	Seite 19
QP_ξ	Quadratisches, parametrisiertes Problem	Seite 41
P	Primales Standardproblem	Seite 97
D	Duales Standardproblem	Seite 98
MCP	Bikriterielles, quadratisches Problem	Seite 110
$T_a M$	Tangentialebene an M in a	Def. A.2.5
$N_a M$	Normalenvektormenge von M in a	Def. A.2.6

Inhaltsverzeichnis

1	Einführung	7
1.1	Motivation	8
1.1.1	Der Zugang über Innere-Punkte Methoden	8
1.1.2	Ein Anwendungsbeispiel: Multikriterielle Optimierung	12
1.2	Inhaltsübersicht	15
I	Theorie und ihre Anwendung	17
2	Parametrisierte Optimierung	18
2.1	Das Standardproblem	19
2.2	Stabilität der Lösungsmenge	22
2.2.1	Stabilität für nichtlineare Probleme	23
2.2.2	Stabilität τ -perturbierter KKT-Systeme	25
2.3	Approximierende Lösungen	29
2.3.1	Die Warmstartabbildung	29
2.3.2	Wiederverwendbarkeit von Warmstartpunkten	34
2.4	Sensitivität τ -perturbierter Systeme	40
2.4.1	Abschätzung zulässiger Parametermengen	46
2.4.2	Sensitivität der Komplementaritätsbedingungen	48
2.4.3	Die Approximation der Lösungsmenge	52

3	Geometrie der Lösungsmenge	60
3.1	Lösungserhaltende Perturbationen	60
3.2	Bestimmung von Parameteränderungen	67
4	Multikriterielle Optimierung	73
4.1	Grundlagen	73
4.2	Optimalitätsbedingungen	75
4.3	Pareto- und Effizienzmengen	78
4.4	Bogenlängen von Effizienzmengen	89
4.5	Approximationen mit τ -perturbierten Systemen	91
II	Software und Tests	94
5	Software und numerische Ergebnisse	95
5.1	QIPP – Ein Innere-Punkte Löser	96
5.1.1	Optimierung quadratischer Aufgaben	97
5.1.2	Das dualisierte Optimierungsproblem	97
5.1.3	Das gestörte KKT-System	100
5.1.4	Das Mehrotra-Verfahren	102
5.1.5	Das erweiterte System	107
5.2	MCQIPP – Ein multikriterieller Optimierer	109
5.3	Numerisches Beispiel: Steuerung von Kraftwerken	113
6	Zusammenfassung und Ausblick	124
A	Grundlagen der Differentialgeometrie	127
A.1	Parametrisierte Kurven	127
A.2	Flächen und Mannigfaltigkeiten	128
B	Ausgaben zu numerischen Auswertungen	131
C	Technische Daten des Rechners	135

Kapitel 1

Einführung

Die Kernidee zur vorliegenden Arbeit basiert auf einem Prinzip der Theorie zu *Inneren-Punkte Methoden* (IP-Methoden), einer recht jungen Klasse von Optimierungsverfahren. Im Jahre 1968 veröffentlichten Fiacco und McCormick ihr Buch [FiaMcCo90] (Neuaufgabe von 1990), welches erstmals die Idee verfolgte, optimale Lösungen konvexer, nichtlinearer Optimierungsprobleme mit Hilfe von inneren Punkten zu ermitteln. Den letzten Anstoß gab nach einer längeren Pause die 1984 erscheinende Arbeit von Karmarkar [Ka84], worin ein Lösungsverfahren für lineare Optimierungsprobleme vorgestellt wurde, das im Gegensatz zur Simplex-Methode eine polynomielle Laufzeit aufwies. Auch in der Anwendung erwies sich das Verfahren als konkurrenzfähig.

Ein weiteres entscheidendes Ereignis war die Veröffentlichung des Buches von Nemerov und Nemirovskii [NeNe94], die mit den beiden Eigenschaften der *Selbstkonkordanz* und *Selbstbeschränkung* das Prinzip der IP-Methoden für konvexe nichtlineare Programme erweiterten.

Seitdem konnten noch andere, teils exotische Erweiterungen hinsichtlich der Aufgabenstellung abgeleitet werden. In dieser Vielfalt liegt sicherlich einer der Vorteile Innerer-Punkte Methoden.

Hinter den entsprechenden IP-Methoden stecken spezielle Barriere-Verfahren, welche die Iterierten bezüglich der Ungleichungsrestriktionen im „Inneren“ der Restriktionsmenge halten, d.h. die Ungleichungsrestriktionen sind stets strikt erfüllt.

1.1 Motivation

1.1.1 Der Zugang über Innere-Punkte Methoden

Der zentrale Begriff des *Warmstartpunkts* im hier vorliegenden Sinne stammt aus der IP-Theorie. Dem Verfasser ist die Existenz einer erschöpfenden Theorie über Forderungen an einen primal-dual zulässigen Startpunkt nicht bekannt, die Hinlänglichkeit bezüglich des Gelingens von IP-Verfahren versprechen.

Zwei Aspekte sind jedoch offensichtlich:

- i) Einerseits sollten die Iterierten möglichst weit von den Grenzen der Ungleichungsrestriktionen liegen, da sich das numerische Verhalten wegen schlecht konditionierter Gleichungssysteme zur Bestimmung der Iterierten i.d.R. verschlechtert, je näher diese an den Restriktionsrand gelangen.
- ii) Andererseits ist die Nähe zur optimalen Lösung von Vorteil, außer wenn große Residuen vorliegen.

Sofern das Optimum am Restriktionsrand liegt, gelangen die Iterierten mit jeder Iteration zwangsläufig näher an den Rand. Gerade hier liegt die Schwierigkeit, weiterhin im relativen Inneren der Restriktionsmenge zu bleiben. Einige IP-Verfahren machen sich einen sehr eleganten Lösungsansatz zu nutze, gemeint ist die Orientierung an einem *zentralen Pfad* (s. [Meg89], [McL80], [Son89], [BaLa89I], [BaLa89II], [BaLa90]).

Zur Begriffserklärung des zentralen Pfads betrachte man beispielhaft das konvex-quadratische Optimierungsproblem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \frac{1}{2}x^\top Qx + c^\top x \\ \text{N.B.} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{1.1}$$

mit positiv semidefiniter und quadratischer Matrix $Q \in \mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Ein Punkt $x^* \in \mathbb{R}^n$ ist genau dann optimale Lösung von (1.1), wenn Vektoren λ^* und s^* existieren, mit denen das System

$$\begin{pmatrix} -Q & A^\top & I \\ A & 0 & 0 \\ S^* & 0 & X^* \end{pmatrix} \begin{pmatrix} x^* \\ \lambda^* \\ s^* \end{pmatrix} = \begin{pmatrix} c \\ b \\ \tau e + X^* S^* e \end{pmatrix}, \quad x^*, s^* \geq 0 \tag{1.2}$$

für $\tau = 0$ erfüllt ist, wobei $S^* = \text{diag}(s^*)$, $X^* = \text{diag}(x^*)$ und $e = (1, \dots, 1)^\top$ gelten.

In diesem Fall entspricht (1.2) den KKT-Bedingungen. Die Komplementaritätsbedingungen $-s^\top x = 0$ werden unter der Voraussetzung strikter Komplementarität $s_i + x_i > 0, \forall i = 1, 2, \dots, n$ mittels des Teilsystems $S^*x^* + X^*s^* = 0$ beschrieben.

Um die natürlich induzierte Halbordnung des \mathbb{R}^n zu beschreiben, werden die Relationen „ $>$ “, „ $<$ “, „ \geq “ und „ \leq “ auch für Vektoren eingeführt. Für x und $y \in \mathbb{R}^n$, die bez. einer dieser Relation in Beziehung zueinander stehen, ist jede Komponente $x_j, j = 1, 2, \dots, n$ entsprechend größer, kleiner, größer oder gleich bzw. kleiner oder gleich als y_j .

Für $\tau > 0$ erfüllen die zulässigen Punkte zu (1.2) die Ungleichungsrestriktionen, hier $x \geq 0$, strikt. Der zentrale Pfad entspricht der Menge

$$\{w = (x, \lambda, s) \mid w \text{ erfüllt (1.2), } \tau > 0\}.$$

Elemente des zentralen Pfads befriedigen Forderung (i) absolut.

Pfadverfolgungsverfahren wie z.B. die *Short-* oder *Long-Step-Methode* (s. [KoMiYo89II], [Wri97], [KoMiYo89I], [MoAd89I]) orientieren sich am zentralen Pfad, indem die Iterierten innerhalb einer Toleranzumgebung des Pfades gehalten werden. In jeder Iteration wird ein zugeordneter Störungsparameter τ gegen Null reduziert. Die Null wird selbstverständlich nie erreicht, jedoch bricht der jeweilige Algorithmus ab, sobald τ hinreichend klein ist. Da für sehr kleine τ auch die Komplementaritätsbedingungen fast erfüllt sind, ist die jeweilige Iterierte eine entsprechend genaue Approximation der Lösung.

Der in Kapitel 5 vorgestellte *Algorithmus von Mehrotra* (Alg. 5.1.3) ist ebenfalls in diese Verfahrensklasse einzuordnen.

Es zeigt sich, dass die zentralen Pfade zweier Optimierungsaufgaben, die sich lediglich durch kleine Perturbationen $\Delta c, \Delta Q, \Delta A, \Delta b$ unterscheiden, entsprechend nahe beieinander liegen (s. [YilWri02], [FliHes02]). Insofern lässt sich für eine Iterierte w der unperturbierten Aufgabe unter Umständen ein Korrekturschritt Δw ansetzen, der es erlaubt, von einem Pfad zum anderen zu springen. Der sich ergebende Punkt \tilde{w} ist bei hinreichend kleinen Perturbationen wieder als Iterierte des perturbierten Problems nutzbar und wird deshalb *Warmstartpunkt* genannt. Abbildung 1.1 visualisiert den zugrundeliegenden Vorgang und beinhaltet die zugehörigen Gleichungssysteme.

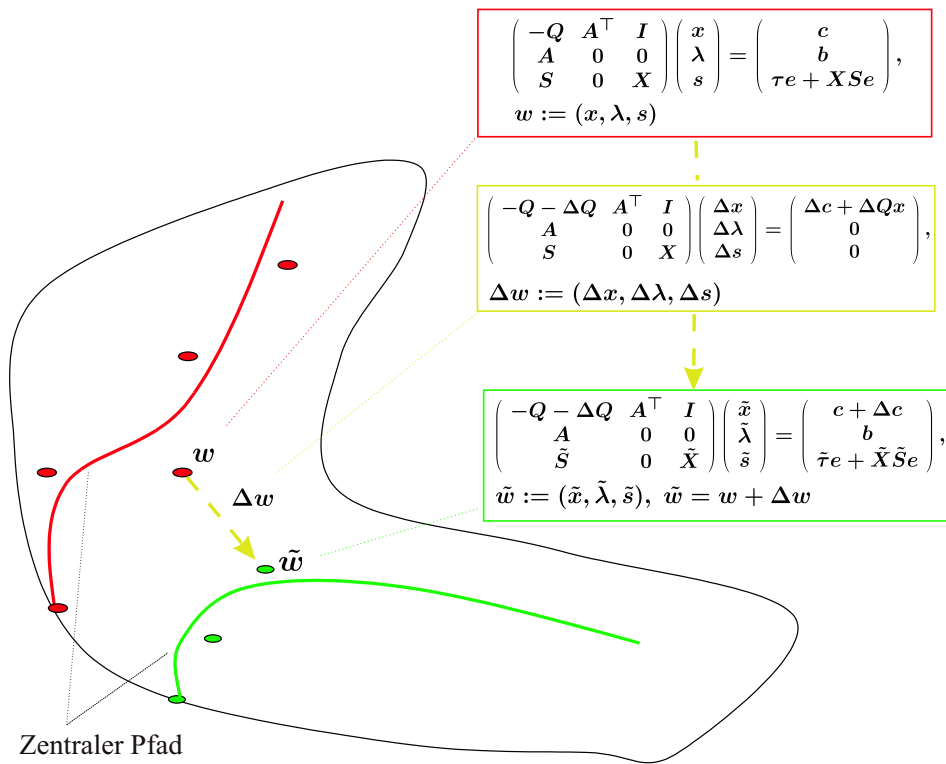


Abbildung 1.1: Berechnung eines Warmstartpunkts

So gewonnene Warmstartpunkte haben den Vorteil, je nach genutzter Iterierten w , sehr nahe an der optimalen Lösung zu liegen. Es ist allerdings nicht trivial, Voraussagen darüber zu treffen, ob der Warmstartpunkt wieder in der Nähe des neuen zentralen Pfads liegt, geschweige denn innerer Punkt ist. Einfaches Ausprobieren ist problematisch, weil ein Versuch dem Aufwand einer Iteration im IP-Verfahren entspricht. Yildirim und Wright [YilWri02] sowie Fliege und Heseler [FliHes02] konnten für den linearen bzw. quadratischen Fall hinreichende Bedingungen an Perturbationen stellen, jedoch ist die Lücke zwischen Theorie und Praxis sehr groß.

Sofern es gelingt, wieder an wohlzentrierte Punkte zu gelangen, ergeben Warmstartpunkte ausgezeichnete Startpunkte im Sinne der oben genannten Forderungen (i) und (ii).

Es finden sich bisher sehr wenige Publikationen zu Warmstartpunkten für IP-Methoden. Anwendungsgebiete sind immer dort gegeben, wo mehrere sich ähnelnde Optimierungsprobleme mit IP-Verfahren gelöst werden. Unter ähnlichen Programmen versteht man solche, die durch Hinzugabe und/oder Wegnahme einiger Restriktionen, ferner durch Perturbationen der Restriktionen und Zielfunktionen entstehen.

Erstere Klasse ergibt sich z.B. zu *Cutting-Plane-Methoden* [Kel60] und *Branch-And-Bound-* bzw. *Branch-And-Cut-Methoden* in der gemischt-ganzzahligen Opti-

mierung [Wol98]. Warmstartpunkte für Cutting-Plane-Aufgaben wurden von Gondzio [Gon98] entwickelt; für Branch-And-Bound-Methoden zu gemischt-ganzzahligen Problemen gibt es von Borchers und Mitchell [BoMi92] erste Ansätze.

Die zweite Klasse, welche durch Perturbationen der Daten charakterisiert ist, entsteht unter anderem bei der *Dekompositionsmethode von Bender* [Ben62] oder in der multikriteriellen Optimierung (s. nächsten Abschnitt). Gondzio [GoVi99] behandelt Warmstartpunkte zur Lösung von Subproblemen der Dekomposition großer Gleichungssysteme.

Yıldırım und Wright [YilWri02] untersuchten allgemein perturbierbare lineare Optimierungsprobleme, deren Perturbation innerhalb eines Newton-Schrittes absorbiert werden konnten und lieferten eine Sensitivitätsanalyse sowie ein hinreichendes Kriterium zur Gewinnung wiederverwendbarer Warmstartpunkte. Gleichzeitig quantifizierten sie den Zusammenhang zwischen der Größe der Perturbation und der notwendigen Entfernung der Iterierten zum Optimum. Gondzio und Grothey [GonGro03] verbesserten diesen Ansatz und wandten die Ergebnisse auf Dekompositionsprobleme an.

Fliege und Hiesler [FliHes02] erweiterten die Resultate von Yıldırım und Wright für konvex-quadratische Optimierungsprobleme und zeigten die Nutzbarkeit von Warmstartpunkten in der multikriteriellen Optimierung durch numerische Tests. Heermann, Fliege und Weyers [HeFlWe04] experimentierten mit einer neuartigen Warmstartstrategie. Von einem Startpunkt ausgehend werden in jeder Iteration Verzweigungen bezüglich der Parameter vorgenommen, so dass eine Baumstruktur entsteht deren Blätter die Effizienzmenge approximiert.

Eine Abhandlung über Warmstartpunkte für konvexe, nichtlineare Optimierungsprobleme in der multikriteriellen Optimierung leistet Fliege [Fli01], [Fli04]. Er betrachtet u.a. Warmstartpunkte für zentrale Pfade primaler Methoden und selbstkonkordante Barrierefunktionen.

1.1.2 Ein Anwendungsbeispiel: Multikriterielle Optimierung

In der *multikriterielle Optimierung* - näher in Kapitel 4 beschrieben - werden vektorwertige Zielfunktionen

$$f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_k \end{pmatrix}$$

unter Nebenbedingungen betrachtet. Im Gegensatz zur *skalaren Optimierung*, d.h. Optimierungsaufgaben mit einer reellwertigen Zielfunktion, ergibt sich aufgrund verschiedener Gewichtungsmöglichkeiten der Ziele f_1, f_2, \dots, f_k eine ganze Menge optimaler Lösungen, der *Effizienzmenge*.

Eine gebräuchliche Methode zur Bestimmung der Effizienzmenge ist die Skalierungsmethode. Legt sich ein Entscheidungsträger auf eine Gewichtung $\xi \in \mathbb{R}_{\geq 0}^k$ der Zielfunktionen f_1, f_2, \dots, f_k fest, so ergibt sich mit $\xi^\top f$ eine reellwertige Zielfunktion und damit ein skalares Optimierungsprogramm, dessen Lösung die optimale Alternative liefert.

Die gesamte Effizienzmenge ergibt sich demnach als Lösungsmenge eines *parametrisierten Optimierungsproblems* und kann approximiert werden, indem für möglichst viele Gewichtungen das entsprechende Ersatzproblem gelöst wird. Die entstehenden Programme unterscheiden sich dabei lediglich durch Perturbationen in der skalarisierten Zielfunktion. Erfüllen die Ersatzzielfunktionen und Restriktionen die Voraussetzungen für die Anwendung von primal-dualen Pfadverfolgungs-IP-Methoden, können für kleine Gewichtsänderungen bereits berechnete Iterationen für die Ermittlung von Warmstartpunkten genutzt werden.

Die anfangs erwähnte Kernidee besteht gerade darin, das Prinzip der Warmstartpunktberechnung für die Approximation der Effizienzmenge zu nutzen. Dazu betrachte man multikriterielle Optimierungsprobleme als Spezialfall der parametrisierten Optimierung. Sofern das Optimierungsproblem so geschaffen ist, dass jedem Parameter mittels der skalarisierten Ersatzfunktion genau ein KKT-Punkt zugeordnet werden kann, ist die Existenz einer Lösungsabbildung von Parametern auf Lösungen gesichert. Über vorauszusetzende Regularitätsbedingungen sind damit notwendige Bedingungen für Optima gegeben. Mit Konvexität gilt sogar die Hinlänglichkeit.

Die Charakterisierung der Lösungen durch die KKT-Punkte hat den Vorteil, dass diese eine Nullstellenabbildung liefern, welche die Lösungsabbildung implizit be-

schreibt. Findet dahingehend der Satz über implizite Funktionen (Satz A.0.1) Anwendung, kann lokal sogar Differenzierbarkeit nachgewiesen werden. Gleichzeitig ergeben sich Ableitungsinformationen, die eine Approximation der Lösungsmenge erlauben, sogenannte *Homotopie-* oder *Pfadverfolgungs-Methoden*.

Man unterscheide die hier angesprochene Pfadverfolgung zu den bereits erwähnten IP-Pfadverfolgungsmethoden. Während die entsprechenden IP-Methoden die Verfolgung des zentralen Pfads zur Gewinnung einer optimalen Lösung anstreben, ist es diesmal die gesuchte Lösungsmenge selbst, die den Pfad bildet. Obwohl sich beide Verfahren derselben Technik bedienen, unterliegen sie doch zweier unterschiedlicher Zielsetzungen und sollten im weiteren Verlauf nicht verwechselt werden.

Bezüglich der Homotopiemethoden wurde von Allgower und Georg [AlGe92] ein Prediktor-Korrektor-Verfahren vorgeschlagen, für Lösungsmengen, die durch nicht-lineare Gleichungen

$$H(u) = 0, \quad H : \mathbb{R}^{N+1} \rightarrow \mathbb{R}^N,$$

H ist glatt, gegeben sind. Guddat et al. [GuVaJo90], [GuVaTaWe85] betrachten die Lösungsmengen aus der Sicht parametrisierter Optimierungsprobleme sowie deren KKT-Systeme und trifft damit auf dasselbe Prinzip wie Allgower. Eine spezielle Methode zur Ermittlung von Effizienzmengen liefert Rakowska et al. [RaHaWa91]. Gemeinsamer Nachteil all dieser Methoden ist die Beschränkung auf einen reellwertigen Parameter und damit auf bikriterielle Probleme.

Einen allgemeineren Ansatz gibt das Buch von Rheinboldt [Rhei86]. Darin werden die implizit gegebenen Lösungsmengen unter recht allgemeinen Voraussetzungen als differenzierbare Mannigfaltigkeiten identifiziert. Dies ermöglicht eine Abbildung auf Karten und damit eine Parametrisierung der Lösungen.

Darauf baut Hillermeier [Hil00] auf und zeigt unter bestimmten Voraussetzungen, dass die Menge der KKT-Punkte skalarisierter, multikriterieller Optimierungsprobleme eine $(k - 1)$ -dimensionale, differenzierbare Untermannigfaltigkeit repräsentiert. Aufgrund dessen kann eine *generalisierte Homotopie-Strategie* für multikriterielle Optimierungsprobleme repräsentiert werden, die auch mehrdimensionale Parameter und damit beliebig viele Zielfunktionen zulässt.

All diese Methoden haben zwei gemeinsame Nachteile. Zum einen besteht die Notwendigkeit, dass neu gewonnene Punkte wieder auf der Mannigfaltigkeit liegen bzw. eine Nullstelle der impliziten Funktion sind. Dies wird mit Newton-Schritten im Korrektorschritt erreicht. Die nötige Anzahl der Schritte hängt davon ab, wie viel

Iterationen das angesetzte Newton-Verfahren benötigt um das zugrunde liegende Nullstellenproblem zu lösen oder zumindest hinreichend zu approximieren. Im Algorithmus von Hillermeier werden zudem für jeden Punkt QR-Zerlegungen mit Householder-Transformationen berechnet, was zusätzlichen Aufwand bereitet. Zum anderen setzen die Algorithmen direkt an der gefragten Lösungsmenge an, was nicht selten numerische Instabilitäten mit sich bringt. So existieren für allgemeine, konvexe Zielfunktionen i.d.R. nicht einmal wohldefinierte Lösungsabbildungen, da einem Parameter teilweise eine ganze Menge von KKT-Punkten zugeordnet werden kann. Erst bei strikt konvexen Zielfunktionen ist die Existenz einer solchen Lösungsabbildung gesichert. Aber selbst strikt konvexe Funktionen, die schwache Krümmungen aufweisen, bringen oft sehr schlecht konditionierte Gleichungssysteme mit sich, die ebenfalls zu numerischen Schwierigkeiten führen.

Beide Nachteile können ausgeräumt werden, wenn die Pfadverfolgung nicht an der eigentlichen Lösungsmenge ansetzt, sondern an der Menge der Punkte, die ein System analog zu (1.2) mit $\tau > 0$ erfüllen. Allgemeiner wird anstatt eines Skalars sogar ein Vektor $\tau \in \mathbb{R}_{>0}^n$ zugelassen, der individuelle Komplementaritätsstörungen in jeder Komponente erfasst. Von jedem Punkt der KKT-ähnlichen Menge findet dann eine „Restoptimierung“ statt, d.h. die gefundenen Punkte sind Warmstartpunkte, die als Startpunkte für Berechnung der exakten Lösung fungieren. Wird die Störung τ klein gewählt, sind die Pfade in die optimale Lösung sehr kurz.

Auch hier kann gezeigt werden, dass eine differenzierbare Abbildung von Parametern auf KKT-ähnlichen Punkten mit gestörten Komplementaritätsbedingungen unter gegebenen Voraussetzungen existiert. Zusätzlich treten sogar bei allgemein konvexen Zielfunktionen mit Krümmung Null keine singulären Gleichungssysteme auf. Ferner sind unter noch allgemeineren Voraussetzungen differenzierbare Untermanigfaltigkeiten nachweisbar.

Allgemein gilt: Die Systeme sind numerisch umso stabiler, je größer die Störung $\tau \in \mathbb{R}_{>0}^n$ gewählt ist. Bezüglich der Wahl von τ muss demnach ausgelotet werden zwischen der Notwendigkeit stabiler Systeme und dem Wunsch schneller Lösungspfade.

Da die Pfadverfolgung im relativen Inneren der Restriktionsmenge statt findet, ist mehr Platz für Approximationsfehler. Es kann eine Toleranzumgebung konstruiert werden, in denen sich die Warmstartpunkte befinden dürfen. Das verringert die notwendige Anzahl von Korrektorschritten enorm.

Ein ähnliches Prinzip wird von Das und Dennis [DasDen96] verfolgt, die in ihrem Ansatz des *Normal-Boundary Intersection* mit der *konvexen Hülle der individuellen Minima* (CHIM) ebenfalls eine Methode liefern, die von vielversprechenden Startpunkten aus in die Effizienzmenge laufen. Die geometrisch motivierte Herkunft der Startpunkte unterscheidet sich jedoch grundsätzlich von der einer Warmstartstrategie.

Eine Verknüpfung heuristischer Ansätze und Pfadverfolgungstechniken für allgemeine, multikriterielle Optimierungsprobleme entwickelte Schütze et al. [DeSchHe05] unter dem Namen *Subdivision Techniques*.

1.2 Inhaltsübersicht

In Kapitel 2 werden gewisse Standardformen parametrisierter Optimierungsprobleme und deren Lösungsmengen betrachtet. Die KKT-Systeme der parametrisierten Aufgaben können als Nullstellenprobleme betrachtet werden, welche die Lösungsmenge zu allen Parametern implizit beschreibt.

Unter recht allgemeinen Bedingungen kann der Satz A.0.1 über implizite Funktionen angewandt werden und liefert so lokal die Existenz einer eindeutig bestimmten Abbildung von Parametern auf KKT-ähnliche Punkte, die in den Komplementaritätsbedingungen mit einem Parameter τ gestört sind. Die gestörten KKT-Systeme werden als τ -*perturbierte KKT-Systeme* definiert.

In Abschnitt 2.3 wird die Idee des Warmstartpunkts für parametrisierte Optimierungsprobleme abstrahiert, indem Abbildungen (sog. *Warmstartabbildungen*) definiert werden, die approximative Lösungen bei Parameterwechseln liefern. Die resultierenden Werte sind potentielle Warmstartpunkte und werden auf Zulässigkeit und Wiederverwendbarkeit untersucht.

Die sinnvolle Wahl möglicher Parameterwechsel wird in Abschnitt 2.4.1 untersucht. Das Kapitel endet mit der Präsentation eines praktizierbaren Algorithmus, der mit dem Programm MCQIPP (s. Kapitel 5) umgesetzt wurde.

Kapitel 3 besteht aus zwei Teilen. Der erste Teil liefert eine geometrische Betrachtung der Lösungsmenge parametrisierter Optimierungsprobleme. Diese werden als parametrisierte Flächen erkannt. Diese Sichtweise ermöglicht die Charakterisierung singulärer Punkte, die sich durch „Knicke“ in der Fläche bemerkbar machen.

Der zweite Teil behandelt zwei gängige Methoden zur Schrittweitenwahl bei der Pfadverfolgung von Kurven. Es werden die *tangentiale Schrittweitensteuerung* mit Informationen 1. Ordnung und die *zirkuläre Schrittweitenbestimmung* mit Informationen 2. Ordnung vorgestellt. Diese Steuerungen werden bei der Umsetzung der Effizienzmengenberechnung im Programm MCQIPP verwendet.

Kapitel 4 gibt eine kurze Einführung in die multikriterielle Optimierung. Es werden Begriffserklärungen, Optimalitätsbedingungen und die Skalarisierungsmethode eingeführt. Danach wird der Zusammenhang der bisher entwickelten Theorie auf den Fall der multikriteriellen Optimierung übertragen. Weiterhin werden die Gestalt der Lösungsmengen sowie die Charakterisierung von Knicken derselben mit differentialgeometrischen Hilfsmitteln hergeleitet.

In Kapitel 5 wird letztendlich das (C++)-Programm QIPP vorgestellt, ein eigens erstellter IP-Solver basierend auf den *Mehrotra-Algorithmus* für quadratische skalare Optimierungsprobleme. Dieser wird zur „Restoptimierung“ im MATLAB-Programm MCQIPP verwendet, was danach besprochen wird. Die in MCQIPP implementierten Methoden sind nicht optimiert, da die Priorität vorerst in der Funktionalität lag. Das Kapitel schließt mit numerischen Tests an realen Daten aus der Kraftwerkoptimierung und Benchmarks mit einer alternativen Strategie und dem Programm NBI von Das und Dennis [DasDen96] ab.

Teil I

Theorie und ihre Anwendung

Kapitel 2

Parametrisierte Optimierung

Die Motivation zur Einführung in die parametrisierte Optimierung liegt darin, Störungen in Ziel- und Restriktionsfunktionen eines Optimierungsproblems zu modellieren um deren Einfluss auf die optimalen Lösungen zu studieren. Störungen bzw. Perturbationen fließen in Form von Parametern in die jeweiligen Funktionen ein und können danach mit Hilfe von Parameteränderungen dargestellt werden. Aufgrund des funktionalen Zusammenhangs lassen sich somit leicht Stetigkeits- und Differenzierbarkeitsaussagen treffen.

Ziel dieses Kapitels ist die Modellierung perturbierter Optimierungsprobleme sowie der approximativen Bestimmung von Lösungen gestörter Probleme auf der Grundlage bereits bekannter Lösungen. Letztendlich ergibt sich ein Algorithmus zur Bestimmung der gesamten Lösungsmenge eines parametrisierten Optimierungsproblems.

Die Innovation gegenüber herkömmlicher Verfahren liegt in der Betrachtung sogenannter τ -perturbierter Systeme. Anstelle von KKT-Punkten wie bei Hillermeier [Hil00] oder Guddat et al [GuVaJo90], [GuVaTaWe85] wird eine erweiterte Form von KKT-Systemen betrachtet, welche in den Komplementaritätsbedingungen zusätzliche Störungen zulassen. Dies ermöglicht die Deaktivierung aller Ungleichungsrestriktionen und vereinfacht numerische Berechnungen. Das Prinzip stammt aus der Theorie der Inneren-Punkte Methoden und führt zu einem Weg, die Parameterwechsel im stabileren „Inneren“ der Restriktionsmenge zu betrachten.

2.1 Das Standardproblem

Ziel- und Restriktionsfunktionen parametrisierter Optimierungsprobleme hängen zusätzlich von einem Parameter

$$\xi \in \Xi \subset \mathbb{R}^k \quad (2.1)$$

ab. Damit können Parametern ganze Optimierungsprobleme und ferner auch deren Lösungen zugeordnet werden.

Im Rahmen dieser Abhandlung werden diesbezüglich Restriktionsmengen der Art

$$\mathcal{F}(\xi) := \{x \in \mathbb{R}^n \mid h(x; \xi) = 0; g(x; \xi) \leq 0\}, \quad (2.2)$$

mit $h \in \mathcal{C}^2(\mathbb{R}^n \times \Xi, \mathbb{R}^{m_h})$ und $g \in \mathcal{C}^2(\mathbb{R}^n \times \Xi, \mathbb{R}^{m_g})$ betrachtet. Allgemein stellt sich somit für jedes $\xi \in \Xi$ folgendes Optimierungsproblem:

$$\min_{x \in \mathcal{F}(\xi)} f(x; \xi). \quad (P_\xi)$$

Die Gesamtheit der Problematik bezeichnet man als **parametrisiertes Optimierungsproblem**. Hierbei gilt $f \in \mathcal{C}^2(\mathbb{R}^n \times \Xi, \mathbb{R})$.

Bezüglich des Parameters $\xi \in \Xi$ genügt einfache Differenzierbarkeit, soll aber hier der Einfachheit halber vernachlässigt werden.

Entsprechend ergibt sich mit den Lagrangemultiplikatoren $\lambda \in \mathbb{R}^{m_h}$ und $s \in \mathbb{R}^{m_g}$ die Lagrangefunktion $\mathcal{L} : \mathbb{R}^{n+m_h+m_g+k} \rightarrow \mathbb{R}$,

$$\mathcal{L}(x, \lambda, s; \xi) := f(x; \xi) + \lambda^\top h(x; \xi) + s^\top g(x; \xi) \quad (2.3)$$

zum parametrisierten Optimierungsproblem (P_ξ) . Mit einer zusätzlichen Schlupfvariablen $\pi \in \mathbb{R}_{\geq 0}^{m_g}$ lauten damit die KKT-Bedingungen

$$\begin{aligned} D_x \mathcal{L}(x, s, \lambda; \xi)^\top &= 0 \\ h(x; \xi) &= 0 \\ g(x; \xi) + \pi &= 0 \\ S\Pi e &= 0 \\ s, \pi &\geq 0. \end{aligned} \quad (2.4)$$

Dabei repräsentieren - hier als fortan geltende Konvention - $S := \text{diag}(s)$ und $\Pi := \text{diag}(\pi)$ Diagonalmatrizen, deren Diagonale von den gleichnamigen, kleingeschriebenen Vektoren beschrieben werden. Der Vektor e entspreche $(1, \dots, 1)^\top \in \mathbb{R}^n$.

Bezeichnung 2.1.1 Für eine kompaktere Schreibweise, stehe der Vektor

$$w := \begin{pmatrix} x \\ \lambda \\ s \\ \pi \end{pmatrix} \quad (2.5)$$

oder in einer vereinfachten Darstellung $w = (x, \lambda, s, \pi)$. Die Dimensionsgrößen n , m_h und m_g werden zu

$$\tilde{n} := n + m_h + 2m_g \quad (2.6)$$

zusammengefasst, womit $w \in \mathbb{R}^{\tilde{n}}$ gilt.

So ersetzt ferner

$$\mathcal{L} : \mathbb{R}^{\tilde{n}+k} \rightarrow \mathbb{R}, (w, \xi) \mapsto \mathcal{L}(x, \lambda, s; \xi) \quad (2.7)$$

die Funktion $\mathcal{L}(x, \lambda, s; \xi)$ aus (2.3). Die Schlupfvariable π ist zwar redundant, da sie durch die Ungleichungsrestriktionen bereits gegeben ist, jedoch bietet sich diese Schreibweise aus ableitungstechnischen Gründen an (s.u.).

Die Berechnung der Lösungsmengen wird sich - wie sich noch heraus stellt - auf die Ermittlung von KKT-Punkten reduzieren. Dahinter verbergen sich Nullstellenprobleme der Abbildung $F : \mathbb{R}^{\tilde{n}} \times \mathbb{R}^k \rightarrow \mathbb{R}^{\tilde{n}}$,

$$F(w; \xi) := \begin{pmatrix} D_x \mathcal{L}(w; \xi)^\top \\ h(x; \xi) \\ g(x; \xi) + \pi \\ S\Pi e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (2.8)$$

mit $s, \pi \geq 0$.

Für spätere Betrachtungen spielen die **Komplementaritätsbedingungen**

$$S\Pi e = 0 \quad (2.9)$$

eine entscheidende Rolle. Deshalb unterscheidet man zwischen dem Teil

$$r(w; \xi) := \begin{pmatrix} D_x \mathcal{L}(w; \xi) \\ h(x; \xi) \\ g(x; \xi) + \pi \end{pmatrix} \quad (2.10)$$

des Residuums und den Komplementaritätsbedingungen, so dass

$$F(w; \xi) = \begin{pmatrix} r(w; \xi) \\ S\Pi e \end{pmatrix} \quad (2.11)$$

gilt.

Ein Punkt $w^* = (x^*, \lambda^*, s^*, \pi^*)$ erfüllt die **strikten Komplementaritätsbedingungen**, sofern

$$s_i^* > 0 \quad \Leftrightarrow \quad \pi_i^* = 0 \quad (2.12)$$

für alle $i = 1, 2, \dots, m_g$ erfüllt ist.

Das folgende Lemma zeigt, dass unter gewissen Voraussetzungen Parameteränderungen vorgenommen werden können, so dass nach der Änderung wieder dieselben Restriktionen aktiv sind, die vorher aktiv waren. Umgekehrt wechseln auch keine inaktiven Restriktionen zu aktiven. Das Resultat wird später u.a. in der multikriteriellen Optimierung von Nutzen sein (siehe z.B. Satz 4.3.4).

Lemma 2.1.2 *Man betrachte das parametrisierte Optimierungsproblem (P_ξ) . Auf einer Umgebung $U \subset \Xi \subset \mathbb{R}^k$ existiere eine \mathcal{C} -Abbildung*

$$w : U \rightarrow \mathbb{R}^{\tilde{n}},$$

die jedem Parameter $\xi \in U$ einen entsprechenden Punkt $w(\xi)$ zuteilt, so dass

$$F(w(\xi); \xi) = 0$$

gilt.

Ist $w^* = w(\xi^*)$, $\xi^* \in U$ ein KKT-Punkt des Problems (P_{ξ^*}) , in dem die strikten Komplementaritätsbedingungen gelten, so existiert eine Umgebung $V \subseteq U$ mit der das Bild der Abbildung $w|_V$ wiederum aus KKT-Punkten besteht, für die wiederum – wie in w^* – dieselben Restriktionen aktiv bzw. inaktiv sind und demnach die strikten Komplementaritätsbedingungen erhalten bleiben.

Beweis:

Die Abbildung w lässt sich in weitere Abbildungen s, λ und π unterteilen, derartig dass $w(\xi) = (x(\xi), \lambda(\xi), s(\xi), \pi(\xi))$ gilt.

Ist $w^* = (x^*, \lambda^*, s^*, \pi^*)$ KKT-Punkt, so erfüllt er zum einen die Bedingung $F(w^*; \xi^*) = 0$ als auch $s^*, \pi^* \geq 0$. Aufgrund der strikten Komplementaritätsbedingungen existieren zusätzlich Indexmengen

$$J_s(\xi) := \{i = 1, 2, \dots, m_g \mid s_i(\xi) > 0\} \quad (2.13)$$

und

$$J_\pi(\xi) := \{1, 2, \dots, m_g\} \setminus J_s(\xi), \quad (2.14)$$

so dass die Einteilung

$$\begin{aligned} s_i^* &> 0, \quad \pi_i^* = 0 \quad \forall i \in J_s(\xi^*) \\ s_i^* &= 0, \quad \pi_i^* > 0 \quad \forall i \in J_\pi(\xi^*) \end{aligned}$$

vorgenommen werden kann.

Aufgrund der strikt positiven s_i^* , $i \in J_s(\xi^*)$ und π_j^* , $j \in J_\pi(\xi^*)$ sowie der Stetigkeit der Abbildung w existiert eine Umgebung W von ξ^* , welche die Indexmengen für alle $\xi \in W$ erhält, d.h. $J_s(\xi) = J_s(\xi^*)$ und $J_\pi(\xi) = J_\pi(\xi^*)$.

Für alle Parameter ξ aus $V = W \cap U$ sind somit die KKT-Bedingungen 2.4 und gleichzeitig die strikten Komplementaritätsbedingungen erfüllt.

□

2.2 Stabilität der Lösungsmenge

Wie bereits erwähnt lässt sich jedem Parameter $\xi \in \Xi$ das Optimierungsproblem (P_ξ) zuordnen und damit auch deren Lösungsmenge. Sofern jedem Parameter ξ ein eindeutiger KKT-Punkt und damit eine eindeutige Lösung der Funktion F zugeordnet werden kann, lassen sich die Menge der KKT-Punkte als Abbildung

$$w : \Xi \rightarrow \mathbb{R}^{\tilde{n}}, \quad \xi \mapsto w(\xi) \quad (2.15)$$

beschreiben. Gleichzeitig ergeben sich damit entsprechende Funktionen $x : \Xi \rightarrow \mathbb{R}^n$, $\lambda : \Xi \rightarrow \mathbb{R}^{m_h}$, $s : \Xi \rightarrow \mathbb{R}^{m_g}$ und $\pi : \Xi \rightarrow \mathbb{R}^n$.

Bemerkung 2.2.1 *Ob die Größen w , x , λ , s und π als Vektoren eines \mathbb{R} -Vektorraums oder als \mathcal{C}^1 -Abbildungen aufgefasst werden, hängt vom Blickwinkel ab. Sofern eine Unterscheidung nicht aus dem Kontext ersichtlich ist, wird diese - wenn nötig - im weiteren Verlauf vorgenommen.*

Ansonsten, wenn keine eindeutige Zuordnung möglich ist, spricht man von mengenwertigen Abbildungen (*engl.: Point-To-Set Maps*), die unter anderem in [SawNakTan85] ausführlich behandelt werden.

Im Folgenden werden Voraussetzungen hergeleitet, welche die Existenz einer Lösungsabbildung w garantieren. Darüber hinaus werden für w Stetigkeits- und Differenzierbarkeitsaussagen getroffen.

2.2.1 Stabilität für nichtlineare Probleme

Die Existenz und Wohldefiniertheit der Abbildung w aus (2.15) kann sehr gut mit Hilfe des Theorems A.0.1 über implizite Funktionen gezeigt werden. Sobald die Funktionalmatrix $D_w F(\hat{w}; \hat{\xi})$ an einer Nullstelle $(\hat{w}, \hat{\xi}) \in \mathbb{R}^{\tilde{n}} \times \Xi$ von F regulär ist, liefert das Theorem die Existenz und Differenzierbarkeit der Abbildung w . Gleichzeitig ergibt sich durch Anwendung der Kettenregel das partielle Differential $D_\xi w$ an der Stelle $\hat{\xi}$ als Lösung des Gleichungssystems

$$D_w F(\hat{w}; \hat{\xi}) \cdot D_\xi w(\hat{\xi}) = -D_\xi F(\hat{w}; \hat{\xi}). \quad (2.16)$$

Speziell für (P_ξ) gilt dabei

$$D_w F(w; \xi) = \begin{pmatrix} D_{xx} \mathcal{L}(w; \xi) & D_x h(x; \xi)^\top & D_x g(x; \xi)^\top & 0 \\ D_x h(x; \xi) & 0 & 0 & 0 \\ D_x g(x; \xi) & 0 & 0 & I \\ 0 & 0 & \Pi & S \end{pmatrix} \quad (2.17)$$

und

$$D_\xi F(w; \xi) = \begin{pmatrix} D_{\xi x} \mathcal{L}(w; \xi) \\ D_\xi h(x; \xi) \\ D_\xi g(x; \xi) \\ 0 \end{pmatrix}. \quad (2.18)$$

Somit liegt bei gegebener Nullstelle die Schwierigkeit darin, zu zeigen, dass die Matrix in (2.17) regulär ist, womit zumindest lokal die Abbildung w existiert.

1968 veröffentlichten Fiacco und McCormick dazu in [FiaMcCo90], Section 3.4, Theorem 6 ein Theorem, aus dem die Existenz eines stetig-differenzierbaren KKT-Tripels $(x(\xi), s(\xi), \lambda(\xi))$ für allgemeine nichtlineare Optimierungsprobleme mit linearen Perturbationen resultierte. Sowohl Robinson [Rob74] als auch Fiacco [Fia76] erweiterten die Aussagen für allgemeinere Perturbationen, wie der folgende Satz zeigt ([Fia83], Section 2.4, Theorem 2.4.4):

Satz 2.2.2 (Fiacco und Robinson) *Sei x^* eine optimale Lösung des Optimierungsproblems $(P_{\xi=0})$ mit entsprechenden Lagrangemultiplikatoren (λ^*, s^*) . Falls für*

- (a) $f, g, h \in C^2$ in (x, ξ) innerhalb einer Umgebung um $(x^*, 0)$
- (b) die hinreichende Optimalitätsbedingung 2. Ordnung in (x^*, λ^*, s^*) ,

(c) die LICQ in x^*

(d) sowie die strikten Komplementaritätsbedingungen erfüllt sind,

dann

(i) ist x^* isolierte, lokale Lösung des Problems $(P_{\xi=0})$ mit eindeutigen Lagrange-multiplikatoren (λ^*, s^*) ,

(ii) existiert eine Nullumgebung $U(0) \subset \Xi$, so dass für alle $\xi \in U(0)$ die die Werte der eindeutig bestimmten C^1 -Abbildung $w(\xi) := (x(\xi), \lambda(\xi), s(\xi))$ der hinreichenden Optimalitätsbedingungen 2. Ordnung für das Problem (P_ξ) genügen. Ferner gilt $w(0) = (x^*, \lambda^*, s^*)$ und $x(\xi)$ bildet eine isolierte, lokale Lösung des Problems P_ξ mit eindeutigen Lagrangemultiplikatoren $(\lambda(\xi), s(\xi))$.

(iii) Außerdem gelten für alle Punkte $w(\xi)$, $\xi \in U(0)$ die LICQ und strikte Komplementaritätsbedingung.

Für quadratische Optimierungsprobleme

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \frac{1}{2}x^\top Qx + c^\top x \\ \text{N.B.} \quad & Ax = b \\ & Cx \leq c_u, \end{aligned} \tag{2.19}$$

wobei Q positiv definit und symmetrisch ist, veröffentlichte Daniel im Jahr 1973 ein Theorem, das die Auswirkungen kleiner Perturbationen in der Zielfunktion auf die optimale Lösung durch eine Abschätzung beschrieb.

Satz 2.2.3 (Daniel) Seien x^* optimale Lösung von 2.19 und x optimale Lösung des Problems mit der perturbierten Zielfunktion $\frac{1}{2}x^\top(Q + \Delta Q)x + (c + \Delta c)^\top x$. Sei weiterhin $\delta := \max\{\|\Delta Q\|, \|\Delta c\|\}$, dann gilt mit $\kappa > 0$ als kleinsten Eigenwert von Q die Ungleichung

$$\|x - x^*\| \leq \frac{\delta}{\kappa - \delta}(1 + \|x^*\|), \tag{2.20}$$

falls $\delta < \kappa$.

Jittorntrum [Jit81] veröffentlichte 1981 weitgreifendere Resultate für allgemeine, nichtlineare Optimierungsprobleme, die zusätzlich Abschätzungen - wie in (2.20) - für Lagrangemultiplikatoren lieferten.

Lipschitzstetigkeit für strikt konvexe Zielfunktionen, konvexen Nebenbedingungen und konvex kompakten Parametermengen werden u.a. von Hagen [Hag79] und Malanowski [Mal87] bewiesen.

2.2.2 Stabilität τ -perturbierter KKT-Systeme

Die dieser Abhandlung zugrunde liegende Problematik entsteht durch die Aufgabe der Approximation der Lösungsmenge parametrisierter, konvexer Optimierungsprobleme. Es wurde bereits gezeigt, dass die Lösungsmenge unter bestimmten Voraussetzungen mit Hilfe einer differenzierbaren Abbildung w beschrieben werden kann, die allerdings nur durch die jeweilige Ableitung eines Punktes erklärt ist. Diese Ableitungen werden mit Hilfe von Taylorentwicklungen zur Approximation der Abbildung w und damit zur Approximation der Lösungsmenge genutzt.

Unter diesem Gesichtspunkt werden nun drei Aspekte genannt, welche den neuen Begriff des τ -perturbierten KKT-Systems motivieren sollen.

1. Durch Approximationen mit Ableitungen können aktive Ungleichungsrestriktionen sehr schnell verletzt werden und so zu unlässigen Punkten führen, mit denen eine Approximation der Lösungsmenge erschwert oder gar praktisch unmöglich wird.
2. Selbst bei Einhaltung der Ungleichungsrestriktionen, entstünden durch die Approximation der Lösungsmenge im Allgemeinen Residuen, so dass sich mit jedem neu gewonnenen Punkt auch eine neue Nullstellenabbildung F ergäbe. In der Nähe der Lösung können diese Residuen numerisch meist nicht mehr kompensiert werden.
3. Zum anderen besitzen allgemein konvexe Optimierungsprobleme - wie z.B. (2.19) mit lediglich positiv semidefiniter Matrix Q - i.A. keine Abbildung w , da die Lösungsmenge von P_ξ zu einem festen Parameter ξ nicht eindeutig sein muss.

Ein Prinzip der Inneren-Punkte Methoden (Stichwort: **Zentraler Pfad**) kann hier Abhilfe schaffen. Die Komplementaritätsbedingungen des KKT-Systems werden perturbiert. Das folgende System (2.22) erweitert das KKT-System für folgende Untersuchungen.

Definition 2.2.4 *Werden die zum parametrisierten Optimierungsproblem (P_ξ) gehörenden Komplementaritätsbedingungen (2.9) positiv gestört, d.h.*

$$S\Pi e = \tau, \quad \tau \in \mathbb{R}_{\geq 0}^{m_g}, \quad (2.21)$$

*so heißt das System τ -perturbiertes **KKT-System**.*

Das entsprechende Nullstellenproblem der Abbildung $F : \mathbb{R}^{\tilde{n}} \times \Xi \times \mathbb{R}_{\geq 0}^{m_g} \rightarrow \mathbb{R}^{\tilde{n}}$ lautet dann

$$F(w; \xi, \tau) := \begin{pmatrix} D_x \mathcal{L}(w; \xi)^\top \\ h(x; \xi) \\ g(x; \xi) + \pi \\ S\Pi e - \tau \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (2.22)$$

mit $s, \pi \geq 0$.

Definition 2.2.5 Der Punkt $w = (x, \lambda, s, \pi) \in \mathbb{R}^{\tilde{n}}$ heißt τ -zulässig bezüglich des Optimierungsproblems (P_ξ) , sofern Bedingung (2.22), $\pi \geq 0$ und $s \geq 0$ erfüllt sind.

Bemerkung 2.2.6

Für Lösungen τ -perturbierter KKT-Systeme mit $\tau > 0$ sind alle Ungleichungsrestriktionen automatisch inaktiv, da für die Schlupfvariable $\pi > 0$ gelten muss. Entsprechend ist auch s strikt positiv.

Definition 2.2.7 Die Menge

$$\mathcal{N}_F(\Xi, \tau) := \{(w; \xi) \mid w \text{ ist } \tau\text{-zulässig bez. } (P_\xi), \xi \in \Xi\}$$

heißt τ -**Urbildmenge** zum parametrisierten Optimierungsproblem (P_ξ) . Die Menge $\mathcal{N}_F(\Xi, \tau)$ ist **vollständig gegeben**, falls zu jedem $\xi \in \Xi$ ein $(w; \xi) \in \mathcal{N}_F(\Xi, \tau)$ existiert.

Ob $\mathcal{N}_F(\Xi, \tau)$ vollständig gegeben ist, hängt u.a. von der Wahl der Parametermenge Ξ bezüglich des Parameters τ ab. Lemma 2.2.8 liefert ein hinreichendes Kriterium dafür, wann $\mathcal{N}_F(\Xi, \tau)$ zumindest lokal vollständig gegeben ist.

Lemma 2.2.8 Gegeben sei ein parametrisiertes Optimierungsproblem (P_ξ) mit den \mathcal{C}^2 -Abbildungen f, g und h sowie einem Parameter-Tupel $\nu = (\xi, \tau) \in \mathbb{R}^k \times \mathbb{R}_{> 0}^{m_g}$ und einem Punkt $w_\nu = (x_\nu, \lambda_\nu, s_\nu, \pi_\nu) \in \mathbb{R}^{\tilde{n}}$. Es sei w_ν τ -zulässig bezüglich (P_ξ) .

Darüber hinaus haben die zugehörigen Funktionalmatrizen $D_x g(x_\nu; \xi) \in \mathbb{R}^{m_g \times n}$ vollen Spaltenrang, $D_x h(x_\nu; \xi) \in \mathbb{R}^{m_g \times n}$ vollen Zeilenrang und die Hessematrix der Lagrangefunktion $D_{xx} \mathcal{L}(w_\nu; \xi) \in \mathbb{R}^{n \times n}$ sei positiv semidefinit.

Dann existieren

- a) Umgebungen $U(\nu)$ von ν und $U(w_\nu)$ von w_ν sowie eine eindeutig bestimmte, stetig-differenzierbare Abbildung

$$w : U(\nu) \rightarrow U(w_\nu), \quad v \mapsto w(v),$$

so dass $F(w(v), v) = 0$ gilt, für alle $v \in U(\nu)$,

- b) für festes τ eine Umgebung $U(\xi)$ von ξ und eine eindeutig bestimmte, stetig und nach ξ differenzierbare Abbildung

$$w(\cdot, \tau) : U(\xi) \rightarrow U(w_\nu), \quad y \mapsto w(y, \tau),$$

so dass $F(w(y), y) = 0$ gilt, für alle $y \in U(\xi)$, und $\mathcal{N}_F(U(\xi), \tau)$ vollständig gegeben ist.

Beweis:

Zu a): Es seien $\nu = (\xi, \tau) \in \mathbb{R}^k \times \mathbb{R}_{>0}^{m_g}$ und $w_\nu = (x_\nu, \lambda_\nu, s_\nu, \pi_\nu) \in \mathbb{R}^{\tilde{n}}$, τ -zulässig bez. eines beliebig gewählten Parameters ξ .

Behauptung (a) folgt sofort aus dem Satz über implizite Funktionen (Satz A.0.1), wenn die Matrix $D_w F(w_\nu; \nu)$ regulär ist und damit das Gleichungssystem

$$D_w F(w_\nu; \nu) \cdot D_\nu w(\nu) = -D_\nu F(w_\nu; \nu) \quad (2.23)$$

eine eindeutige Lösung $D_\nu w(\nu)$ besitzt.

Die Matrix $D_w F(w_\nu; \nu)$ entspricht weiterhin der Funktionalmatrix aus (2.17), da τ diesbezüglich keine Auswirkungen hat. Es bleibt zu zeigen, dass $D_w F(w; \nu)$ regulär ist.

Zur Vereinfachung der Notation, setze man

$$\begin{aligned} L &:= D_{xx} \mathcal{L}(w_\nu; \xi), & A &:= D_x h(x_\nu; \xi), & C &:= D_x g(x_\nu; \xi), \\ x' &:= D_\nu x(\nu), & \lambda' &:= D_\nu \lambda(\nu), & s' &:= D_\nu s(\nu) & \pi' &:= D_\nu \pi(\nu) \\ r_L &:= D_{\nu x} \mathcal{L}(w_\nu; \xi), & r_A &:= D_\nu h(x_\nu; \xi), & r_C &:= D_\nu g(x_\nu; \xi) & r_{S\Pi} &:= D_\nu (S_\nu \Pi_\nu e - \tau). \end{aligned}$$

Es ist dann

$$D_w F(w_\nu; \nu) = \begin{pmatrix} L & A^\top & C^\top & 0 \\ A & 0 & 0 & 0 \\ C & 0 & 0 & I \\ 0 & 0 & \Pi_\nu & S_\nu \end{pmatrix}.$$

Wegen des zusätzlichen Parameters τ ergibt sich im Vergleich zu (2.18) die rechte Seite

$$D_\nu F(w_\nu; \nu) = \begin{pmatrix} D_{\xi x} \mathcal{L}(w_\nu; \xi) & 0 \\ D_\xi h(x_\nu; \xi) & 0 \\ D_\xi g(x_\nu; \xi) & 0 \\ 0 & -I \end{pmatrix} = \begin{pmatrix} r_L \\ r_A \\ r_C \\ r_{S\Pi} \end{pmatrix}$$

Insgesamt erhält man äquivalent zu (2.23) das Gleichungssystem

$$\begin{pmatrix} L & A^\top & C^\top & 0 \\ A & 0 & 0 & 0 \\ C & 0 & 0 & I \\ 0 & 0 & \Pi_\nu & S_\nu \end{pmatrix} \begin{pmatrix} x' \\ \lambda' \\ s' \\ \pi' \end{pmatrix} = - \begin{pmatrix} r_L \\ r_A \\ r_C \\ r_{S\Pi} \end{pmatrix}. \quad (2.24)$$

Es wird nun versucht, das System nach und nach aufzulösen.

Als erstes wird π' eliminiert. Nach Voraussetzung wähle man w_ν , so dass $F(w_\nu; \nu) = 0$ und damit $S_\nu \Pi_\nu e - \tau = 0$ erfüllt sein müssen. Da $\tau \in \mathbb{R}_{>0}^{m_g}$ gilt und aufgrund der Positivitätsbedingung $s_\nu, \pi_\nu \geq 0$ können die Komponenten der Vektoren s_ν und π_ν somit nur strikt positiv sein (s. Bemerkung 2.2.6). Die Diagonalmatrizen S_ν und Π_ν sind folglich positiv definit und damit regulär.

Ferner gilt wegen $S_\nu \pi' + \Pi_\nu s' = -r_{S\Pi}$ das Ergebnis

$$\pi' = -S_\nu^{-1} (r_{S\Pi} + \Pi_\nu s').$$

Unter Ausnutzung dieser Information reduziert sich das Gleichungssystem (2.24) somit auf das System

$$\begin{pmatrix} L & A^\top & C^\top \\ A & 0 & 0 \\ C & 0 & -S_\nu^{-1} \Pi_\nu \end{pmatrix} \begin{pmatrix} x' \\ \lambda' \\ s' \end{pmatrix} = - \begin{pmatrix} r_L \\ r_A \\ r_C - S_\nu^{-1} r_{S\Pi} \end{pmatrix}$$

Analog kann auch $s' = \Pi_\nu^{-1} S_\nu (r_C - S_\nu^{-1} r_{S\Pi}) + \Pi_\nu^{-1} S_\nu C x'$ beschrieben werden und man erhält

$$\begin{pmatrix} L + C^\top \Pi_\nu^{-1} S_\nu C & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} x' \\ \lambda' \end{pmatrix} = - \begin{pmatrix} r_L + C^\top \Pi_\nu^{-1} S_\nu (r_C - S_\nu^{-1} r_{S\Pi}) \\ r_A \end{pmatrix}.$$

Da C vollen Spaltenrang besitzt ist $C^\top \Pi_\nu^{-1} S_\nu C$ positiv definit. Nach Voraussetzung ist L positiv semidefinit und so die Matrix

$$Q := L + C^\top \Pi_\nu^{-1} S_\nu C$$

regulär. Dann gilt aber mit

$$r := - (r_L + C^\top \Pi_\nu^{-1} S_\nu (r_C - S_\nu^{-1} r_{S\Pi}))$$

die Gleichung

$$x' = Q^{-1}r - Q^{-1}A^\top \lambda'.$$

Einsetzen in das Gleichungssystem $Ax' = -r_A$ liefert

$$AQ^{-1}A^\top \lambda' = r_A + AQ^{-1}r$$

und da A vollen Zeilenrang besitzt, ist die Matrix $AQ^{-1}A^\top$ invertierbar. Das obige Gleichungssystem kann nach λ' aufgelöst werden.

Damit ist die Funktionalmatrix $D_w F(w; \nu)$ regulär, was zu zeigen war.

Zu b): Auch für ein festes τ kann der Satz über implizite Funktionen bezüglich des Parameters ξ angewendet werden, so dass wie im Beweis zu (a) verfahren werden kann, womit zunächst eine Umgebung V_ξ existiert, so dass die Bildmenge $w(V_\xi)$ mit Parametern aus V_ξ Nullstellen von F bildet. Die Bildelemente sind im Allgemeinen jedoch nicht alle τ -zulässig, da die Bilder der Teilabbildungen s und π auch negativ werden können. Es reicht allerdings die offene Menge

$$U(\xi) := V_\xi \cap \{\phi \in V_\xi \mid w(\phi) \text{ ist } \tau\text{-zulässig}\} \neq \emptyset$$

zu betrachten. Dann ist $\mathcal{N}_F(U(\xi), \tau)$ vollständig gegeben.

□

Bemerkung 2.2.9 *Alternativ muss die Matrix C aus dem Beweis zu Lemma 2.2.8 nicht unbedingt vollen Spaltenrang besitzen. Sofern L sogar positiv definit ist, bleibt die Regularität der Funktionalmatrix $D_w F(w; \nu)$ weiter erhalten.*

2.3 Approximierende Lösungen

2.3.1 Die Warmstartabbildung

Im vorherigen Abschnitt wurden Existenz- und Eindeutigkeitsaussagen für eine stetig-differenzierbare Abbildung w getroffen, die Parameter $\xi \in \Xi$ und zuletzt auch

$\tau \in \mathbb{R}_{\geq 0}^{m_g}$ auf Lösungen des τ -perturbierten Systems, d.h. $(w(\xi, \tau), \xi) \in \mathcal{N}_F(\Xi, \tau)$, abbildet.

In diesem Abschnitt werden gegebene Punkte $(w, \xi) \in \mathcal{N}_F(\Xi, \tau)$ und damit bestehende Ableitungsinformationen erster Ordnung für eine approximative Bestimmung weiterer Nullstellen von F aus (2.22) genutzt, die durch einen Parameterwechsel von ξ auf $\tilde{\xi} \in \Xi$ bei festem τ beschrieben werden.

Dies führt zur Definition der Warmstartabbildung bzw. des Warmstartpunktes, deren Namensgebung sich durch spätere Anwendungen zur Berechnung der Lösungsmenge der Optimierungsprobleme P_ξ zu Parametern $\xi \in \Xi$ motiviert, in denen die ermittelten Warmstartpunkte als Startpunkte eines Optimierungsverfahrens genutzt werden (s. Algorithmus 2.4.9).

Der Parameter τ wird fest gewählt sein, deshalb benutze man anstelle der Nullstellenabbildung aus (2.22) kurz

$$F_\tau : \mathbb{R}^{\tilde{n}+k} \rightarrow \mathbb{R}^{\tilde{n}}, \quad (w; \xi) \mapsto F(w; \xi, \tau). \quad (2.25)$$

Da sich die Abbildungswerte von $F_\tau(w; \xi)$ und $F(w; \xi)$ aus (2.8) lediglich durch die perturbierten Komplementaritätsbedingungen unterscheiden, kann auch hier die Residuenabbildung $r(w; \xi)$ aus (2.10) genutzt werden, so dass gleichermaßen

$$F_\tau(w; \xi) = \begin{pmatrix} r(w; \xi) \\ S\Pi e - \tau \end{pmatrix} \quad (2.26)$$

gilt.

Definition 2.3.1 *Gegeben seien die Menge $\Xi \subset \mathbb{R}^k$ sowie die Umgebung $U \subseteq \Xi$, so dass für ein festes $\tau \in \mathbb{R}_{\geq 0}^{m_g}$ die τ -Urbildmenge $\mathcal{N}_F(U, \tau)$ vollständig gegeben ist. Dazu existiere eine eindeutig bestimmte, stetig-differenzierbare Abbildung $w : U \rightarrow \mathbb{R}^{\tilde{n}}$, welche τ -zulässige Werte erzeuge.*

Eine Abbildung

$$\rho : U \times \Xi \rightarrow \mathbb{R}^{\tilde{n}} \quad (2.27)$$

heißt **Warmstartabbildung**, falls sie die Bedingungen

$$F_\tau(\rho(\xi, \tilde{\xi}); \tilde{\xi}) = o\left(\|\tilde{\xi} - \xi\|\right), \quad (2.28)$$

für alle $\xi, \tilde{\xi} \in U$, $\tilde{\xi} \rightarrow \xi$ erfüllt.

Der Wert $v := (x, \lambda, s, \pi) := \rho(\xi, \tilde{\xi})$ heißt **Warmstartpunkt** zum Parameter $\tilde{\xi}$ bezüglich ξ .

Ein Warmstartpunkt v heißt **eigentlich**, wenn der Wert der Residuenabbildung (2.10) dem Nullvektor entspricht, d.h. $r(v; \tilde{\xi}) = 0$.

Ein Warmstartpunkt heißt **zulässig**, falls zusätzlich

$$s \geq 0 \text{ und } g(x, \tilde{\xi}) \leq 0 \quad (2.29)$$

gilt.

Die Abbildung

$$\Delta w : U \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}, \quad (\xi, \tilde{\xi}) \mapsto \rho(\xi, \tilde{\xi}) - w(\xi) \quad (2.30)$$

ist die zugehörige **Differenzenabbildung**. Deren Abbildungswerte werden **Differenzenvektoren** genannt.

Es folgen nun zwei Beispiele für Warmstartabbildungen, die auch im weiteren Verlauf dieses Kapitels ihre Anwendung finden.

Beispiel 2.3.2 Für dieses Beispiel sei

$$F_\tau : \mathbb{R}^{\tilde{n}} \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}$$

Lipschitzstetig bez. des ersten Arguments mit Lipschitzkonstante L . Ferner existiere auf einer Umgebung U die Abbildung

$$w \in \mathcal{C}^1(U, \mathbb{R}^{\tilde{n}}),$$

so dass für alle $\xi \in U$ die Gleichung $F_\tau(w(\xi); \xi) = 0$ erfüllt ist.

Für beliebig gewählte Parameter ξ und $\tilde{\xi}$ aus U kann mit Hilfe der Taylorentwicklung

$$w(\tilde{\xi}) = w(\xi) + D_\xi w(\xi) \cdot \Delta\xi + o(\|\Delta\xi\|), \quad (2.31)$$

wobei $\Delta\xi = \tilde{\xi} - \xi$ gilt, die Warmstartabbildung

$$\rho_w(\xi, \tilde{\xi}) := w(\xi) + D_\xi w(\xi) \cdot \Delta\xi. \quad (2.32)$$

und die Differenzenabbildung

$$\Delta w_{\rho_w}(\xi, \tilde{\xi}) := D_\xi w(\xi) \cdot \Delta\xi \quad (2.33)$$

definiert werden.

Denn mit $v := \rho_w(\xi, \tilde{\xi})$ führt die Folgerung

$$\begin{aligned} \left\| F_\tau(v; \tilde{\xi}) \right\| &\stackrel{F_\tau(w(\tilde{\xi}); \tilde{\xi})=0}{=} \left\| F_\tau(w(\tilde{\xi}); \tilde{\xi}) - F_\tau(v; \tilde{\xi}) \right\| \\ &\stackrel{L\text{-Stetig}}{\leq} L \left\| w(\tilde{\xi}) - v \right\| \\ &\stackrel{(2.31)}{=} o(\|\Delta\xi\|). \end{aligned}$$

zu Bedingung (2.28).

Die Abbildung w ist nur lokal auf U definiert; eine Einschränkung des ersten Arguments auf U ist notwendig. Hinsichtlich des zweiten Arguments ist eine Fortsetzung auf $\Xi \supset U$ möglich und sinnvoll, so dass $\rho_w : U \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}$ betrachtet wird. Die Berücksichtigung der Fortsetzung ist wichtig, da die Umgebung U im Allgemeinen nicht bekannt ist und somit angesetzte Parameterwechsel besser erfasst werden können. Per Definition ist die Warmstartabbildung ρ_w weiterhin wohldefiniert, wenn das zweite Argument aus $\Xi \subseteq \mathbb{R}^k$ stammt.

Beispiel 2.3.3 Es seien die Größen und Voraussetzungen aus Beispiel 2.3.2 gegeben. Diesmal gelte zusätzlich $F_\tau \in C^1$ und für beliebige Parameter ξ und $\tilde{\xi}$ aus U betrachte man die Taylorentwicklung

$$F_\tau(w(\tilde{\xi}); \tilde{\xi}) = F_\tau(w(\xi); \tilde{\xi}) + D_w F_\tau(w(\xi); \tilde{\xi}) \cdot (w(\tilde{\xi}) - w(\xi)) + R,$$

wobei $R = o(\|w(\tilde{\xi}) - w(\xi)\|)$ das Taylorrestglied beschreibt. Sofern die Funktionalmatrix $D_w F_\tau(w(\xi); \tilde{\xi})$ regulär ist, folgt

$$(w(\tilde{\xi}) - w(\xi)) \stackrel{F_\tau(w(\tilde{\xi}); \tilde{\xi})=0}{=} -D_w F_\tau(w(\xi); \tilde{\xi})^{-1} \left(F_\tau(w(\xi); \tilde{\xi}) + R \right)$$

für $\Delta\xi = \tilde{\xi} - \xi$.

Dann setze man

$$\rho_F(\xi, \tilde{\xi}) := w(\xi) - D_w F_\tau(w(\xi); \tilde{\xi})^{-1} \cdot F_\tau(w(\xi); \tilde{\xi}) \quad (2.34)$$

als Warmstartabbildung mit der Differenzenabbildung

$$\Delta w_{\rho_F}(\xi, \tilde{\xi}) := -D_w F_\tau(w(\xi); \tilde{\xi})^{-1} \cdot F_\tau(w(\xi); \tilde{\xi}), \quad (2.35)$$

wenn Bedingung (2.28) erfüllt ist. Dies ist unter anderem mit den folgenden Voraussetzungen der Fall:

- a) F_τ ist Lipschitz bez. des ersten Arguments mit Lipschitzkonstante L_F und w ebenfalls Lipschitzstetig mit Konstante L_w . Für $U \subset \Xi$ gebe es eine offene Menge $V \supset \bar{U}$, so dass $w \in \mathcal{C}^2(V, \mathbb{R}^{\tilde{n}})$,
- b) die Parametermengen \bar{U} und Ξ seien kompakt und
- c) die Funktionalmatrix $F_\tau(w(\xi); \tilde{\xi})$ sei regulär für alle $\xi \in \bar{U}$ und $\tilde{\xi} \in \Xi$.

Es gelte $v := \rho_F(\xi, \tilde{\xi})$. Um zu zeigen, dass ρ_F eine Warmstartabbildung ist, betrachte man zuerst

$$\begin{aligned} \left\| F_\tau(v; \tilde{\xi}) \right\| &\stackrel{F_\tau(w(\tilde{\xi}); \tilde{\xi})=0}{=} \left\| F_\tau(w(\tilde{\xi}); \tilde{\xi}) - F_\tau(v; \tilde{\xi}) \right\| \\ &\stackrel{L\text{-Stetig}}{\leq} L_F \cdot \left\| w(\tilde{\xi}) - v \right\| \\ &\leq L_F \cdot \left\| D_w F_\tau(w(\xi); \tilde{\xi})^{-1} R \right\| \end{aligned}$$

Da $\left\| D_w F_\tau(w(\xi); \tilde{\xi})^{-1} \right\|$ stetig auf dem Kompaktum $\bar{U} \times \Xi$ ist, wobei $\xi \in \bar{U}$ und $\tilde{\xi}$, kann

$$D_F := \max_{v \in \Xi, u \in \bar{U}} \left\| D_w F_\tau(w(u); v)^{-1} \right\|$$

definiert werden. Mit $\tilde{L}_F := L_F \cdot D_F$ lässt sich

$$\left\| F_\tau(v; \tilde{\xi}) \right\| \leq \tilde{L}_F \cdot \|R\| = o\left(\left\| w(\tilde{\xi}) - w(\xi) \right\|\right)$$

festhalten.

Bleibt zu zeigen, dass $o\left(\left\| w(\tilde{\xi}) - w(\xi) \right\|\right) = o(\|\Delta\xi\|)$ gilt. Die Gleichheit ist jedoch aufgrund der Lipschitzstetigkeit von w gegeben, denn es gilt die Abschätzung

$$\left(\left\| w(\tilde{\xi}) - w(\xi) \right\|\right) \leq L_w \|\Delta\xi\|.$$

Somit ist $\rho_F : U \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}$ eine Warmstartabbildung. Wegen der Entwicklung in F_τ entspricht ρ_F den Newton-Schritten der Abbildung $F_\tau(\cdot, \tilde{\xi}) : \mathbb{R}^{\tilde{n}} \rightarrow \mathbb{R}^{\tilde{n}}$ zur Reduzierung des Residuums $F_\tau(w(\xi); \tilde{\xi}) \neq 0$.

Korollar 2.3.4 (Existenz von ρ_w) Gegeben sei das parametrisierte Optimierungsproblem (P_ξ) . Zu einem $\tau \in \mathbb{R}_{\geq 0}^{m_g}$ existiere ein bezüglich ξ τ -zulässiger Punkt $(w_\xi, \xi) \in \mathbb{R}^{\tilde{n}} \times \mathbb{R}^k$, so dass $D_w F_\tau(w_\xi; \xi)$ regulär ist. Dann existiert auch eine Umgebung U_ξ von ξ sowie eine eindeutig bestimmte Abbildung $w \in \mathcal{C}^1(U_\xi, \mathbb{R}^{\tilde{n}})$

mit τ -zulässigen Werten für alle $\xi \in U_\xi$. D.h. die τ -Urbildmenge $\mathcal{N}_F(U_\xi, \tau)$ ist vollständig gegeben, ebenso die Warmstartabbildung

$$\rho_w : U_\xi \times \mathbb{R}^k \rightarrow \mathbb{R}^{\tilde{n}} \quad (2.36)$$

aus Beispiel (2.3.2).

Beweis:

Folgt sofort aus Satz A.0.1 über implizite Funktionen und Beispiel (2.3.2), denn wegen der regulären Funktionalmatrix gilt

$$D_\xi w(\xi) = D_w F_\tau(w_\xi, \xi)^{-1} D_\xi F_\tau(w_\xi, \xi).$$

□

Korollar 2.3.5 (Existenz von ρ_F) Gegeben sei das parametrisierte Optimierungsproblem (P_ξ) . Zu einem $\tau \in \mathbb{R}_{\geq 0}^{m_g}$ existiere ein bezüglich ξ τ -zulässiger Punkt $(w_\xi, \xi) \in \mathbb{R}^{\tilde{n}} \times \mathbb{R}^k$, so dass $D_w F_\tau(w; \xi)$ regulär ist. Dann existiert auch eine Umgebung V_ξ von ξ sowie eine eindeutig bestimmte Abbildung $w \in \mathcal{C}^1(V_\xi, \mathbb{R}^{\tilde{n}})$ mit τ -zulässigen Werten für alle $\xi \in V_\xi$.

Darüber hinaus seien die Voraussetzungen (a)-(c) aus Beispiel 2.3.3 erfüllt. Damit existieren die kompakten Mengen Ξ und \bar{U} und die Warmstartabbildung

$$\rho_F : V_\xi \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}, \quad \bar{U} \supseteq V_\xi \quad (2.37)$$

ist wohldefiniert.

2.3.2 Wiederverwendbarkeit von Warmstartpunkten

Es stellt sich die Frage, inwiefern die Warmstartabbildungen der Beispiele 2.3.2 und 2.3.3 Parameteränderungen von ξ auf $\tilde{\xi}$ auf wiederverwendbare Werte abbilden. Das heißt z.B., wie nahe der ermittelte Warmstartpunkt an der zu bestimmenden Nullstellenmenge zu F_τ liegt. Darüber hinaus ist interessant, wie gut der neue Punkt als Iterierte des parametrisierten Optimierungsproblems $(P_{\tilde{\xi}})$ geeignet ist. Insofern sind die Aspekte Zulässigkeit, Residuen und die Nähe zur Lösung von $(P_{\tilde{\xi}})$ entscheidend. Eben diese Fragestellungen sind in diesem Abschnitt Gegenstand der Untersuchung.

Bereits per Definition ist klar, dass Warmstartpunkte nicht wieder τ -zulässig sein müssen und damit i.A. keine neuen Nullstellen der Abbildung F_τ liefern. Stattdessen werden lokal bei einer Parameterperturbation $\Delta\xi = \tilde{\xi} - \xi$ nur Residuen der Größenordnung $o(\|\Delta\xi\|)$ zugelassen.

Diesbezüglich mache man sich klar, wie die Werte der Abbildungen (2.33) und (2.35) berechnet werden.

Der Differenzenvektor $\Delta w_{\rho_w} := \Delta w_{\rho_w}(\xi, \tilde{\xi})$ aus Beispiel (2.3.2) ergibt sich für beliebige Parameter $\xi, \tilde{\xi} \in \Xi$ aus dem Gleichungssystem

$$D_w F_\tau(w(\xi); \xi) \cdot \Delta w_{\rho_w} = -D_\xi F_\tau(w(\xi); \xi) \cdot (\Delta\xi). \quad (2.38)$$

Mit der Aufteilung $\Delta w_{\rho_w} = (\Delta x_{\rho_w}, \Delta \lambda_{\rho_w}, \Delta s_{\rho_w}, \Delta \pi_{\rho_w})$ gilt äquivalent

$$\begin{pmatrix} D_{xx}\mathcal{L}(w; \xi) & D_x h(x; \xi)^\top & D_x g(x; \xi)^\top & 0 \\ D_x h(x; \xi) & 0 & 0 & 0 \\ D_x g(x; \xi) & 0 & 0 & I \\ 0 & 0 & \Pi & S \end{pmatrix} \begin{pmatrix} \Delta x_{\rho_w} \\ \Delta \lambda_{\rho_w} \\ \Delta s_{\rho_w} \\ \Delta \pi_{\rho_w} \end{pmatrix} = - \begin{pmatrix} D_\xi \mathcal{L}(w; \xi) \\ D_\xi h(x; \xi) \\ D_\xi g(x; \xi) \\ 0 \end{pmatrix} \cdot (\Delta\xi).$$

Für den Differenzenvektor $(\Delta x_{\rho_F}, \Delta \lambda_{\rho_F}, \Delta s_{\rho_F}, \Delta \pi_{\rho_F}) = \Delta w_{\rho_F} := \Delta w_{\rho_F}(\xi, \tilde{\xi})$ aus Beispiel (2.3.3) stellt sich das Gleichungssystem

$$D_w F_\tau(w(\xi); \tilde{\xi}) \cdot \Delta w_{\rho_F} = -F_\tau(w(\xi); \tilde{\xi}) \quad (2.39)$$

und damit

$$\begin{pmatrix} D_{xx}\mathcal{L}(w; \tilde{\xi}) & D_x h(x; \tilde{\xi})^\top & D_x g(x; \tilde{\xi})^\top & 0 \\ D_x h(x; \tilde{\xi}) & 0 & 0 & 0 \\ D_x g(x; \tilde{\xi}) & 0 & 0 & I \\ 0 & 0 & \Pi & S \end{pmatrix} \begin{pmatrix} \Delta x_{\rho_F} \\ \Delta \lambda_{\rho_F} \\ \Delta s_{\rho_F} \\ \Delta \pi_{\rho_F} \end{pmatrix} = - \begin{pmatrix} r(w(\xi); \tilde{\xi}) \\ 0 \end{pmatrix}.$$

Beide Gleichungssysteme besitzen ein Teilsystem der Art

$$\Pi \Delta s + S \Delta \pi = 0, \quad (2.40)$$

anhand dessen sich die folgenden Resultate fest machen lassen.

Das erste Lemma beleuchtet die Änderungen in den Komplementaritätsbedingungen und somit die Nähe zur Lösung aber gleichzeitig auch zur τ -Urbildmenge.

Lemma 2.3.6 *Für gegebene Vektoren $w = (x, \lambda, s, \pi)$, $s, \pi > 0$ und $\Delta w = (\Delta x, \Delta \lambda, \Delta s, \Delta \pi)$ gelte Bedingung (2.40). Es seien $\tau := \Pi S e$ und $\tilde{w} = w + \Delta w$ sowie $\tilde{w} = (\tilde{x}, \tilde{\lambda}, \tilde{s}, \tilde{\pi})$, dann gilt*

$$\tilde{\tau} := \tilde{\Pi} \tilde{S} e = \tau + \Delta \tau$$

mit $\Delta \tau := \Delta \Pi \Delta S e$ und letztlich

$$\tau \geq \tilde{\tau}.$$

Beweis: Es gilt

$$\begin{aligned}
\tilde{\Pi}\tilde{S}e &= (\Pi + \Delta\Pi)(S + \Delta S)e \\
&= \Pi Se + \underbrace{\Pi\Delta Se + \Delta\Pi Se}_{=0} + \Delta\Pi\Delta Se \\
&= \tau + \underbrace{\Delta\Pi\Delta Se}_{=\Delta\tau} \\
&= \tau + \Delta\tau = \tilde{\tau}.
\end{aligned}$$

Ferner zeigt sich

$$\Delta\Pi\Delta Se = -\Pi^{-1}S\Delta\Pi^2e \leq 0, \quad (2.41)$$

da Π, S und $\Delta\Pi^2$ nur positive Einträge besitzen. Das bedeutet letztlich $\tau \geq \tilde{\tau}$.

□

Damit ist klar, dass die resultierenden Warmstartpunkte $w + \Delta w$ i.A. nicht mehr in derselben τ -Urbildmenge liegen. Tendenziell wird die Perturbation τ der Komplementaritätsbedingungen verkleinert. Der neu gewonnene Punkt liegt hinsichtlich der Komplementaritätsbedingungen „näher“ an der optimalen Lösung des parametrisierten Optimierungsproblems ($P_{\tilde{\xi}}$) als der ursprüngliche Punkt w der Lösung von (P_{ξ}). Andererseits, gemessen an der Komplementaritätsbedingung, lässt sich der Abstand von $w + \Delta w$ zur τ -Urbildmenge mit Hilfe von $\Delta\tau$ beschreiben.

Als nächstes geht es um die Zulässigkeit der Warmstartpunkte. Per Definition ist ein Warmstartpunkt genau dann zulässig, wenn $s \geq 0$ und $\pi \geq 0$ erfüllt ist (s. Definition 2.3.1).

Lemma 2.3.7 *Für die Vektoren $w = (x, \lambda, s, \pi)$ und $\Delta w = (\Delta x, \Delta\lambda, \Delta s, \Delta\pi)$ gelte Bedingung (2.40). Mit $\pi > 0$ und $s > 0$ und einem festen $\varepsilon \in (0, 1)$ folgt dann*

(a)

$$\|S^{-1}\Delta s\|_{\infty} = \|\Pi^{-1}\Delta\pi\|_{\infty}, \quad (2.42)$$

(b)

$$\|S^{-1}\Delta s\|_{\infty} \leq 1 - \varepsilon \quad \Leftrightarrow \quad \begin{cases} s + \Delta s \geq \varepsilon s > 0 \\ \pi + \Delta\pi \geq \varepsilon\pi > 0 \end{cases}. \quad (2.43)$$

Beweis:

Zu a): Aus Voraussetzung (2.40) folgt

$$\frac{\Delta\pi_i}{\pi_i} = -\frac{\Delta s_i}{s_i} \quad \forall i \in \{1, \dots, n\} \quad (2.44)$$

und damit sofort $\Pi^{-1}\Delta\pi = -S^{-1}\Delta s$.

Zu b): Es gelte $\|S^{-1}\Delta s\|_\infty \leq 1 - \varepsilon$. Daraus ergibt sich

$$(-1 + \varepsilon)s_i \leq \Delta s_i \leq (1 - \varepsilon)s_i, \quad \forall i \in \{1, \dots, n\}. \quad (2.45)$$

Die Summe $s_i + \Delta s_i$ ist nur dann kleiner oder gleich Null, wenn $\Delta s_i < 0$. Das führte jedoch zu

$$s_i + \Delta s_i \geq s_i + (-1 + \varepsilon)s_i = \varepsilon s_i > 0. \quad (2.46)$$

Da aus (a) die Gleichheit der beiden Normen gegeben ist, kann die Einhaltung der Positivität für π analog gezeigt werden.

Bleibt noch die Rückrichtung der Äquivalenz zu zeigen. Seien dazu $\pi + \Delta\pi > \varepsilon\pi$ und $s + \Delta s > \varepsilon s$ vorausgesetzt. Diese Bedingungen sind äquivalent zu

$$\frac{\Delta\pi_i}{\pi_i} \geq -1 + \varepsilon \quad (2.47)$$

$$\frac{\Delta s_i}{s_i} \geq -1 + \varepsilon \quad (2.48)$$

für alle $i = 1, \dots, n$ und ein geeignetes $\varepsilon > 0$. Ferner gilt

$$\frac{\Delta s_i}{s_i} \leq 1 - \varepsilon, \quad (2.49)$$

wegen $-1 + \varepsilon \leq \frac{\Delta\pi_i}{\pi_i} = -\frac{\Delta s_i}{s_i}$, folgend aus Gleichung (2.44).

Insgesamt folgt daraus die Behauptung $\|S^{-1}\Delta s\|_\infty \leq 1 - \varepsilon$.

□

Als Resultat halte man fest:

Korollar 2.3.8 *Gegeben seien die Mengen Ξ und $U \subset \Xi$. Für ein parametrisiertes Optimierungsproblem (P_ξ) sei der bezüglich $\xi \in \Xi$ τ -zulässige Punkt $w = (x, \lambda, s, \pi) \in \mathbb{R}^{\tilde{n}}$ gegeben. Ferner existiere eine Warmstartfunktion $\rho : U \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}$ sowie resultierend eine entsprechende Differenzenabbildung $\Delta w = (\Delta x, \Delta \lambda, \Delta s, \Delta \pi)$.*

Erfüllen w und $\Delta w(\xi, \tilde{\xi})$ für alle $\xi \in U, \tilde{\xi} \in \Xi$ die Bedingung (2.40), so ergeben sich für alle Parameter der Menge

$$Y := \{v \in \Xi \mid \|S^{-1}\Delta s(\xi, v)\|_\infty < 1\} \quad (2.50)$$

zulässige Warmstartpunkte.

Beweis:

Ergibt sich aus Lemma 2.3.7

□

Für eine Nullstellen-Abbildung $w \in \mathcal{C}^1(U, \mathbb{R}^{\tilde{n}})$ von F_τ existiert nach Korollar 2.3.4 die Differenzenabbildung

$$\Delta w_{\rho_w} : U \times \mathbb{R}^k \rightarrow \mathbb{R}^{\tilde{n}}, \quad (\xi, \tilde{\xi}) \mapsto D_\xi w(\xi) \cdot (\tilde{\xi} - \xi)$$

insbesondere an einer beliebigen Stelle $\xi \in U$, wenn $D_w F_\tau(w(\xi), \xi)$ regulär ist. Ein Differenzenvektor $\Delta w_{\rho_w}(\xi, \tilde{\xi})$, $\tilde{\xi} \in \mathbb{R}^k$ und damit $\Delta s_{\rho_w}(\xi, \tilde{\xi})$ ergibt sich somit als Linearkombination der Spalten der Funktionalmatrix $D_\xi w(\xi)$, deren Koeffizienten der Vektor $\tilde{\xi} - \xi$ liefert. So kann demnach sehr leicht die Menge Y der zulässigen Parameter mittels Ermittlung geeigneter Linearkombinationen bestimmt werden.

Schwieriger gestaltet sich die Bestimmung von Y für die Differenzenabbildung (2.35)

$$\Delta w_{\rho_F} : U \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}, \quad (\xi, \tilde{\xi}) \mapsto -D_w F_\tau(w(\xi); \tilde{\xi})^{-1} \cdot F_\tau(w(\xi); \tilde{\xi}),$$

wobei die Menge Ξ so zu wählen war, dass sämtliche Funktionalmatrizen $D_w F_\tau(w(\xi); \tilde{\xi})$, $\tilde{\xi} \in \Xi$ regulär sind.

Eine Berechnung des Differenzenvektors $\Delta w_{\rho_F}(\xi, \tilde{\xi})$ erfordert die Lösung eines Gleichungssystems mit der Funktionalmatrix $D_w F_\tau(w(\xi); \tilde{\xi})$, die sich für jedes $\tilde{\xi} \in \Xi$ ändert und somit neu berechnet werden müsste. D.h. eine Bestimmung der Menge Y durch geschicktes Probieren mit mehreren Parametern, z.B. nach dem Prinzip eines Bisektionsverfahrens, ist sehr aufwendig.

Bleibt die Klärung des letzten Aspekts, das Verhalten der Residuen. Die Veränderung in den Komplementaritätsbedingungen konnte bereits in Lemma 2.3.6 beschrieben werden. Das Augenmerk liegt jetzt auf der Residuenabbildung r aus (2.10)

Lemma 2.3.9 *Zu einem festen Perturbationsparameter $\tau \in \mathbb{R}_{>0}^{m_g}$ gebe es die Parametermengen U und Ξ , so dass für die Abbildung $F_\tau \in \mathcal{C}^1(\mathbb{R}^{m_g} \times \Xi, \mathbb{R}^{\tilde{n}})$ die Warmstartabbildung $\rho_F : U \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}$ existiert. Aus F_τ resultiere für beliebige aber feste $\xi \in \Xi$ eine affin lineare Residuenabbildung $r(\cdot, \xi) : \mathbb{R}^{\tilde{n}} \rightarrow \mathbb{R}^{\tilde{n}-n}$ wie unter (2.10).*

Dann ist ρ_F eine eigentliche Warmstartabbildung.

Beweis: Aufgrund der existierenden Warmstartabbildung ρ_F gibt es per Definition eine vollständig gegebene τ -Urbildmenge $\mathcal{N}_F(U, \tau)$ sowie eine eindeutig bestimmte Nullstellen-Abbildung $w \in \mathcal{C}^1(U, \mathbb{R}^{\tilde{n}})$.

Für eine Nullstelle $(w(\xi), \xi) \in \mathcal{N}_F(U, \tau)$, $w(\xi) = (x(\xi), \lambda(\xi), s(\xi), \pi(\xi))$ ist dann auch $r(w(\xi); \xi) = 0$.

Die Warmstartabbildung ρ_F ist genau dann eigentlich, wenn zu einem Parameterwechsel von $\xi \in U$ auf $\tilde{\xi} \in \Xi$ mit dem zugehörigen Differenzenvektor $\Delta w_{\rho_F}(\xi, \tilde{\xi}) = \rho_F(\xi, \tilde{\xi}) - w(\xi)$ die Bedingung

$$r(w(\xi) + \Delta w(\xi, \tilde{\xi}); \tilde{\xi}) = 0$$

erfüllt ist.

Zusätzlich muss Δw_{ρ_F} , wegen (2.35) und da

$$F_\tau(w(\xi); \tilde{\xi}) = \begin{pmatrix} r(w(\xi); \tilde{\xi}) \\ \text{diag}(s(\xi)) \cdot \text{diag}(\pi(\xi)) \cdot e - \tau \end{pmatrix}$$

gilt, insbesondere dem Gleichungssystem

$$D_w r(w(\xi); \tilde{\xi}) \cdot \Delta w_{\rho_F}(\xi, \tilde{\xi}) = -r(w(\xi); \tilde{\xi}) \quad (2.51)$$

genügen.

Nach Voraussetzung ist $r(\cdot, \xi)$ für alle $\xi \in \Xi$ affin linear. Damit kann $r(\cdot, \xi)$ mit Hilfe der Matrix $B(\xi) \equiv B(w, \xi) = D_w r(w; \xi)$, $w \in \mathbb{R}^{\tilde{n}}$ und einem Vektor $d(\xi)$ beschrieben werden, nämlich

$$r(w; \xi) = B(\xi) w + d(\xi). \quad (2.52)$$

Entsprechend folgt

$$r(w(\xi) + \Delta w_{\rho_F}(\xi, \tilde{\xi}); \tilde{\xi}) \stackrel{(2.52)}{=} B(\tilde{\xi}) w(\xi) + B(\tilde{\xi}) \Delta w_{\rho_F}(\xi, \tilde{\xi}) + d(\tilde{\xi}) \quad (2.53)$$

$$\stackrel{(2.52)}{=} r(w(\xi); \tilde{\xi}) + B(\tilde{\xi}) \Delta w_{\rho_F}(\xi, \tilde{\xi}) \quad (2.54)$$

$$\stackrel{(2.51)}{=} 0. \quad (2.55)$$

□

Bemerkung 2.3.10 Für lineare und quadratische Optimierungsprobleme bildet die Residuenfunktion $r(\cdot; \xi)$ eine affin lineare Abbildung für alle $\xi \in \Xi$. In diesen Fällen ist ρ_F eine eigentliche Warmstartabbildung (s. Beispiel 2.4.1).

Korollar 2.3.11 *Es sei $\tau \in \mathbb{R}_{>0}^{m_g}$ gegeben. Dazu existiere eine eigentliche Warmstartabbildung $\rho : U \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}$ und für eine Parameteränderung von $\xi \in U$ auf $\tilde{\xi} \in \Xi$ der resultierende Differenzvektor $\Delta w = (\Delta x, \Delta \lambda, \Delta s, \Delta \pi)$. Dann existiert auch eine Abbildung $w_\tau \in \mathcal{C}^1(U, \mathbb{R}^{\tilde{n}})$, die τ -zulässige Werte liefert.*

Bedingung (2.40) sei für $w_\tau(\xi)$ und Δw erfüllt. Sofern die Parameteränderung zu einem zulässigen Warmstartpunkt

$$v := w_\tau(\xi) + \Delta w$$

führt, existiert ein $\tilde{\tau} \in \mathbb{R}_{>0}^{m_g}$, $\tilde{\tau} \leq \tau$, so dass $F_{\tilde{\tau}}(v; \tilde{\xi}) = 0$ gilt.

Ist darüber hinaus $D_w F_{\tilde{\tau}}(v; \tilde{\xi})$ regulär, existiert eine Umgebung $U_{\tilde{\xi}}$ von $\tilde{\xi}$ und eine eindeutig bestimmte Abbildung $w_{\tilde{\tau}} \in \mathcal{C}^1(U_{\tilde{\xi}}, \mathbb{R}^{\tilde{n}})$, die für alle $\phi \in U_{\tilde{\xi}}$ τ -zulässige Werte liefert.

Beweis:

Aufgrund der gegebenen Warmstartabbildung ρ gibt es per Definition 2.3.1 lokal eine Nullstellen-Abbildung w_τ . Der Perturbationsparameter $\tilde{\tau} \leq \tau$ folgt aus Lemma 2.3.6. Da v wieder zulässig ist, muss $\tilde{\tau} > 0$ gelten. Nach Voraussetzung ist ρ eigentlich und damit $r(v; \tilde{\xi}) = 0$, wodurch $F_{\tilde{\tau}}(v; \tilde{\tau}) = 0$ gelten muss.

□

2.4 Sensitivität τ -perturbierter Systeme

Hinsichtlich der Schwierigkeiten einen zulässigen Parameterwechsel für die Warmstartabbildung ρ_F aus Beispiel 2.3.3 zu finden, wird in diesem Abschnitt basierend auf Lemma 2.3.7 eine a-priori-Abschätzung für zulässige Wechsel durchgeführt.

Dazu betrachte man eine spezielle Variante des parametrisierten Optimierungsproblems (P_ξ) mit $\xi \in \Xi \subset \mathbb{R}$,

$$\begin{aligned} \min \quad & f(x) + \xi \Delta f(x) \\ \text{NB.} \quad & h(x) + \xi \Delta h(x) = 0 \\ & g(x) + \xi \Delta g(x) + \pi = 0 \\ & \pi \geq 0 \end{aligned} \tag{2.56}$$

und entsprechende Abbildungen $h, \Delta h \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^{m_h})$ und $g, \Delta g \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^{m_g})$ sowie $f, \Delta f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$.

So besitzt auch die Lagrangefunktion die Form

$$\mathcal{L}(w; \xi) = \mathcal{L}(w) + \xi \Delta \mathcal{L}(w) \quad (2.57)$$

mit $\Delta \mathcal{L}(w) = \Delta f(x) + \lambda^\top \Delta h(x) + s^\top \Delta g(x)$. Letztendlich unterliegt F_τ aus (2.25) ebenfalls der Struktur

$$F_\tau(w; \xi) = F_\tau(w; 0) + \xi \Delta F(w), \quad (2.58)$$

wobei

$$\Delta F(w) := \begin{pmatrix} D_x \Delta \mathcal{L}(w)^\top \\ \Delta h(x) \\ \Delta g(x) \\ 0 \end{pmatrix} \quad (2.59)$$

gilt.

Dies wirkt sich weiter auf die Funktionalmatrizen $D_w F_\tau(w; \xi)$ aus. Definiert man

$$M(w; \xi) := D_w F_\tau(w; \xi), \quad (2.60)$$

$$\Delta M(w) := D_w \Delta F(w) \quad (2.61)$$

$$= \begin{pmatrix} D_{xx} \Delta \mathcal{L}(w) & D_x \Delta h(x)^\top & D_x \Delta g(x)^\top & 0 \\ D_x \Delta h(x) & 0 & 0 & 0 \\ D_x \Delta g(x) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

so folgt

$$D_w F_\tau(w; \xi) = M(w; \xi) = M(w; 0) + \xi \Delta M(w). \quad (2.62)$$

Beispiel 2.4.1 Das Verständnis für die neu definierten Größen soll nun anhand eines Beispiels vertieft werden. Dazu betrachte man das konvex quadratische Optimierungsproblem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \frac{1}{2} x^\top (Q + \xi \Delta Q) x + (c + \xi \Delta c)^\top x \\ \text{N.B.} \quad & (A + \xi \Delta A) x = b + \xi \Delta b \\ & x \geq 0, \end{aligned} \quad (QP_\xi)$$

wobei $\xi \in [0, 1]$. Für alle $\xi \in [0, 1]$ seien die Matrizen $Q + \xi \Delta Q \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit; die Matrizen $A + \xi \Delta A \in \mathbb{R}^{m \times n}$ haben vollen Zeilenrang. Die Schlupfvariable π entspricht hier aufgrund der Ungleichungsrestriktion $g(x) = x$ der Variablen x und wird deshalb nicht weiter mit aufgeführt.

Die zweite Ableitung der Lagrangefunktion $D_{xx}\mathcal{L}(w; \xi)$ besteht lediglich aus der Matrix $Q + \xi\Delta Q$. Damit sind die Voraussetzungen des Lemmas 2.2.8 erfüllt, d.h. um jeden Parameter ξ mit τ -zulässigen Punkt $w = (x, \lambda, s)$ existiert eine Umgebung $U(\xi)$ auf der eine eindeutig bestimmte, stetig-differenzierbare Abbildung w die Nullstellen von

$$F_\tau(w; \xi) = \begin{pmatrix} -(Q + \xi\Delta Q)x - (c + \xi\Delta c) + (A + \xi\Delta A)^\top \lambda + s \\ (A + \xi\Delta A)x - (b + \xi\Delta b) \\ SXe - \tau \end{pmatrix}$$

liefert.

Entsprechend ergeben sich die Größen und Funktionen:

$$\Delta F(w) = \begin{pmatrix} -\Delta Qx - \Delta c + \Delta A^\top \lambda \\ \Delta Ax - \Delta b \\ 0 \end{pmatrix} \quad (2.63)$$

$$M(w; \xi) = D_w F_\tau(w; \xi) = \begin{pmatrix} -(Q + \xi\Delta Q) & (A + \xi\Delta A)^\top & I \\ A + \xi\Delta A & 0 & 0 \\ S & 0 & X \end{pmatrix} \quad (2.64)$$

$$\Delta M \equiv \Delta M(w) = D_w \Delta F(w) = \begin{pmatrix} -\Delta Q & \Delta A^\top & 0 \\ \Delta A & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.65)$$

Die Warmstartabbildung ρ_w im Punkt w zum Parameter $\tilde{\xi} = 0$ bezüglich ξ leitet sich demnach aus dem Gleichungssystem

$$\begin{pmatrix} -Q & A^\top & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x_{\rho_w} \\ \Delta \lambda_{\rho_w} \\ \Delta s_{\rho_w} \end{pmatrix} = \xi \cdot \begin{pmatrix} \Delta Qx + \Delta c - \Delta A^\top \lambda \\ \Delta b - \Delta Ax \\ 0 \end{pmatrix} \quad (2.66)$$

her. Ebenso ergibt sich die Warmstartabbildung ρ_F aus dem Gleichungssystem

$$\begin{pmatrix} -(Q + \xi\Delta Q) & (A + \xi\Delta A)^\top & I \\ A + \xi\Delta A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x_{\rho_F} \\ \Delta \lambda_{\rho_F} \\ \Delta s_{\rho_F} \end{pmatrix} = \xi \cdot \begin{pmatrix} \Delta Qx + \Delta c - \Delta A^\top \lambda \\ \Delta b - \Delta Ax \\ 0 \end{pmatrix}. \quad (2.67)$$

Die Warmstartabbildungen ρ_F und ρ_w sind wohldefiniert, da es sich nach Lemma 2.2.8 bei den benötigten Matrizen $M(w; \xi)$ und $M(w; 0)$ stets um reguläre handelt.

Gegenüber dem allgemeinen nichtlinearen Fall treten im quadratischen Fall günstige Bedingungen auf, die zu weiteren Aussagen führen. Wie bereits erwähnt, liefert ρ_F

für lineare und quadratische Probleme eigentliche Warmstartpunkte, sofern die Parameter zulässig sind. Denn setzt man den Warmstartpunkt $w_{\rho_F} = w + \Delta w_{\rho_F}$ in F_τ ein, wobei $\Delta w_{\rho_F} = (x_{\rho_F}, \lambda_{\rho_F}, s_{\rho_F})$, so gilt

$$F_\tau(w_{\rho_F}; \xi) = F_\tau(w_{\rho_F}; 0) + \xi \Delta F(w_{\rho_F}) \quad (2.68)$$

$$= \begin{pmatrix} -Q(x + \Delta x_{\rho_F}) - c + A^\top(\lambda + \Delta \lambda_{\rho_F}) + s + \Delta s_{\rho_F} \\ A(x + \Delta x_{\rho_F}) - b \\ (S + \Delta S_{\rho_F})(X + \Delta X_{\rho_F})e - \tau \end{pmatrix} \quad (2.69)$$

$$+ \xi \cdot \begin{pmatrix} -\Delta Q(x + \Delta x_{\rho_F}) - \Delta c + \Delta A^\top(\lambda + \Delta \lambda_{\rho_F}) \\ \Delta A(x + \Delta x_{\rho_F}) - \Delta b \\ 0 \end{pmatrix}$$

$$\stackrel{F_\tau(w; 0)=0}{=} \begin{pmatrix} -Q\Delta x_{\rho_F} + A^\top \Delta \lambda_{\rho_F} + \Delta s_{\rho_F} \\ A\Delta x_{\rho_F} \\ \underbrace{S\Delta X_{\rho_F}e + X\Delta S_{\rho_F}e + \Delta S_{\rho_F}\Delta X_{\rho_F}e}_{=0} \end{pmatrix} \quad (2.70)$$

$$+ \xi \cdot \begin{pmatrix} -\Delta Q(x + \Delta x_{\rho_F}) - \Delta c + \Delta A^\top(\lambda + \Delta \lambda_{\rho_F}) \\ \Delta A(x + \Delta x_{\rho_F}) - \Delta b \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} -(Q + \xi \Delta Q)\Delta x_{\rho_F} + (A + \xi \Delta A)^\top \Delta \lambda_{\rho_F} + \Delta s_{\rho_F} \\ (A + \xi \Delta A)\Delta x_{\rho_F} \\ \Delta S_{\rho_F} \Delta X_{\rho_F} e \end{pmatrix} \quad (2.71)$$

$$+ \xi \cdot \begin{pmatrix} -\Delta Qx - \Delta c + \Delta A^\top \lambda \\ \Delta Ax - \Delta b \\ 0 \end{pmatrix}$$

$$\stackrel{(2.67)}{=} \begin{pmatrix} 0 \\ 0 \\ \Delta S_{\rho_F} \Delta X_{\rho_F} e \end{pmatrix}. \quad (2.72)$$

Genauso kann das Residuum des Warmstartpunktes $\rho_w(w; \xi) =: w_{\rho_w}$ mit dem Differenzenvektor Δw quantifiziert werden.

$$F_\tau(w_{\rho_w}; 1) \stackrel{F_\tau(w; 0)=0}{=} \begin{pmatrix} -Q\Delta x_{\rho_w} + A^\top \Delta \lambda_{\rho_w} + \Delta s_{\rho_w} \\ A\Delta x_{\rho_w} \\ \Delta S_{\rho_w} \Delta X_{\rho_w} e \end{pmatrix} \quad (2.73)$$

$$+ \begin{pmatrix} -\Delta Q(x + \Delta x_{\rho_w}) - \Delta c + \Delta A^\top(\lambda + \Delta \lambda_{\rho_w}) \\ \Delta A(x + \Delta x_{\rho_w}) - \Delta b \\ 0 \end{pmatrix}$$

$$\stackrel{(2.66)}{=} \begin{pmatrix} -\Delta Q\Delta x_{\rho_w} + \Delta A^\top \Delta \lambda_{\rho_w} \\ \Delta A\Delta x_{\rho_w} \\ \Delta S_{\rho_w} \Delta X_{\rho_w} e \end{pmatrix}. \quad (2.74)$$

Da sich der Differenzenvektor von ρ_w proportional zu ξ verhält, entsteht das Resi-

duum zum Parameter ξ

$$F(w_{\rho_w}; \xi) = \xi^2 \cdot \begin{pmatrix} -\Delta Q \Delta x_{\rho_w} + \Delta A^\top \Delta \lambda \\ \Delta A \Delta x_{\rho_w} \\ \Delta S_{\rho_w} \Delta X_{\rho_w} e \end{pmatrix}. \quad (2.75)$$

Um den Einfluss von Parameteränderungen zu untersuchen, gehe man davon aus, dass die Nullstelle $(w_0; 0)$ mit $F_\tau(w_0; 0) = 0$, $w_0 = (x_0, \lambda_0, s_0, \pi_0)$ bekannt und deren Funktionalmatrix $D_w F_\tau(w_0; 0) = M(w_0; 0)$ regulär ist. Wegen Lemma 2.2.8(b) existiert eine Umgebung $U \subset \mathbb{R}$ und eine eindeutig bestimmte, stetig-differenzierbare Abbildung $w : U \rightarrow \mathbb{R}^{\tilde{n}}$ mit τ -zulässigen Werten.

Schon wegen Korollar 2.3.4 ist

$$\rho_w : U \times \mathbb{R} \rightarrow \mathbb{R}^{\tilde{n}}$$

als Warmstartabbildung wohldefiniert. Zusätzlich gebe es eine nichtleere Menge

$$\Xi = \{\xi \in \mathbb{R} \mid M(w(\phi); \xi) \text{ regulär } \forall \phi \in U\}$$

mit der auch

$$\rho_F : U \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}$$

Warmstartabbildung sei.

Entsprechend können auch die Differenzenabbildungen

$$\Delta w_{\rho_w} : U \times \mathbb{R} \rightarrow \mathbb{R}^{\tilde{n}}, \quad (\xi, \tilde{\xi}) \mapsto -(\tilde{\xi} - \xi) \cdot M(w(\xi); \xi)^{-1} D_\xi F_\tau(w(\xi); \xi) \quad (2.76)$$

mit den Unterteilungen

$$\Delta w_{\rho_w} = (\Delta x_{\rho_w}, \Delta \lambda_{\rho_w}, \Delta s_{\rho_w}, \Delta \pi_{\rho_w}) \quad (2.77)$$

und

$$\Delta w_{\rho_F} : U \times \Xi \rightarrow \mathbb{R}^{\tilde{n}}, \quad (\xi, \tilde{\xi}) \mapsto -M(w(\xi); \tilde{\xi})^{-1} F_\tau(w(\xi); \tilde{\xi}) \quad (2.78)$$

mit der Unterteilung

$$\Delta w_{\rho_F} = (\Delta x_{\rho_F}, \Delta \lambda_{\rho_F}, \Delta s_{\rho_F}, \Delta \pi_{\rho_F}). \quad (2.79)$$

eingeführt werden.

Für den zulässigen Parameterbereich hinsichtlich ρ_w ergibt sich wegen Lemma 2.3.7 mit

$$\tilde{\xi} := \frac{1}{\|S^{-1}\Delta s_{\rho_w}(0, 1)\|_\infty}. \quad (2.80)$$

das Intervall $Y := [0, \tilde{\xi})$, so dass ρ_w für alle $\xi \in Y$ zulässige Warmstartpunkte liefert.

Die spezielle Wahl des parametrisierten Optimierungsproblems (2.56) ermöglicht die Herleitung eines Zusammenhangs zwischen den Differenzenvektoren von ρ_w und ρ_F .

Lemma 2.4.2 *Man betrachte das parametrisierte Optimierungsproblem (2.56) zu einem festen Parameter $\tau \in \mathbb{R}_{\geq 0}^{mg}$. Der Punkt $(w_0; 0)$ sei τ -zulässig. Zu einer Menge $V \subset \mathbb{R}$, $0 \in V$ seien die Funktionalmatrizen $M(w_0; \xi)$ für alle $\xi \in V$ regulär.*

Gegeben seien die beiden Differenzenabbildungen aus Δw_{ρ_w} aus (2.77) und Δw_{ρ_F} aus (2.78). Dann folgt

$$\Delta w_{\rho_F}(0, \xi) = (I + \xi M(w_0; 0)^{-1} \Delta M(w_0)) \Delta w_{\rho_w}(0, \xi) \quad (2.81)$$

für alle $\xi \in V$.

Beweis:

Wegen (2.58) folgt zum einen für beliebige $w \in \mathbb{R}^{\tilde{n}}$ und $\xi \in \mathbb{R}$,

$$D_\xi F_\tau(w; \xi) = \Delta F(w), \quad (2.82)$$

zum anderen für die Nullstelle $(w_0; 0)$ und beliebige $\xi \in \mathbb{R}$

$$\begin{aligned} F_\tau(w_0; \xi) &= \underbrace{F_\tau(w_0; 0)}_{=0} + \xi \Delta F(w_0) \\ &= \xi \cdot \Delta F(w_0) \end{aligned} \quad (2.83)$$

Per Definition (2.76) folgt

$$\begin{aligned} \Delta w_{\rho_w}(0; \xi) &= -\xi \cdot M(w_0; 0)^{-1} D_\xi F_\tau(w_0; 0) \\ &\stackrel{(2.82)}{=} -\xi \cdot M(w_0; 0)^{-1} \Delta F(w_0) \end{aligned} \quad (2.84)$$

Es bleibt Behauptung (2.81) zu zeigen. Dafür ist

$$\begin{aligned} \Delta w_{\rho_F}(0, \xi) &= -D_w F_\tau(w_0; \xi)^{-1} \cdot F_\tau(w_0; \xi) \\ &\stackrel{(2.62), (2.83)}{=} \xi \cdot (M(w_0; 0) + \xi \Delta M(w_0))^{-1} \cdot \Delta F(w_0) \\ &= \xi \cdot (I + \xi M(w_0; 0)^{-1} \Delta M(w_0))^{-1} \cdot M(w_0; 0)^{-1} \cdot \Delta F(w_0) \\ &\stackrel{(2.84)}{=} (I + \xi M(w_0)^{-1} \Delta M(w_0))^{-1} \cdot \Delta w_{\rho_w}(0, \xi). \end{aligned}$$

□

Sobald $M(w_0; 0)^{-1}\Delta M(w_0)$ der Nullmatrix entspricht, sind die entsprechenden Differenzvektoren Δw_{ρ_w} und Δw_{ρ_F} gleich. Die Bestimmung einer zulässigen Parametermenge für ρ_F ist dann trivial.

Der folgende Abschnitt beinhaltet einen möglichen Ansatz, wenn $M(w_0; 0)^{-1}\Delta M(w_0)$ eben nicht der Nullmatrix entspricht.

2.4.1 Abschätzung zulässiger Parametermengen

Mit Hilfe der bisher gewonnenen Resultate lässt sich eine a-priori Abschätzung für eine Menge $Y \subset \Xi$ konstruieren, die nur Parameter enthält, die zu zulässigen Warmstartpunkten bezüglich der Warmstartabbildung ρ_F führen. Insbesondere Lemma 2.3.7 und 2.4.2 spielen hierbei eine große Rolle. So werden Informationen zulässiger Parameter der Warmstartabbildung ρ_w benutzt, um entsprechende Parameter für ρ_F zu finden.

Für die Abschätzung der Parametermenge Y in Lemma 2.4.4 wird folgendes Lemma benötigt:

Lemma 2.4.3 (Neumannsche Reihe) *Sei $H \in \mathbb{R}^{n \times n}$ und $\|H\| < 1$, dann ist $I - H$ regulär und es gilt*

$$(I - H)^{-1} = \sum_{k=1}^{\infty} H^k \quad (2.85)$$

mit der Abschätzung

$$\|(I - H)^{-1}\| \leq \frac{1}{1 - \|H\|}. \quad (2.86)$$

Beweis: S. [GolVLo89], Lemma 2.3.3.

Satz 2.4.4 (a-priori Abschätzung) *Wie in Lemma 2.4.2 betrachte man das parametrisierte Optimierungsproblem (2.56) zu einem festen Parameter $\tau \in \mathbb{R}_{\geq 0}^{m_g}$. Der Punkt $(w_0; 0)$, $w_0 = (x_0, \lambda_0, s_0, \pi_0)$ sei wieder τ -zulässig. Zu einer nichtleeren Menge $V \subset \mathbb{R}$, $0 \in V$ seien die Funktionalmatrizen $M(w_0; \xi)$ für alle $\xi \in V$ regulär, diesmal mit $M(w_0; 0)^{-1}\Delta M(w_0) \neq 0$.*

Entsprechend seien die beiden Differenzenabbildungen aus Δw_{ρ_w} aus (2.77) und Δw_{ρ_F} aus (2.78) wohldefiniert. Für ein $\varepsilon \in]0, 1[$ definiere man den Parameter

$$\xi_{\rho_w} := \frac{1 - \varepsilon}{\|S_0^{-1} \Delta s_{\rho_w}(0, 1)\|_\infty} \quad (2.87)$$

und $\kappa := \varepsilon \cdot \min_{i=\{1, \dots, n\}} \{(s_0)_i, (\pi_0)_i\}$.

Dann kann mit $H := M(w_0; 0)^{-1} \Delta M(w_0)$ sowie dem Parameter

$$\xi_{\rho_F} := \min \left\{ 1, \xi_{\rho_w}, \frac{1}{\|H\|_\infty}, \frac{\kappa}{\|\Delta w_{\rho_w}(0, 1)\|_\infty + \kappa \|H\|_\infty} \right\}. \quad (2.88)$$

die Menge

$$Y := [0, \xi_{\rho_F}[\quad (2.89)$$

definiert werden, so dass mit der Differenzenabbildung $\Delta s_{\rho_F}(0, \cdot) : V \rightarrow \mathbb{R}^{m_g}$ und $\Delta \pi_{\rho_F}(0, \cdot) : V \rightarrow \mathbb{R}^{m_g}$ aus (2.79)

$$s_0 + \Delta s_{\rho_F}(0, \xi) > 0 \quad (2.90)$$

$$\pi_0 + \Delta \pi_{\rho_F}(0, \xi) > 0 \quad (2.91)$$

für alle $\xi \in Y \cap V$ gilt.

Beweis: Man verwende $\|\cdot\| := \|\cdot\|_\infty$. Die Existenz und Eindeutigkeit der Differenzenabbildungen Δw_{ρ_F} und Δw_{ρ_w} , deren Unterabbildungen Δs_{ρ_w} , Δs_{ρ_F} , $\Delta \pi_{\rho_w}$, $\Delta \pi_{\rho_F}$ usw. ist wegen der Invertierbarkeit der Funktionalmatrizen $M(w; \xi)$ gegeben. Angesichts des Lemmas 2.4.2 besteht dann der Zusammenhang

$$\Delta w_{\rho_F}(0, \xi) = \xi (I + \xi H)^{-1} \Delta w_{\rho_w}(0, 1). \quad (2.92)$$

Infolge der Bedingung (2.87) und Lemma 2.3.7 gelten für alle $\xi \in [0, \xi_{\rho_w}]$ die Abschätzungen

$$s_0 + \Delta s_{\rho_w}(0, \xi) \geq \varepsilon s_0,$$

$$\pi_0 + \Delta \pi_{\rho_w}(0, \xi) \geq \varepsilon \pi_0.$$

D.h. es existiert eine kleinste Differenz $\kappa = \varepsilon \cdot \min_{i=\{1, \dots, n\}} \{(s_0)_i, (\pi_0)_i\}$. Die Bedingungen (2.90) und (2.91) sind somit auf jeden Fall erfüllt, sofern für einen Parameter $\xi \in [0, \xi_{\rho_w}[$ der Abstand $\|\Delta w_{\rho_F}(0, \xi) - \Delta w_{\rho_w}(0, \xi)\|$ echt kleiner κ ist. Denn

$$s_0 + \Delta s_{\rho_F}(0, \xi) = s_0 + \Delta s_{\rho_w}(0, \xi) - \Delta s_{\rho_w}(0, \xi) + \Delta s_{\rho_F}(0, \xi) \geq 0,$$

da $\|s_0 + \Delta s_{\rho_w}(0, \xi)\|_\infty \geq \kappa$ und $\|\Delta s_{\rho_F}(0, \xi) - \Delta s_{\rho_w}(0, \xi)\|_\infty \leq \kappa$ gelten. Für π_0 folgt das gleiche.

Sei also $\xi \in [0, \xi_{\rho_F}[$ beliebig gewählt. Nach Voraussetzung ist $\xi < \frac{1}{\|H\|}$ und somit $\xi\|H\| < 1$, weshalb Lemma 2.4.3 anwendbar ist. Der Abstand zwischen den Werten $\Delta w_{\rho_F}(0, \xi)$ und $\Delta w_{\rho_w}(0, \xi) = \xi \Delta w_{\rho_w}(0, 1)$ verhält sich wie folgt:

$$\begin{aligned}
\|\Delta w_{\rho_F}(0, \xi) - \Delta w_{\rho_w}(0, \xi)\| &\stackrel{(2.92)}{=} \|(I + \xi H)^{-1} \xi \Delta w_{\rho_w}(0, 1) - \xi \Delta w_{\rho_w}(0, 1)\| \\
&= \xi \|(I + \xi H)^{-1} - I\| \|\Delta w_{\rho_w}(0, 1)\| \\
&= \xi \|(I + \xi H)^{-1} (I - I - \xi H)\| \|\Delta w_{\rho_w}(0, 1)\| \\
&= \xi^2 \|(I + \xi H)^{-1} H\| \|\Delta w_{\rho_w}(0, 1)\| \\
&\stackrel{Lem.2.4.3}{\leq} \frac{\xi^2 \|H\|}{1 - \xi \|H\|} \|\Delta w_{\rho_w}(0, 1)\| \\
&\stackrel{\xi < \frac{1}{\|H\|}}{<} \frac{\xi}{1 - \xi \|H\|} \|\Delta w_{\rho_w}(0, 1)\| \\
&\stackrel{\xi < \xi_{\rho_F}}{\leq} \frac{\xi_{\rho_F}}{1 - \xi_{\rho_F} \|H\|} \|\Delta w_{\rho_w}(0, 1)\| \\
&\stackrel{(2.88)}{<} \frac{\kappa \|\Delta w_{\rho_w}(0, 1)\|}{\left(\|\Delta w_{\rho_w}(0, 1)\| + \kappa \|H\|\right) \left(1 - \frac{\kappa \|H\|}{\|\Delta w_{\rho_w}(0, 1)\| + \kappa \|H\|}\right)} \\
&= \kappa.
\end{aligned}$$

Damit ist die Behauptung gezeigt. □

2.4.2 Sensitivität der Komplementaritätsbedingungen

Die Änderungen des Perturbationsparameters τ und damit in den Komplementaritätsbedingungen ist nun Gegenstand der Betrachtungen. Lemma 2.3.6 beschrieb bereits mittels der Differenzvektoren die Veränderung $\Delta\tau = \Delta S \Delta \Pi$ a-posteriori.

Je nach Warmstartabbildung ρ_w oder ρ_F entstehen unterschiedliche Parameteränderungen $\Delta\tau$. Die Differenz kann jedoch eingegrenzt werden.

Lemma 2.4.5 *Es seien S und Π zwei positiv definite Diagonalmatrizen und es gebe Vektoren $\Delta\pi^1$, Δs^1 , $\Delta\pi^2$ und Δs^2 , so dass*

$$\begin{aligned}
S \Delta\pi^1 + \Pi \Delta s^1 &= 0 \\
S \Delta\pi^2 + \Pi \Delta s^2 &= 0
\end{aligned}$$

gilt. Ferner definiere man die Differenzvektoren

$$\begin{aligned} d_\pi &:= \Delta\pi^2 - \Delta\pi^1 \\ d_s &:= \Delta s^2 - \Delta s^1, \end{aligned}$$

welche beschränkt seien, d.h. $\delta := \max\{\|d_\pi\|_\infty, \|d_s\|_\infty\}$. Dann gilt

$$\|\Delta\Pi_2\Delta s^2 - \Delta\Pi_1\Delta s^1\|_\infty \leq \delta(\|\Delta\pi^1\|_\infty + \|\Delta s^1\|_\infty) + \delta^2.$$

Beweis: Es ist

$$\begin{aligned} \Delta\Pi_2\Delta s^2 &= (\Delta\Pi_1 + \text{diag}(d_\pi))(\Delta s^1 + d_s) \\ &= \Delta\Pi_1\Delta s^1 + \Delta\Pi_1d_s + \Delta S_1d_\pi + \text{diag}(d_\pi)d_s \end{aligned}$$

und deshalb

$$\|\Delta\Pi_2\Delta s^2 - \Delta\Pi_1\Delta s^1\|_\infty = \|\text{diag}(d_\pi)d_s + \Delta\Pi_1d_s + \Delta S_1d_\pi\|_\infty \quad (2.93)$$

$$\leq \delta(\|\Delta\pi^1\|_\infty + \|\Delta s^1\|_\infty) + \delta^2 \quad (2.94)$$

□

Korollar 2.4.6 *Unter den Voraussetzungen aus Satz 2.4.4 werden für den Punkt $w_0 = (x_0, \lambda_0, s_0, \pi_0)$ die Differenzvektoren $\Delta w_{\rho_F}(0, \xi) = (\Delta x_{\rho_F}, \Delta \lambda_{\rho_F}, \Delta s_{\rho_F}, \Delta \pi_{\rho_F})$ und $\Delta w_{\rho_w}(0, \xi) = (\Delta x_{\rho_w}, \Delta \lambda_{\rho_w}, \Delta s_{\rho_w}, \Delta \pi_{\rho_w})$ für ein $\xi \in [0, \xi_{\rho_F}[$ berechnet.*

Der Parameter $\varepsilon > 0$ war vorgegeben, so dass

$$\begin{aligned} s_0 + \Delta s_{\rho_w} &> \varepsilon s_0, \\ \pi_0 + \Delta \pi_{\rho_w} &> \varepsilon \pi_0 \end{aligned} \quad (2.95)$$

galten, was im Beweis zu Satz 2.4.4 deutlich wurde. Ausserdem wurde $\kappa = \varepsilon \cdot \min_{i=1, \dots, n} \{(s_0)_i, (\pi_0)_i\}$ eingeführt.

Dann gilt:

$$\|\Delta\Pi_{\rho_F}\Delta s_{\rho_F} - \Delta\Pi_{\rho_w}\Delta s_{\rho_w}\|_\infty \leq \kappa \cdot (\|\Delta\pi_{\rho_w}\|_\infty + \|\Delta s_{\rho_w}\|_\infty) + \kappa^2. \quad (2.96)$$

Beweis: Die Differenzvektoren erfüllen die Voraussetzungen für Lemma 2.4.5 und mit $\delta = \kappa$ folgt die Behauptung. □

Das letzte Resultat legt die Definition einer Umgebung nahe.

Definition 2.4.7 Die Menge

$$\mathcal{T}(\tau, \theta) := \{w = (x, \lambda, s, \pi) \mid \pi, s \geq 0, -\theta\tau \leq \tau - \Pi s \leq \theta\tau\} \quad (2.97)$$

heie θ -**Umgebung** um die τ -Urbilder eines parametrisierten Optimierungsproblems (P_ξ).

Satz 2.4.8 (a-priori Abschtzung II) Es sei $0 < \theta < 1$.

In Satz 2.4.4 gelte zustzlich

$$\varepsilon \in \left[0, \min \left\{ 1, \frac{\sqrt{\theta\tau_{\min}}}{\min_{i=1, \dots, m_g} \{(s_0)_i, (\pi_0)_i\}} \right\} \right], \quad (2.98)$$

mit $\tau_{\min} := \min_{i=1, \dots, n} \{\tau_i\}$ und

$$\alpha := \frac{\theta\tau_{\min} - \kappa^2}{\kappa(\|\Delta\pi_{\rho_w}(0, 1)\|_\infty + \|\Delta s_{\rho_w}(0, 1)\|_\infty)}. \quad (2.99)$$

Dann gilt

$$\rho_F(0, \xi) \in \mathcal{T}(\tau, \theta) \quad \forall \xi \in [0, \min\{\xi_{\rho_F}, \alpha \cdot \xi_{\rho_F}\}]. \quad (2.100)$$

Beweis: Es gilt

$$\begin{aligned} \kappa^2 &= \varepsilon^2 \cdot \min\{(s_0)_i, (\pi_0)_i\}^2 \\ &\stackrel{(2.98)}{<} \theta\tau_{\min} \end{aligned}$$

und damit wegen $\theta\tau_{\min} - \kappa^2 > 0$ auch $\alpha > 0$. Es handelt sich somit um einen zulssigen Multiplikator α .

Da smtliche Voraussetzungen aus Satz 2.4.4 erfllt sind, knnen die Werte

$$w_{\rho_F}(\xi) = (x_{\rho_F}(\xi), \lambda_{\rho_F}(\xi), s_{\rho_F}(\xi), \pi_{\rho_F}(\xi)) := \rho_F(0, \xi)$$

fr alle $\xi \in [0, \xi_{\rho_F}[$ nur zulssige Warmstartpunkte sein. Mit

$$\Delta\tau_{\rho_F}(\xi) := \underbrace{\text{diag}(s_{\rho_F}(0)) \cdot \pi_{\rho_F}(0)}_{=\tau} - \text{diag}(s_{\rho_F}(\xi)) \cdot \pi_{\rho_F}(\xi)$$

bleibt zu zeigen, dass $w_{\rho_F}(\xi)$ Bedingung (2.100) erfllt, d.h. hier

$$-\theta\tau \leq \Delta\tau_{\rho_F}(\xi) \leq \theta\tau \quad \forall \xi \in [0, \min\{\xi_{\rho_F}, \alpha \cdot \xi_{\rho_F}\}]. \quad (2.101)$$

Der Vektor $\Delta\tau_{\rho_F}(\xi)$ kann wie in Lemma 2.3.6 auch mit den Differenzenvektoren

$$\Delta w_{\rho_F}(\xi) = (\Delta x_{\rho_F}(\xi), \Delta \lambda_{\rho_F}(\xi), \Delta s_{\rho_F}(\xi), \Delta \pi_{\rho_F}(\xi)) := \rho_F(0, \xi) - \rho_F(0, 0)$$

beschrieben werden, nämlich durch $\Delta\tau_{\rho_F}(\xi) = \text{diag}(\Delta s_{\rho_F}(\xi)) \cdot \Delta\pi_{\rho_F}(\xi)$.

Andererseits existiert auch der Warmstartpunkt

$$w_{\rho_w}(\xi) = (x_{\rho_w}(\xi), \lambda_{\rho_w}(\xi), s_{\rho_w}(\xi), \pi_{\rho_w}(\xi)) := \rho_w(0, \xi)$$

bzw. der zugehörige Differenzenvektor

$$\Delta w_{\rho_w}(\xi) = (\Delta x_{\rho_w}(\xi), \Delta \lambda_{\rho_w}(\xi), \Delta s_{\rho_w}(\xi), \Delta \pi_{\rho_w}(\xi)) := \rho_w(0, \xi) - \rho_w(0, 0)$$

und mit $\Delta\tau_{\rho_w}(\xi) := \text{diag}(\Delta s_{\rho_w}(\xi)) \cdot \Delta\pi_{\rho_w}(\xi)$ sowie Korollar 2.4.6 gilt die Abschätzung

$$\begin{aligned} \|\Delta\tau_{\rho_F}(\xi) - \Delta\tau_{\rho_w}(\xi)\|_\infty &\leq \kappa(\|\pi_{\rho_w}(\xi)\|_\infty + \|s_{\rho_w}(\xi)\|_\infty) + \kappa^2 \\ &= \kappa \cdot \frac{\xi}{\xi_{\rho_F}} (\|\pi_{\rho_w}(\xi_{\rho_F})\|_\infty + \|s_{\rho_w}(\xi_{\rho_F})\|_\infty) + \kappa^2 \\ &\stackrel{\xi < \alpha \xi_{\rho_F}}{\leq} \kappa \cdot \alpha (\|\pi_{\rho_w}(\xi_{\rho_F})\|_\infty + \|s_{\rho_w}(\xi_{\rho_F})\|_\infty) + \kappa^2 \\ &\stackrel{(2.99)}{\leq} \theta \tau_{\min} \end{aligned}$$

für alle $\xi \in [0, \alpha \cdot \xi_{\rho_F}[$.

□

Letztendlich ist man mit den bisherigen Resultaten in der Lage mittels der Warmstartabbildung ρ_w a-priori eine Parametermenge Y zu bestimmen, für die $\rho_F(0, \xi)$ für alle $\xi \in Y$ zulässige Warmstartpunkte sind. Sogar die θ -Umgebung nach Definition 2.4.7 kann dabei festgelegt werden

Als Fazit können drei Punkte festgehalten werden.

1. Ein a-priori-Kriterium für die Berechnung eines zulässigen Warmstartpunktes mit ρ_F existiert, ist jedoch unbefriedigend, da die Schrittweitenberechnung sehr pessimistisch ist. Tatsächlich gelingen in Praxi, im Allgemeinen weit größere Perturbationen. Dennoch zeigen die obigen Resultate die Durchführbarkeit solcher Berechnungen für jeden Punkt.
2. Die τ -Perturbation bleibt nicht konstant. Bei der Ermittlung der Warmstartpunkte treten Änderungen in den gestörten Komplementaritätsbedingungen auf. Bei mehrmaliger Anwendung addieren sich diese Änderungen und führen, sofern sie nicht korrigiert werden, zu einer ungewollt großen τ -Perturbation.

3. Auftretende Rundungsfehler und Residuen $r(w; \xi)$ müssen wieder reduziert werden.

Im folgenden Abschnitt wird ein Algorithmus vorgestellt, der unter Beachtung der obigen drei Punkte einen praktikablen Ansatz zur Berechnung einer Lösungsmenge vorschlägt.

2.4.3 Die Approximation der Lösungsmenge

Die bisherigen Ausführungen zeigten prinzipiell die Problematik und eine mögliche Methodik bei der Berechnung eines Warmstartpunktes. Das Ziel ist die Approximation der optimalen Lösungen, die aus einem parametrisierten Optimierungsproblem hervorgehen. Dabei besitzt der Warmstartpunkt eine Doppelfunktion. Zum einen ist er nutzbar für die Berechnung weiterer Warmstartpunkte, also weitere Parameterverschiebungen. Zum anderen fungiert er als Startpunkt für ein Optimierungsverfahren, das die Lösung des Optimierungsproblems zum entsprechenden Parameter berechnet.

Abbildung 2.1 visualisiert den Berechnungsvorgang. Zu Anfang wird zu einem Parameter ξ_1 ein τ -zulässiger Punkt w_1 gefordert. Der Störungsvektor der Komplementaritätsbedingungen heißt τ_1 und um die Abhängigkeiten hervorzuheben wird w_1 durch $w(\xi_1, \tau_1)$ beschrieben. Da $w(\xi_1, \tau_1)$ noch nicht die eigentliche Lösung zum Parameter ξ_1 liefert, wird er vorerst als Startpunkt für ein Optimierungsverfahren genutzt, welches das gefragte Optimum $w(\xi_1, 0)$ ermittelt. Für $w(\xi_1, \tau_1)$ existiert unter o.g. Voraussetzungen eine Umgebung $U(\xi_1)$ von ξ_1 , auf der eine Abbildung definiert ist, die Parameter aus $U(\xi_1)$ auf Punkte abbildet, welche zusammen mit dem Parameter eine Nullstelle der Abbildung F_{τ_1} bilden. Wie oben besprochen lässt sich über den Satz der impliziten Funktionen eine Ableitung in $w(\xi_1, \tau_1)$ berechnen, welche die Konstruktion einer Warmstartabbildung erlaubt, mit der ein neuer Parameter ξ_2 und ein neuer Punkt w_2 ermittelt werden kann. Dieser weist im Allgemeinen ebenfalls einen neuen Störungsvektor τ_2 bezüglich der Komplementaritätsbedingungen auf. In der Abbildung wurde der neue Warmstartpunkt mit $w(\xi_2, \tau_2)$ bezeichnet. Damit die Unterschiede zwischen den Störungsvektoren τ_1 und τ_2 kontrollierbar bleiben, muss darauf geachtet werden, dass $w(\xi_2, \tau_2)$ innerhalb der θ -Umgebung $\mathcal{T}(\tau_1, \theta)$ bleibt. Nachdem das nächste Optimum $w(\xi_2, 0)$ ermittelt ist, wird der nächste Warmstartpunkt $w(\xi_3, \tau_3)$ berechnet usw.

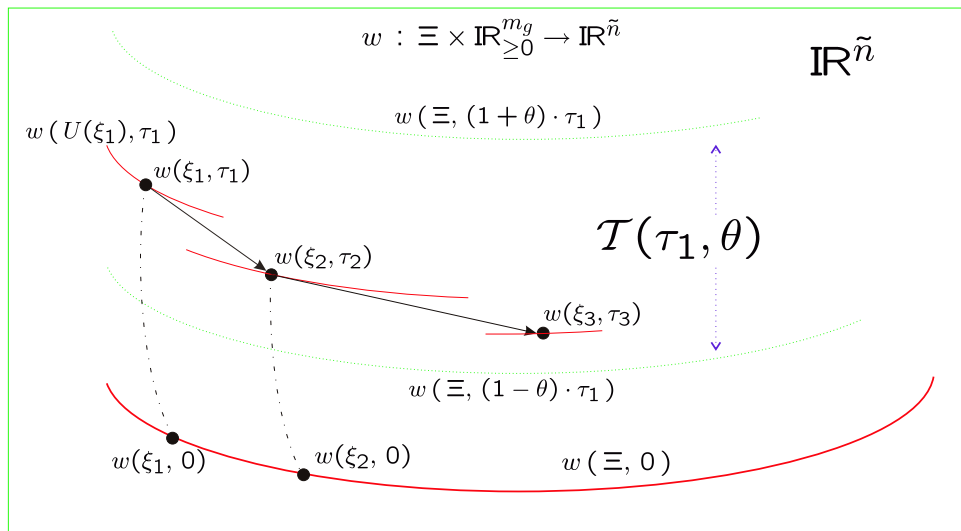


Abbildung 2.1: Approximation der Lösungsmenge

Die Wahl des τ und einer zugehörigen θ -Umgebung verhilft der Berechnung der Warmstartpunkte zu mehr Stabilität. Andererseits, je größer das τ , desto weiter sind die Warmstartpunkte von den optimalen Lösungen entfernt. Zwischen Stabilität und Aufwand muss hier abgewägt werden.

Mit der in Abbildung 2.1 dargestellten Strategie kommt es bei der Approximation der Lösungsmenge u.a. darauf an, welche Parameterwechsel durchgeführt werden, wie die sich aufbauenden Residuen behandelt werden und wie die Warmstartpunkte innerhalb einer θ -Umgebung gehalten werden.

Die Wahl des Optimierungsverfahrens, das die Warmstartpunkte als Iterierte zum Auffinden der optimalen Lösung verwendet, ist hier erst einmal nachrangig und wird in Kapitel 5 besprochen.

Aufgrund der problematischen Parameterwechselbestimmung bei der Warmstartfunktion ρ_F , liefert der folgende Algorithmus einen praktikablen Ansatz mit ρ_w für Optimierungsprobleme wie (2.56.)

Man betrachte die Parametermenge $\Xi = [0, 1]$. Die Komplementaritätsperturbation τ wird mittels eines gegebenen, τ -zulässigen Startpunktes $w_0 = (x_0, \lambda_0, s_0, \pi_0)$ berechnet, indem $\tau = S_0 \Pi_0 e$ berechnet wird. Die Warmstartpunkte werden mit der Warmstartfunktion ρ_w ermittelt. Dabei wird der Parameterwechsel so gewählt, dass der neue Warmstartpunkt wieder zulässig ist und innerhalb einer vorbestimmten θ -Umgebung $\mathcal{T}(\tau, \theta)$ liegt.

Die entstehenden Residuen werden mit einem gedämpften Newtonschritt für F_τ abgeschwächt. Sobald diese nicht mehr durch einen ganzen Newtonschritt korrigiert werden können, wird so lange auf Parameterperturbationen verzichtet, bis das einsetzende gedämpfte Newtonverfahren das Residuum hinreichend reduziert hat.

Algorithmus 2.4.9

INPUT: $w^0 = (x^0, \lambda^0, s^0, \pi^0)$ τ^0 -zulässig bezüglich $(P_{\xi=0})$ mit $\tau^0 := S_0 \Pi_0 e > 0$

SETZE: $\Xi = [0, 1]$, $\theta \in (0, 1)$, $\varepsilon \in (0, 1)$, $\mathcal{W} = \emptyset$.

REPEAT

1. **INITIALISIERE:** $k = 0$, $\xi_k = 0$, $\Delta w^k = 0$, RESID_TERMINATED = FALSE

2. $r^k = F_\tau(w^k; \xi^k)$

3. $\Delta w^{\text{res}} = -M(w^k; \xi_k)^{-1} \cdot r^k$ (Reduzierung des Residuums)

4. **IF** $((s^k + \Delta s^{\text{res}} \geq \varepsilon s^k) \text{ AND } (\pi^k + \Delta \pi^{\text{res}} \geq \varepsilon \pi^k))$

4.a. $\hat{w}^k = \text{OPT}(w^k, \xi^k)$ (\hat{w}^k opt. Lösung von (P_{ξ_k}))

4.b. $\mathcal{W} = \mathcal{W} \cup (\hat{w}^k, \xi^k)$

4.c. $\Delta w^k = -M(w^k; \xi^k)^{-1} \cdot \Delta F(w^k)$ (Differenzvektor Δw_{ρ_w})

4.d. RESID_TERMINATED = TRUE

END IF

5. $\Delta w^k = \Delta w^k + \Delta w^{\text{res}}$

6. **BERECHNE** $\Delta \xi_k$, so dass die Bedingungen

6.a. $s^k + \Delta \xi_k \Delta s^k \geq \varepsilon s^k$ und $\pi^k + \Delta \xi_k \Delta \pi^k \geq \varepsilon \pi^k$

6.b. $w^k + \Delta \xi_k \Delta w^k \in \mathcal{T}(\tau^0, \theta)$

6.c. $\begin{cases} \Delta \xi_k \leq 1 - \xi_k, & \text{IF (RESID_TERMINATED)} \\ \Delta \xi_k \leq 1, & \text{ELSE} \end{cases}$

erfüllt sind.

7. **UPDATE:**

7.a. $w^{k+1} = w^k + \Delta \xi_k \Delta w^k$

7.b. **IF** (RESID_TERMINATED) $\xi_{k+1} = \xi_k + \Delta \xi_k$

ELSE $\xi_{k+1} = \xi_k$

7.c. $k = k + 1$

UNTIL $(\xi_k = 1)$

OUTPUT: \mathcal{W}

Bemerkung 2.4.10 Die Methode OPT in Schritt (4.a) des obigen Algorithmus steht für ein geeignetes Optimierungsverfahren, das mit dem Startpunkt w^k die Lösung \hat{w}^k des parametrisierten Optimierungsproblems zum Parameter ξ^k berechnet. In Kapitel 4 wird diesbezüglich die Mehrotra-Methode, ein Innere-Punkte-Algorithmus, verwendet. Es sind jedoch auch andere Verfahren denkbar.

Lemma 2.4.11 Für jede Iteration k des Algorithmus 2.4.9 existiert eine Schrittweite $\Delta\xi_k > 0$, so dass die Iterierte w^{k+1} den Bedingungen (6.a) und (6.b) des Algorithmus entspricht.

Beweis:

Zu 6.a) Bedingung (6.a) ist erfüllt, wenn $\Delta\xi_k$ so gewählt ist, dass

$$\begin{aligned}\Delta\xi_k &\geq (\varepsilon - 1) \frac{s_i^k}{\Delta s_i^k} && \forall i = 1, \dots, m_g, \\ \Delta\xi_k &\geq (\varepsilon - 1) \frac{\pi_i^k}{\Delta\pi_i^k} && \forall i = 1, \dots, m_g,\end{aligned}$$

gilt. Dabei spielen nur die Komponenten $\Delta\pi_i^k < 0$ und $\Delta s_i^k < 0$ eine Rolle, da ansonsten jedes $\Delta\xi_k \in [0, \infty[$ genügt. Seien also $J_\pi = \{j \mid \Delta\pi_j^k < 0\}$ und $J_s = \{j \mid \Delta s_j^k < 0\}$ die Indexmengen der negativen Komponenten. Dementsprechend definiere man

$$\begin{aligned}\xi_\pi &:= (\varepsilon - 1) \min_{i \in J_\pi} \frac{\pi_i^k}{\Delta\pi_i^k} \quad \text{und} \\ \xi_s &:= (\varepsilon - 1) \min_{i \in J_s} \frac{s_i^k}{\Delta s_i^k}\end{aligned}\tag{2.102}$$

So ist (6.a) für alle

$$\Delta\xi_k \leq \xi_\varepsilon := \min\{\xi_\pi, \xi_s\}$$

erfüllt.

Zu 6.b) Für alle Iterationen gelte $\tau^j := S_j \Pi_j e$, $\forall j = 1, 2, \dots$ und

$$\Delta\tau^j := \tau^j - \tau^{j-1} = \Delta S_j \Delta \Pi_j e.$$

O.B.d.A. gelte $\Delta\tau^k \neq 0$, ansonsten gilt Bedingung (6.b) für beliebige $\Delta\xi_k$ und es gäbe nichts zu zeigen. Für die Iterierte w^{k+1} ist Bedingung (6.b) erfüllt, wenn $\Delta\xi_k$ so gewählt werden kann, dass mit

$$\tau^{k+1} = (\Pi_k + \Delta\xi_k \Delta \Pi_k)(S_k + \Delta\xi_k \Delta S_k)e\tag{2.103}$$

die Abschätzung

$$(1 - \theta)\tau^0 \leq \tau^{k+1} \leq (1 + \theta)\tau^0 \quad (2.104)$$

gilt.

Sofern die boolsche Variable `RESID_TERMINATED = TRUE` gilt, gilt für den Differenzenvektor

$$\Delta w^k = -M(w^k; \xi_k)^{-1} \cdot (r^k + \Delta F(w^k)). \quad (2.105)$$

Dabei sind

$$\Delta F_\tau(w^k) = \begin{pmatrix} D_x \Delta \mathcal{L}(w^k) \\ \Delta h(x^k) \\ \Delta g(x^k) \\ 0 \end{pmatrix}, \quad r^k = F_\tau(w^k) = \begin{pmatrix} r(w^k; \xi_k) \\ S_k \Pi_k e - \tau^0 \end{pmatrix}.$$

Insbesondere besteht wegen (2.105) das Teilsystem

$$\Pi_k \Delta S_k e + S_k \Delta \Pi_k e = \Pi_k S_k e - \tau^0,$$

dass übrigens auch für Δw^k erfüllt ist, wenn `RESID_TERMINATED = FALSE` ist, da die entsprechende rechte Seite aus $\Delta F(w^k)$ der Nullvektor ist.

Dementsprechend ergibt sich mit

$$\begin{aligned} \tau^{k+1} &= \Pi_k S_k e + \Delta \xi_k (\Pi_k \Delta S_k e + S_k \Pi_k e) + \Delta \xi_k^2 \Delta \Pi_k \Delta S_k e \\ &= (1 - \Delta \xi_k) \tau^k + \Delta \xi_k \tau^0 + \Delta \xi_k^2 \Delta \Pi_k \Delta S_k e \\ &\stackrel{\tau^k \in \mathcal{I}(\tau^0, \theta)}{\leq} (1 - \Delta \xi_k) (1 + \theta) \tau^0 + \Delta \xi_k \tau^0 + \Delta \xi_k^2 \Delta \Pi_k \Delta S_k e \\ &= (1 + \theta) \tau^0 - \Delta \xi_k \theta \tau^0 + \Delta \xi_k^2 \Delta \Pi_k \Delta S_k e \end{aligned}$$

eine obere Abschätzung aus (2.103); eine untere Abschätzung folgt analog. Man überlege sich, dass damit bezüglich (2.104) die Abschätzungen

$$(1 - \theta)\tau^0 + \Delta \xi_k \theta \tau^0 + \Delta \xi_k^2 \Delta \tau^k \leq \tau^{k+1} \leq (1 + \theta)\tau^0 - \Delta \xi_k \theta \tau^0 + \Delta \xi_k^2 \Delta \tau^k$$

gelten müssen. Bedingung (2.104) ist erfüllt, wenn

$$\begin{aligned} \Delta \xi_k^2 \Delta \tau^k &\leq \Delta \xi_k \theta \tau^0 \quad \text{und} \\ \Delta \xi_k^2 \Delta \tau^k &\geq -\Delta \xi_k \theta \tau^0, \end{aligned}$$

d.h. sofern

$$\Delta \xi_k \leq \xi_\tau := \theta \cdot \frac{\min_{i=1, \dots, m_g} \tau_i^0}{\|\Delta \tau^k\|_\infty}. \quad (2.106)$$

gilt. Somit existiert ein $\Delta\xi_k \in [0, \xi_\tau[$, so dass Bedingung (6.b) erfüllt ist.

Insgesamt sind mit

$$\Delta\xi_k \in [0, \min\{\xi_\tau, \xi_\varepsilon}\{$$

beide Bedingungen erfüllt.

□

Erläuterung: Algorithmus 2.4.9 beginnt mit einem τ -zulässigen Startpunkt w^0 zum Parameter $\xi = 0$. Anfangs wird die Berechnung von Δw^{res} den Nullvektor ergeben und folglich die boolsche Variable RESID_TERMINATED auf TRUE gesetzt sein. Mit Hilfe der Warmstartfunktion ρ_w wird der Warmstartpunkt $\rho_w(0, 1)$ berechnet.

Lemma 2.4.11 sichert, dass eine Schrittweite $\Delta\xi_k > 0$ existiert, die Bedingung (6.a) und (6.b) erfüllt. Bedingung (6.c) wirkt entweder bei zu großen Residuen, in denen kein Parameterwechsel vollzogen wird oder falls $\Delta\xi_k$ größer als der maximal mögliche Wechsel $1 - \xi_k$ ist. Im zweiten Fall wird der letzte Parameterwechsel erzwungen.

Theoretisch können die Schrittweiten beliebig klein werden. Um eine Mindestschrittweite für alle Iterationen ermitteln zu können, bedarf es einer Abschätzung der Terme ξ_π , ξ_s und ξ_τ über alle Iterationen. Dies ist problematisch, da es dann gilt die Iterierten Δw^k und w^k in der Norm nach oben abzuschätzen. Bei einer zu kleinen Schrittweite sollte deshalb zur Sicherheit ein Neustart mit neuem Startpunkt ermöglicht werden.

Nach der ersten Iteration entsteht in der Regel ein Residuum aufgrund des Perturbationsschritts $\Delta\xi_k$. Diese werden dann zugelassen, wenn sie durch einen ganzen Newtonschritt, d.h. die Berechnung von Δw^{res} vollständig reduziert werden können, wobei $s^k + \Delta s^{\text{res}} \geq \varepsilon s^k$ und $\pi^k + \Delta \pi^{\text{res}} \geq \varepsilon \pi^k$ einzuhalten ist. Diese Bedingung bewirkt, dass Residuen nur so wachsen können, dass sie innerhalb einer Iteration vollständig reduziert werden.

Sofern diese Reduktion nicht vollständig möglich ist, d.h. die Bedingungen in Schritt 4 nicht erfüllt sind, geht der Algorithmus in ein gedämpftes Newtonverfahren zur Reduzierung des Residuums über, indem $\Delta w^k = \Delta w^{\text{res}}$ gesetzt wird. Schritt 3 bildet somit das Newtongleichungssystem bei zu großen Residuen. Die nötige Schrittweite ergibt sich mit $\Delta\xi_k$. Auf Perturbationen des Parameters ξ_k wird solange verzichtet, bis das Residuum wieder hinreichend verkleinert wurde. In der Praxis zeigt sich jedoch, dass in diesen Fällen ein Neustart zu besseren Ergebnissen führt.

In Kapitel 5 werden numerische Tests mit Algorithmus 2.4.9 durchgeführt sowie weitere Modifikationen vorgenommen.

Kapitel 3

Geometrie der Lösungsmenge

Das letzte Kapitel handelte von Lösungen τ -perturbierter KKT-Systeme eines parametrisierten Optimierungsproblems. Unter bestimmten Voraussetzungen konnte für eine nicht näher bestimmte Umgebung U die Existenz und Eindeutigkeit einer C^1 -Abbildung $w : U \rightarrow \mathbb{R}^{\tilde{n}}$ gezeigt werden, die Parametern deren Lösungen zuordnet. Dabei wurde w zusätzlich in ebenfalls stetig-differenzierbare Abbildungen x, λ, s und π unterteilt.

Im Besonderen mit Blick auf die Definitionen A.2.1 und A.1.1 sind diese Abbildungen gleichermaßen parametrisierte Flächen bzw. Kurven. Diese Sichtweise ermöglicht zusätzliche Einblicke in das Verhalten der Lösungsmengen bei Parameterwechseln. Von Interesse sind vor allem singuläre Punkte der Fläche oder Kurve x , weil sich hier die Lösung trotz kleiner Parameterwechsel nicht verändert. Diese Thematik wird in Abschnitt 3.1 behandelt.

Abschnitt 3.2 stellt zwei Methoden zur Bestimmung geeigneter Parameterwechsel bei Kurven dar. Die Parameter sollten idealer Weise so gewählt sein, dass möglichst äquidistante Wegstrecken zwischen den jeweiligen Punkten liegen. Die Ergebnisse werden in Kapitel 5 zur Schrittweitenbestimmung benötigt.

3.1 Lösungserhaltende Perturbationen

Für die Charakterisierung singulärer Punkte betrachte man die differenzierbare, parametrisierte Fläche

$$x : U \rightarrow \mathbb{R}^n$$

mit einer offenen Parametermenge $U \subset \mathbb{R}^k$. Dazu überlege man sich zuerst den Fall, dass x auf die Lösungsmenge eines τ -perturbierten Systems abbildet.

Satz 3.1.1 *Das parametrisierte Optimierungsproblem (P_ξ) sei durch die \mathcal{C}^2 -Zielfunktion $f : \mathbb{R}^n \times \Xi \rightarrow \mathbb{R}$, $\Xi \subseteq \mathbb{R}^k$ und parameterunabhängigen \mathcal{C}^2 -Restriktionsabbildungen $g : \mathbb{R}^n \rightarrow \mathbb{R}^{m_g}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^{m_h}$ gegeben.*

Ferner sei $w_\xi = (x_\xi, \lambda_\xi, s_\xi, \pi_\xi)$ ein τ -zulässiger Punkt zum Parameter $\xi \in \Xi$, so dass die Funktionalmatrizen $D_x g(x_\xi)$ und $D_x h(x_\xi)$ vollen Spalten- bzw. Zeilenrang besitzen und $D_{xx} \mathcal{L}(w_\nu; \xi)$ positiv definit ist.

Dann existiert eine Umgebung U_ξ von ξ und eine \mathcal{C}^1 -Abbildung

$$x : U_\xi \rightarrow \mathbb{R}^n \quad (3.1)$$

für dessen Werte $x(\tilde{\xi})$, $\tilde{\xi} \in U_\xi$ es Vektoren $w_{\tilde{\xi}}, \lambda_{\tilde{\xi}}, s_{\tilde{\xi}}, \pi_{\tilde{\xi}}$ gibt, so dass $w_{\tilde{\xi}} = (x_{\tilde{\xi}}, \lambda_{\tilde{\xi}}, s_{\tilde{\xi}}, \pi_{\tilde{\xi}})$ τ -zulässig ist. Insbesondere ist $x(\xi) = x_\xi$.

Es gilt

$$D_\xi x(\xi) = 0 \quad \Leftrightarrow \quad (D_{\xi x} f(x_\xi; \xi))_{\cdot i} \in \text{Bild}(D_x h(x_\xi)^\top) \quad \forall i = 1, 2, \dots, k.$$

Beweis:

Die Existenz einer Umgebung U_ξ und der Abbildung x folgt direkt aus Lemma 2.2.8(b).

Ähnlich der Notation im Beweis zu Lemma 2.2.8 setze man

$$\begin{aligned} L &:= D_{xx} \mathcal{L}(w_\xi; \xi), & A &:= D_x h(x_\xi), & C &:= D_x g(x_\xi), \\ x' &:= D_\xi x(\xi), & \lambda' &:= D_\xi \lambda(\xi), & s' &:= D_\xi s(\xi) & \pi' &:= D_\xi \pi(\xi) \\ r_L &:= D_{\xi x} \mathcal{L}(w_\xi; \xi), & r_A &:= D_\xi h(x_\xi), & r_C &:= D_\xi g(x_\xi) & r_{S\Pi} &:= D_\xi (S_\xi \Pi_\xi e - \tau). \end{aligned}$$

Da h und g nicht von ξ abhängen und w_ξ τ -zulässig ist, sind r_A , r_C und $r_{S\Pi}$ Nullvektoren und

$$r_L = D_{\xi x} \mathcal{L}(w_\nu; \xi) = D_{\xi x} f(x_\xi; \xi). \quad (3.2)$$

Nach Voraussetzung besitzen A und C vollen Zeilen- bzw. Spaltenrang, L ist positiv semidefinit. Damit kann auch in diesem Fall wie im Beweis von Lemma 2.2.8

$$x' = Q^{-1} r - Q^{-1} A^\top \lambda' \quad (3.3)$$

und

$$\lambda' = (AQ^{-1}A^\top)^{-1}(r_A + AQ^{-1}r) \quad (3.4)$$

berechnet werden, wobei - zur Erinnerung -

$$Q = L + C^\top \Pi_\xi^{-1} S_\xi C$$

und

$$r = -r_L - C^\top \Pi_\nu^{-1} S_\xi (r_C - S_\xi^{-1} r_{S\Pi})$$

eingeführt wurden. Für das Residuum r folgt hier sogar $r = -r_L$, da $r_C = 0$ und $r_{S\Pi} = 0$ gelten, womit wegen (3.3)

$$x' = -Q^{-1}r_L - Q^{-1}A^\top \lambda'. \quad (3.5)$$

und wegen (3.4) und $r_A = 0$

$$\lambda' = -(AQ^{-1}A^\top)^{-1}AQ^{-1}r_L \quad (3.6)$$

folgen.

„ \Rightarrow “

Sei also $x' = 0$, dann folgt

$$-D_{\xi x}f(x_\xi; \xi) \stackrel{(3.2)}{=} -r_L \stackrel{(3.5)}{=} A^\top \lambda'. \quad (3.7)$$

Die Spalten von $D_{\xi x}f(x_\xi; \xi)$ liegen somit im Bild von A^\top .

„ \Leftarrow “

Liegen umgekehrt die Spalten von $D_{\xi x}f(x_\xi; \xi)$ im Bild von $D_x h(x_\xi)^\top = A^\top$, d.h.

$$(D_{\xi x}f(x_\xi; \xi))_{\cdot i} \in \text{Bild}(D_x h(x_\xi)^\top) \quad \forall i = 1, 2, \dots, k, \quad (3.8)$$

so existieren $y_1, y_2, \dots, y_k \in \mathbb{R}^{m_h}$, welche die Gleichungen

$$A^\top y_i = (D_{\xi x}f(x_\xi; \xi))_{\cdot i}, \quad \forall i = 1, 2, \dots, k$$

erfüllen. Es lässt sich daher die Matrix $Y \in \mathbb{R}^{m_h \times k}$, $Y = (y_1, y_2, \dots, y_k)$ definieren, so dass

$$D_{\xi x}f_\tau(w_\xi; \xi) = A^\top Y \quad (3.9)$$

gilt.

Demnach ergibt sich

$$\begin{aligned}
\lambda' &\stackrel{(3.6)}{=} -(AQ^{-1}A^\top)^{-1}AQ^{-1}r_L \\
&= -(AQ^{-1}A^\top)^{-1}AQ^{-1}D_{\xi x}f(x_\xi; \xi) \\
&\stackrel{(3.9)}{=} -(AQ^{-1}A^\top)^{-1}AQ^{-1}A^\top Y \\
&= -Y.
\end{aligned} \tag{3.10}$$

und

$$\begin{aligned}
x' &\stackrel{(3.5)}{=} -Q^{-1}r_L - Q^{-1}A^\top \lambda' \\
&\stackrel{(3.2)}{=} -Q^{-1}D_{\xi x}f(w_\xi; \xi) - Q^{-1}A^\top \lambda' \\
&\stackrel{(3.9)}{=} -Q^{-1}A^\top Y - Q^{-1}A^\top \lambda' \\
&\stackrel{(3.10)}{=} -Q^{-1}A^\top Y + Q^{-1}A^\top Y \\
&= 0.
\end{aligned}$$

□

Singuläre Punkte der parametrisierten Fläche x treten allerdings auch dann auf, wenn x keine Immersion ist, d.h. es existiert ein Parameter ξ , so dass $D_\xi x(\xi)$ keinen vollen Spaltenrang besitzt. Demnach sind die Spaltenvektoren von $D_\xi x(\xi)$ linear abhängig und es muss ein $u \in \mathbb{R}^k$, $u \neq 0$ geben, für das $D_\xi x(\xi) \cdot u = 0$ gilt.

Wegen (3.5) folgt

$$0 = D_\xi x(\xi) \cdot u = -Q^{-1}(r_L - A^\top \lambda') \cdot u,$$

was äquivalent dazu ist, dass die $(n \times k)$ -Matrix

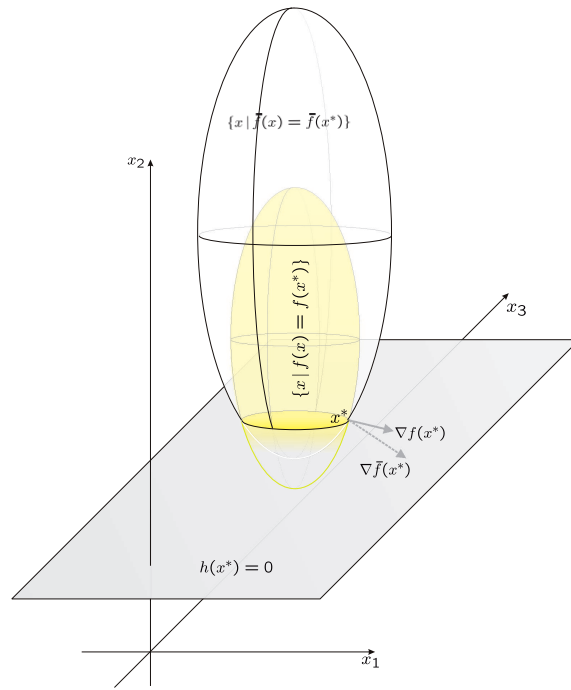
$$(r_L - A^\top \lambda') \stackrel{perDef.}{=} (D_{\xi x}f(x_\xi; \xi) - D_x h(x_\xi)^\top \lambda')$$

ebenfalls keinen vollen Spaltenrang besitzt.

Bei singulären Punkten im mehrdimensionalen Fall kann es von der betrachteten Richtungsableitung abhängen, ob es zu einer Lösungsänderung kommt. Allerdings lässt sich das mehrdimensionale Problem gerade mit Hilfe der Richtungsableitung auf ein 1-dimensionales reduzieren.

Satz 3.1.1 betrachtet ausschliesslich τ -perturbierte Systeme mit strikt positiven τ . Das Resultat wird nun für $\tau = 0$ erweitert. Dafür betrachte man Optimierungsprobleme, bei denen im Optimum $w^* = (x^*, \lambda^*, s^*, \pi^*)$ die strenge Komplementaritätsbedingungen gelten. D.h., wenn

$$s_i^* > 0 \quad \Leftrightarrow \quad g(x_i^*) = \pi_i^* = 0. \tag{3.11}$$

Abbildung 3.1: Lösungserhaltende Perturbation für τ -perturbierte Systeme

Aufgrund dieser Bedingung lassen sich analog zu Lemma 2.1.2 die Indexmengen $J := \{1, \dots, n\}$,

$$J_\pi(w^*) := \{i \in J \mid s_i = 0\} \quad (3.12)$$

$$J_s(w^*) := J \setminus J_0(w^*) \quad (3.13)$$

definieren. Dementsprechend setze man die Variablen

$$s_N = (s_i^*)_{i \in J_\pi(w^*)}, \quad \pi_N = (\pi_i^*)_{i \in J_\pi(w^*)}, \quad (3.14)$$

$$s_B = (s_i^*)_{i \in J_s(w^*)}, \quad \pi_B = (\pi_i^*)_{i \in J_s(w^*)}. \quad (3.15)$$

Ohne Einschränkung kann angenommen werden, dass die Komponenten so geordnet sind, dass

$$s = \begin{pmatrix} s_N \\ s_B \end{pmatrix}, \quad \pi = \begin{pmatrix} \pi_N \\ \pi_B \end{pmatrix} \quad (3.16)$$

gelten. Genauso teilen sich

$$A = (A_N, A_B) \quad \text{und} \quad D_{\xi x} f(x^*; \xi) = \begin{pmatrix} \Delta f_N(x^*; \xi) \\ \Delta f_B(x^*; \xi) \end{pmatrix} \quad (3.17)$$

auf.

Satz 3.1.2 *Wie in Satz 3.1.1 setze sich das parametrisierte Optimierungsproblem (P_ξ) aus den C^2 -Abbildungen $f : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^{m_g}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^{m_h}$, zusammen.*

Der KKT-Punkt $w^ = (x^*, \lambda^*, s^*, \pi^*)$ zum Parameter ξ sei bekannt und erfülle die strengen Komplementaritätsbedingungen. Die Funktionalmatrizen $D_x g(x^*)$ und $D_x h(x^*)$ besitzen vollen Spalten- bzw. Zeilenrang und $D_{xx} \mathcal{L}(w^*; \xi)$ sei positiv definit.*

Dann existiert eine Umgebung U_ξ von ξ und eine C^1 -Abbildung

$$x : U_\xi \rightarrow \mathbb{R}^n \quad (3.18)$$

dessen Werte $x(\tilde{\xi})$, wiederum zu KKT-Punkten zum Parameter $\tilde{\xi}$ gehören und $x(\xi) = x^$ gilt.*

Es gilt

$$D_\xi x(\xi) = 0 \quad \Leftrightarrow \quad (\Delta f_N(x^*; \xi))_{.i} \in \text{Bild}(A_N^\top) \quad \forall i = 1, 2, \dots, k. \quad (3.19)$$

Beweis: Die Existenz einer Abbildung x folgt sofort aus Lemma 2.2.8.

O.B.d.A sei $g(x) = -x$, womit $x = \pi$ folgt. Es gelte wieder die Notation aus dem Beweis von Lemma 2.2.8 und Satz 3.1.2, womit sich insbesondere $x' = D_\xi x(\xi)$ berechnen lässt. Nach Gleichungssystem (2.24) und weil $L = D_{xx} F(w^*; \xi)$ positiv definit ist, ergibt sich

$$x'(\xi) = L^{-1}(-A^\top \lambda' + s'(\xi) - D_{\xi x} f(x^*; \xi))$$

Da die strengen Komplementaritätsbedingungen gelten, können die Variablen mit Hilfe der Mengen $J_\pi(w)$ und $J_s(w)$ aufgeteilt werden. Zusätzlich definiere man

$$x'(\xi) = \begin{pmatrix} x'_N \\ x'_B \end{pmatrix} \quad \text{und} \quad s'(\xi) = \begin{pmatrix} s'_N \\ s'_B \end{pmatrix}.$$

Es folgt dann

$$Sx'(\xi) + Xs'(\xi) = \begin{pmatrix} S_N x'_N + X_N s'_N \\ S_B x'_B + X_B s'_B \end{pmatrix} \stackrel{x_B = s_N = 0}{=} \begin{pmatrix} X_N s'_N \\ S_B x'_B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Da s_N und x_B Nullvektoren sowie x_N und s_B strikt positiv sind, müssen $s'_N = 0$ und $x'_B = 0$ gelten. Die Matrizen s'_B und x'_N könnten beliebig gewählt sein.

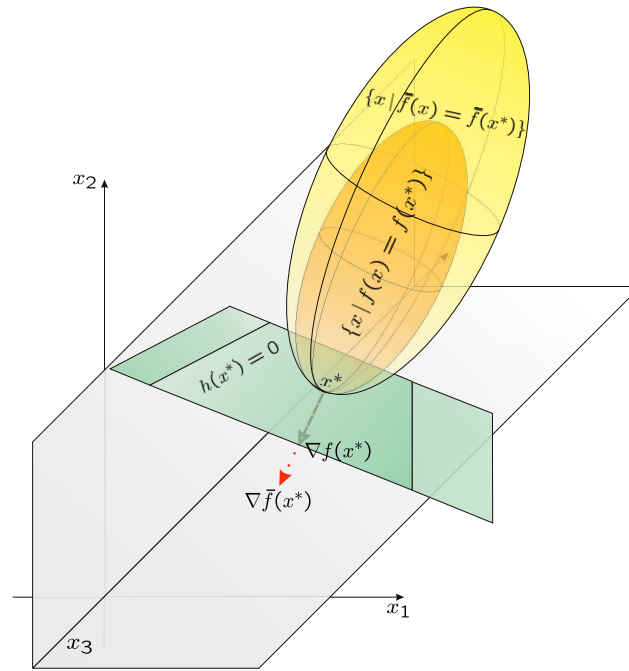


Abbildung 3.2: Lösungserhaltende Perturbation für innere Punkte

Die Ableitung $x'(\xi)$ ist wegen der Regularität von L genau dann Null, wenn der Vektor $-A^\top \lambda'(\xi) + s'(\xi) - D_{\xi x} f(x^*; \xi)$ ebenfalls Null ist. Genauer heißt das

$$\begin{aligned} A_N^\top \lambda'(\xi) &= -\Delta f_N(x^*; \xi) \\ A_B^\top \lambda'(\xi) - s'_B(x^*; \xi) &= -\Delta f_B. \end{aligned} \quad (3.20)$$

Also müssen die Spalten von $\Delta f_N(x^*; \xi)$ im Bild von A_N^\top liegen. Zwingend ergibt sich

$$s'_B = A_B^\top \lambda'(\xi) + \Delta f_B(x^*; \xi).$$

Liegen die Spalten von $\Delta f_N(x^*; \xi)$ nicht im Bild von A^\top , kann Gleichung (3.20) nicht erfüllt sein, was wiederum $x' \neq 0$ nach sich zieht.

□

Bemerkung 3.1.3

Ist die Matrix $L = D_w F(w; \xi)$ positiv semidefinit, können mehrere Lösungen zu einem Parameter ξ existieren. In diesem Fall reicht es nicht aus, dass $-A^\top \lambda'(\xi) + s'(\xi) - \Delta f(w; \xi)$ wie im obigen Beweis gleich Null ist, da auch jeder angenommene Wert aus dem Kern von L diese Bedingung erfüllte.

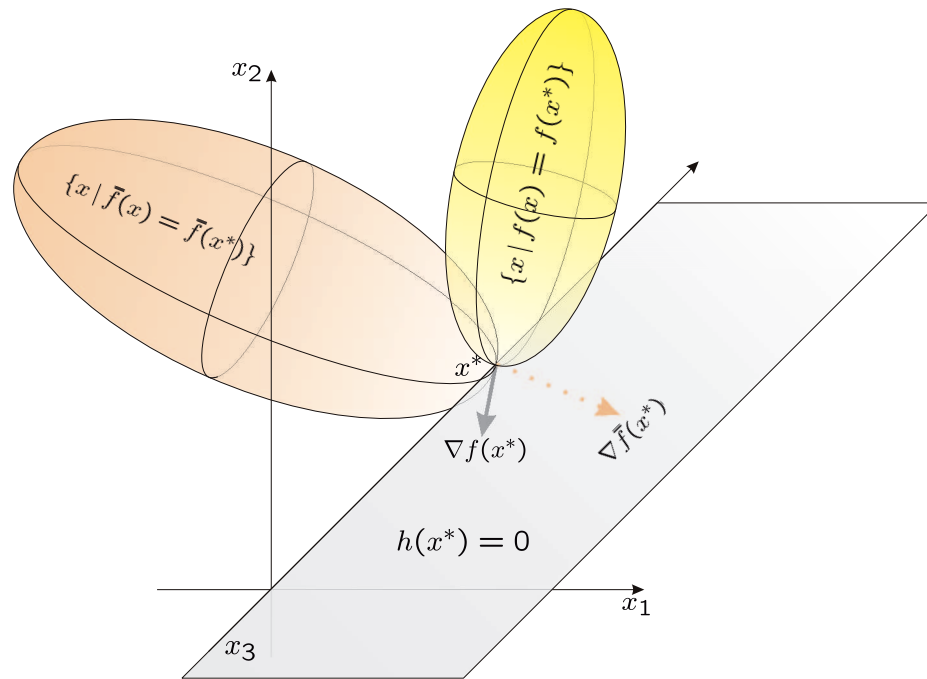


Abbildung 3.3: Allgemeine lösungserhaltende Perturbation

Allerdings ergibt sich mit der Aufteilung $L = (L_N, L_B)$ das Gleichungssystem

$$L_N x_N = -A_N^\top \lambda'(\xi) - \Delta f_N(x^*; \xi). \quad (3.21)$$

Sofern L_N vollen Spaltenrang besitzt, ist x_N genau dann Null, wenn $\Delta f_N(x^*; \xi)$ im Bild von A_N^\top liegt.

Bei nicht regulären Flächen $\varphi : U \rightarrow \mathbb{R}^n$ sind insbesondere zusammenhängende Mengen $K \subset U$ von Interesse, deren Elemente zu singulären Punkten von φ führen. Gerade in den folgenden Kapiteln wird die Erkennung solcher Kompakta eine Rolle spielen.

Definition 3.1.4 Eine differenzierbare Fläche $\varphi : U \rightarrow \mathbb{R}^n$ besitzt genau dann einen **Knick** auf einer zusammenhängenden, nichtleeren Menge $K \subset U$, wenn φ in allen Punkten $\xi \in K$ singulär ist. Der Knick heißt **echt**, falls K aus mehr als einem Punkt besteht.

3.2 Bestimmung von Parameteränderungen

Bei der Darstellung von Kurven mit Hilfe einer endlichen Anzahl von Punkten, werden zu verschiedenen Parametern Auswertungen vorgenommen, so dass die ent-

stehende Punktwolke ein möglichst gutes Bild der Kurve wiedergibt. Wieviele und welche Parameter nötig sind, damit die approximierende Punktwolke der tatsächlichen Menge gerecht wird, hängt von der gewünschten Genauigkeit ab. Sicherlich ist unabhängig davon eine äquidistante Kurvenstrecke zwischen den jeweiligen Punkten wünschenswert.

Wie jedoch bereits Satz 3.1.1 und 3.1.2 zeigen, existieren auch Beispiele, in denen Parameteränderungen zu keinerlei Veränderungen führen. Umgekehrt können sehr kleine Parameterperturbationen zu verhältnismäßig großen Änderungen bezüglich der sich ergebenden Werte bzw. Punkte führen. Letztendlich ist dies eine Frage der Geschwindigkeit, der ersten Ableitung der Kurve im jeweiligen Punkt. Die Sensitivität kann also - zumindest für kleine Perturbationen - mit Hilfe von Ableitungsinformationen gemessen werden.

Dazu werden differenzierbare und reguläre parametrisierte Kurven

$$\varphi : I \rightarrow \mathbb{R}^n \quad \xi \mapsto \varphi(\xi) \quad (3.22)$$

betrachtet. So können Wege auf Flächen durch Kurven beschrieben werden. Tatsächlich erlauben Richtungsableitungen das Abtasten der Fläche mit Hilfe von Kurven.

Eine praktikable jedoch unter Umständen ungenaue Messung der Änderung $\Delta\varphi := \varphi(\xi) - \varphi(\tilde{\xi})$ ist mittels einer Norm $\|\cdot\|$ des zugrundeliegenden Vektorraums gegeben. Ein genaueres Maß liefert sicherlich das Bogenmaß

$$L_{\xi}^{\tilde{\xi}}(\varphi) := \int_{\xi}^{\tilde{\xi}} \|\varphi'(t)\| dt, \quad (3.23)$$

da es die Form und den Verlauf der Kurven beachtet.

Dabei ist klar, dass die Sekante zwischen zwei Punkten kleiner oder gleich der Bogenlänge sein muss, denn es gilt

$$\|\varphi(\tilde{\xi}) - \varphi(\xi)\| = \left\| \int_{\xi}^{\tilde{\xi}} \varphi'(t) dt \right\| \quad (3.24)$$

$$\leq \int_{\xi}^{\tilde{\xi}} \|\varphi'(t)\| dt \quad (3.25)$$

$$= L_{\xi}^{\tilde{\xi}}(\varphi) \quad (3.26)$$

Allerdings erweist sich die Berechnung der Bogenlänge meist aufwendiger als das eigentlich zu lösende Problem.

Eine weitere Schwierigkeit liegt darin, die Bogenlänge L der gesamten Kurve zu schätzen. Denn erst damit kann eine Maßeinheit definiert werden, mit der sich die Kurve in sinnvolle Abstände unterteilen lässt. Soll die Kurve mittels k äquidistanter Punkte dargestellt werden, liegen diese mit einer Länge von $l := \frac{k}{L}$ voneinander entfernt.

Wie aber berechnet man eine Parameteränderung $\Delta\xi$, so dass die Differenzbogenlänge $L_{\xi}^{\xi+\Delta\xi}(\varphi) \approx l$ beträgt?

Tangentiale Schrittweitenbestimmung: Der Tangentialvektor $\varphi(\xi)$ liefert eine Schätzung der zu erwartenden Sekantenlänge $\|\varphi(\tilde{\xi}) - \varphi(\xi)\|$ einer zugrundeliegenden Parameterperturbation von ξ auf $\tilde{\xi}$.

Aufgrund der Taylorapproximation erster Ordnung mit Restglied $R = o(\|\tilde{\xi} - \xi\|)$ gilt für die Sekantenlänge

$$\|\varphi(\tilde{\xi}) - \varphi(\xi)\| = \|(\tilde{\xi} - \xi) \cdot \varphi'(\xi) + R\| \quad (3.27)$$

$$= |\tilde{\xi} - \xi| \cdot \|\varphi'(\xi)\| + o(\|\tilde{\xi} - \xi\|). \quad (3.28)$$

Mit einer vorgegebenen Mindeststrecke l zwischen den Punkten $\varphi(\xi)$ und $\varphi(\tilde{\xi})$ beträgt die nötige Parameteränderung $\Delta\xi = \tilde{\xi} - \xi$ somit ungefähr

$$\Delta\xi = \frac{l}{\|\varphi'(\xi)\|}. \quad (3.29)$$

Ist die sich ergebene Sekantenlänge größer dem vorgegebenen l , d.h.

$$\|\varphi(\xi + \Delta\xi) - \varphi(\xi)\| \geq l,$$

können beispielsweise durch ein Bisektionsverfahren weitere Punkte berechnet werden, die auf der Kurve zwischen $\varphi(\xi)$ und $\varphi(\xi + \Delta\xi)$ liegen. Alternativ verwende man den Quotienten

$$\alpha := \frac{\|\varphi(\xi + \Delta\xi) - \varphi(\xi)\|}{\|\varphi'(\xi)\|} \quad (3.30)$$

und berechne den Punkt $\varphi(\xi + \alpha\Delta\xi)$.

Die tangentielle Schrittweitenbestimmung ist umso genauer, je mehr die Kurve einer Geraden gleicht.

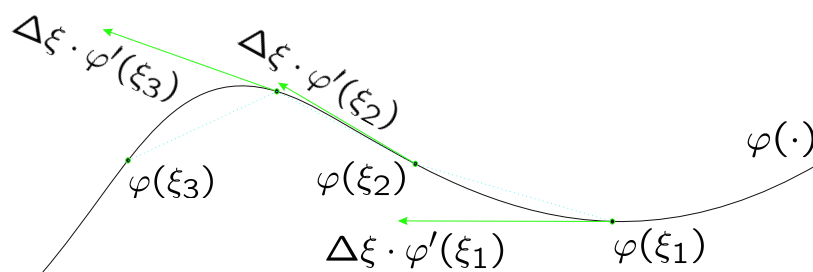


Abbildung 3.4: Tangentiale Schrittweiten

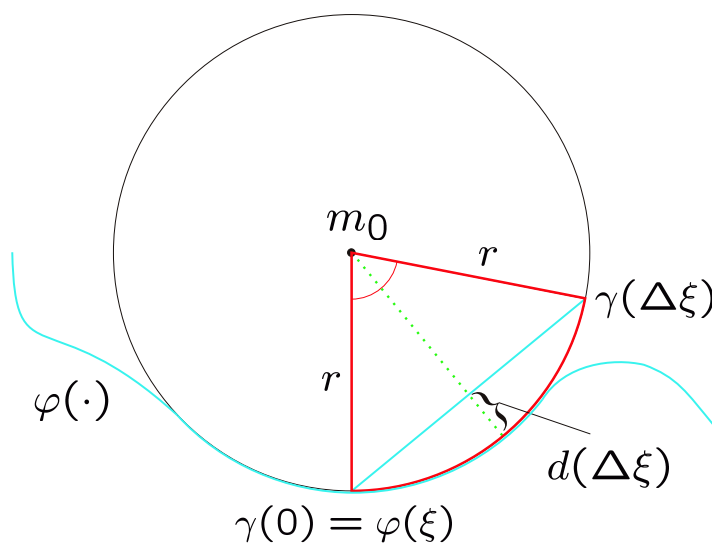


Abbildung 3.5: Zirkuläre Schrittweiten

Zirkuläre Schrittweitenbestimmung: Eine verbesserte Methode für die Abschätzung von Parameterperturbationen ergibt sich durch Hinzunahme von Informationen zweiter Ordnung, der Krümmung $\kappa(\xi)$ einer Kurve in einer Ebene. In der Differentialgeometrie wurde dieser Ansatz u.a. von Wu und Andrade [WuAn99] oder Stoyanov [Sto92] beschrieben.

Sofern die zweite Ableitung $\varphi''(\xi)$ bekannt ist und damit auch die Krümmung, lassen sich Schmiegekreise (s. Definition A.1.5) mit Radius $r := \frac{1}{\kappa}(\xi)$ in den jeweiligen Punkten $\varphi(\xi)$ anlegen. Der Schmiegekreis

$$\gamma : [0, 2\pi r] \rightarrow \mathbb{R}^2, \quad \phi \mapsto r \cdot \begin{pmatrix} \cos\left(\frac{\phi}{r}\right) \\ \sin\left(\frac{\phi}{r}\right) \end{pmatrix}$$

ist wiederum eine parametrisierte Kurve und es gilt $\varphi(\xi) = \gamma(0)$. Der Mittelpunkt m_0 aus Definition A.1.5 sei o.B.d.A. der Nullpunkt.

Die Bogenlänge eines Kreisstücks $L_0^\psi(\gamma)$, $\psi \in [0, 2\pi r]$ ist sehr leicht zu bestimmen, da die Kurve γ nach Bogenmaß parametrisiert ist, d.h.

$$\|\gamma'(\psi)\|_2 = \sqrt{\sin^2\left(\frac{\psi}{r}\right) + \cos^2\left(\frac{\psi}{r}\right)} = 1$$

für alle $\psi \in [0, 2\pi r]$. Dementsprechend gilt

$$\begin{aligned} L_0^\psi(\gamma) &= \int_0^\psi \|\gamma'(t)\|_2 dt \\ &= \psi. \end{aligned} \tag{3.31}$$

Dieses Konstrukt (s. Abbildung 3.5) ermöglicht zum einen die Abschätzung der Sekantenlänge $\|\varphi(\xi + \Delta\xi) - \varphi(\xi)\|$ einer Parameteränderung $\Delta\xi$, indem die leicht berechenbare Sekantenlänge $\|\gamma(\Delta\xi) - \gamma(0)\|$ ermittelt wird. Abbildung 3.6 verdeutlicht die Bestimmung der Sekantenlänge l , die sich bei einem Winkel α wegen $\sin\left(\frac{\alpha}{2}\right) = \frac{l}{2r}$ durch

$$l = 2r \sin\left(\frac{\alpha}{2}\right)$$

errechnet. Der Winkel α ergibt sich mit

$$\alpha = \frac{L_0^\psi(\gamma)}{r} = \frac{\psi}{r}.$$

Eine vorgelegte Sekantenlänge \bar{l} , ermöglicht damit die Schätzung einer maximalen Perturbation von

$$\Delta\xi = 2r \cdot \arcsin\left(\frac{\bar{l}}{2r}\right).$$

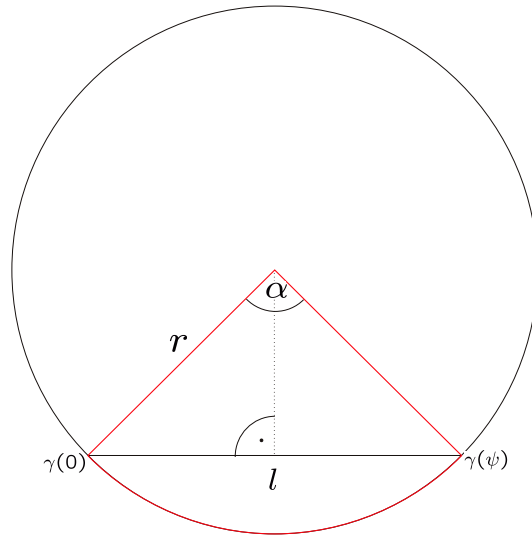


Abbildung 3.6: Die Kreissekante

Zum anderen kann ein maximaler Abstand

$$d(\Delta\xi) = \max_{t \in [0, \Delta\xi]} \min_{\lambda \in [0, 1]} \|\gamma(t) - \lambda\gamma(0) - (1 - \lambda)\gamma(\Delta\xi)\|_2$$

der Sekante zwischen $\gamma(0)$ und $\gamma(\Delta\xi)$ zum Kreisrand vorgegeben werden, der es verhindert, dass die Sekantenlänge eine zu schlechte Abschätzung liefert. Geometrisch ist klar, dass

$$d(\Delta\xi) = r - \left\| m_0 + \frac{1}{2}(\gamma(0) - \gamma(\Delta\xi)) \right\|_2 \quad (3.32)$$

$$\stackrel{m_0=0}{=} \frac{r^2}{2} \cdot \left\| \begin{pmatrix} \cos\left(\frac{0}{r}\right) - \cos\left(\frac{\Delta\xi}{r}\right) \\ \sin\left(\frac{0}{r}\right) - \sin\left(\frac{\Delta\xi}{r}\right) \end{pmatrix} \right\|_2 \quad (3.33)$$

$$= \frac{r}{2} \cdot \sqrt{1 + 2 \cos\left(\frac{\Delta\xi}{r}\right) + \underbrace{\cos^2\left(\frac{\xi}{r}\right) + \sin^2\left(\frac{\Delta\xi}{r}\right)}_{=1}} \quad (3.34)$$

$$= \frac{r}{\sqrt{2}} \cdot \sqrt{1 + \cos\left(\frac{\Delta\xi}{r}\right)} \quad (3.35)$$

gelten muss.

Je mehr die Krümmung gegen Null geht, desto mehr verhält sich φ lokal entsprechend einer Geraden. Der Kreis - gegeben durch die Parameterdarstellung γ - entartet, da der Radius gegen unendlichen strebt. In diesem Fall kann die zirkuläre Schrittweitenbestimmung nicht ohne Weiteres genutzt werden. Hier wäre z.B. eine vorgegebene Beschränkung der Schrittweite nötig. Ein fallsensitives Ausweichen auf die tangentielle Methode böte sich ebenso an.

Kapitel 4

Multikriterielle Optimierung

Die Theorie der vorherigen Kapitel wird nun am Beispiel der multikriteriellen Optimierung vertieft. Es wird sich zeigen, dass diese Optimierungsaufgaben eine spezielle Problemklasse der parametrisierten Optimierung bilden. Diese Sichtweise wurde bisher insbesondere von Guddat [GuVaJo90], [GuVaTaWe85] verfolgt. Allgemeinere Ansätze liefert Rheinboldt [Rhei86], [Rhei74].

Wie sich in Kapitel 3 herausstellte, können Lösungsmengen parametrisierter Optimierungsprobleme unter bestimmten Voraussetzungen als parametrisierte Fläche interpretiert werden. Dieser Aspekt führt speziell in der multikriteriellen Optimierung zu interessanten Blickwinkeln und Resultaten. Vor allem Hillermeier [Hil00] und Rheinboldt [Rhei86] geben den Anstoß zu einer differentialgeometrischen Betrachtung.

4.1 Grundlagen

Unter multikriterieller Optimierung, auch Mehrziel- oder Vektoroptimierung genannt, versteht man die Ermittlung optimaler Lösungen einer vektorwertigen Zielfunktion

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^k, \quad x \mapsto \begin{pmatrix} f_1(x) \\ \vdots \\ f_k(x) \end{pmatrix} \quad (4.1)$$

bezüglich einer Restriktionsmenge X . Eine optimale Lösung entsteht durch simultanes minimieren der Funktionen f_1 bis f_k über X . Dies führt in der Regel nicht

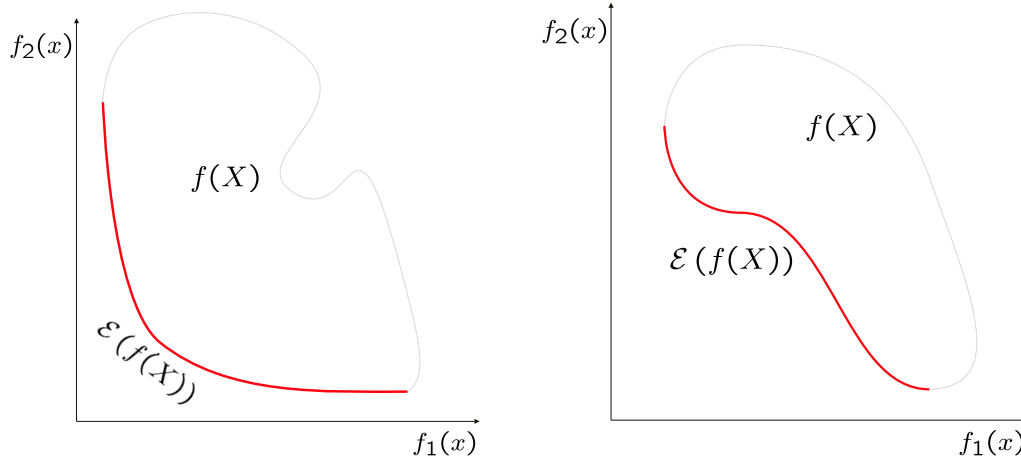


Abbildung 4.1: Effizienzmengen multikriterieller Optimierungsprobleme

auf eindeutige Lösungen. Vielmehr treten im Allgemeinen negative Korrelationen zwischen den k Zielfunktionen auf, d.h. die Verbesserung einer Zielfunktion zieht die Verschlechterung mindestens einer anderen nach sich.

Funktionswerte einer reellwertigen Funktion lassen sich durch die kanonische totale Ordnung zwischen reellen Zahlen vergleichen. Für zwei Werte $z_1, z_2 \in \mathbb{R}$ gilt stets

$$z_1 \leq z_2 \quad \text{oder} \quad z_2 \leq z_1. \quad (4.2)$$

Die Funktionswerte von f bilden k -dimensionale Vektoren zu deren Vergleich die natürlich induzierte Halbordnung ' \leq ' des \mathbb{R}^k zur Verfügung steht. Für solche Halbordnungen gelten Reflexivität, Transitivität und Antisymmetrie, jedoch nicht Bedingung (4.2). Folglich existiert oft keine optimale Lösung im Sinne der skalaren Optimierung. Alternativ sind Punkte $x \in X$, welche die Bedingung

$$f(x) \leq f(y) \quad \text{für alle } y \in X \quad (4.3)$$

erfüllen, interessant.

Die folgende Definition komplettiert den Lösungsbegriff in der multikriteriellen Optimierung.

Definition 4.1.1 Sei Z eine beliebige Menge, halbgeordnet durch " \leq ". Ein Element $\bar{z} \in Z$ heißt **effizient**, sofern kein weiteres Element $z \in Z$, $z \neq \bar{z}$ existiert, so dass

$$z \leq \bar{z}. \quad (4.4)$$

Die Menge

$$\mathcal{E}(Z) := \{z \in Z \mid z \text{ ist effizient}\}$$

heißt **Effizienzmenge** (s. Abbildung 4.1). Ist die Menge Z eindeutig vorgegeben, genügt die Schreibweise $\mathcal{E} = \mathcal{E}(Z)$.

Die **Pareto**menge zu einer gegebenen Abbildung f definiert sich durch

$$\mathcal{P}(f, X) := \{x \in X \mid f(x) \in \mathcal{E}(f(X))\}.$$

Entsprechend heißt ein Element $x \in \mathcal{P}(f, X)$ **Paretopunkt**.

Die Darstellungsform eines multikriteriellen Optimierungsproblems sei für den weiteren Verlauf durch die Schreibweise

$$\begin{aligned} f(x) \rightarrow \min \\ x \in X. \end{aligned} \tag{4.5}$$

gegeben.

4.2 Optimalitätsbedingungen

Für die multikriterielle Optimierung lassen sich - wie in der skalaren Optimierung - ebenfalls Optimalitäts- und Regularitätsbedingungen aufstellen. Der Begriff des KKT-Punkts kann entsprechend verallgemeinert werden, so dass eine Charakterisierung der Paretopunkte ermöglicht wird. Für eine ausführliche Besprechung zu Optimalitätsbedingungen in der multikriteriellen Optimierung siehe z.B. [GoeNeh90], [Mie99] und bezüglich der Optimierung mit reellwertiger Zielfunktion sei hier [GeiKan02] genannt.

Die zulässige Restriktionsmenge X ist jetzt durch die \mathcal{C}^1 -Abbildungen $g : \mathbb{R}^n \rightarrow \mathbb{R}^{m_g}$ und $h : \mathbb{R}^n \rightarrow \mathbb{R}^{m_h}$ gegeben, so dass

$$X := \{x \in \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0\} \tag{4.6}$$

gilt. Für den weiteren Verlauf sei X wie oben definiert.

Definition 4.2.1 (Abadie CQ) Zu einer gegebenen Menge $X \subset \mathbb{R}^n$, $X \neq \emptyset$ und einem Punkt $x^* \in X$ heißt die Menge

$$T(X, x^*) := \text{cone} \left\{ d \in \mathbb{R}^n \mid \exists \{x_n\}_{n \in \mathbb{N}} \subset X, x_n \rightarrow x^*, \frac{x_n - x^*}{\|x_n - x^*\|} \rightarrow d \right\} \tag{4.7}$$

Tangentialkegel an X im Punkt x^* .

Die Menge

$$T_{\text{lin}}(X, x^*) := \text{cone} \{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^\top d \leq 0, i \in I(x^*), \nabla h(x^*)^\top d = 0\} \quad (4.8)$$

heißt **Linearisierungskegel**, wobei $I(x^*) = \{i \in \{1, \dots, m_g\} \mid g_i(x^*) = 0\}$ der Indexmenge der aktiven Ungleichungsrestriktionen entspricht.

Der zulässige Punkt x^* genügt der **Regularitätsbedingung von Abadie** (Abadie-Constraint-Qualification), wenn

$$T(X, x^*) = T_{\text{lin}}(X, x^*) \quad (4.9)$$

gilt.

Die Abadie CQ wird oft auch als Kuhn-Tucker CQ bezeichnet. Aus der Optimierungstheorie ist bekannt, dass stets

$$T(X, x^*) \subseteq T_{\text{lin}}(X, x^*) \quad (4.10)$$

gilt. Mit erfüllten Regularitätsbedingungen wie die Abadie CQ sind die Kegel gleich, was zu folgenden notwendigen und hinreichenden Optimalitätskriterien für multi-kriterielle Optimierungsprobleme führt.

Satz 4.2.2 (Optimalitätsbedingungen)

Gegeben sei $x^* \in X$.

Es sei $x^* \in X$ ein Paretopunkt bezüglich (4.5). Ist in x^* die Abadie CQ erfüllt, so existieren die Vektoren

$$\xi \in \mathbb{R}_{\geq 0}^k \setminus \{0\}, \quad \lambda \in \mathbb{R}^{m_h}, \quad s \in \mathbb{R}_{\geq 0}^{m_g} \quad (4.11)$$

mit den Eigenschaften

$$\nabla f(x^*)^\top \xi + \nabla h(x^*)^\top \lambda + \nabla g(x^*)^\top s = 0 \quad (4.12)$$

$$s^\top g(x^*) = 0. \quad (4.13)$$

Sind f und g konvex, sowie h affin linear, und es existieren die oben angeführten Vektoren $\xi \in \mathbb{R}_{>0}^k$, λ , s , mit denen die Bedingungen (4.11), (4.12) und (4.13) für $x \in X$ erfüllt sind, so ist x ein Paretopunkt.

Beweis: siehe [GoeNeh90], Satz 2.35.

Hillermeier führt in [Hil00] die gleiche Optimalitätsbedingung ein, jedoch basierend auf der Mangasarian-Fromowitz CQ (MFCQ), d.h. die Gradienten der aktiven Restriktionen sind linear unabhängig. Da indes die MFCQ-Bedingung die Abadie CQ impliziert, ist die Aussage in Satz 4.2.2 allgemeiner.

Definition 4.2.3 *Mit der Restriktionsmenge X bildet*

$$\mathcal{K}(f, X) := \{(x, \xi, \lambda, s) \mid x \in X, (x, \xi, \lambda, s) \text{ erfüllt (4.11) – (4.13)}\} \quad (4.14)$$

*die Menge aller **KKT-Punkte** des multikriteriellen Optimierungsproblems (4.5).*

Bemerkung 4.2.4 *Existiert ein Vektor $(x, \xi, \lambda, s) \in \mathcal{K}(f, X)$, so erfüllt auch $(x, \beta \cdot \xi, \beta \cdot \lambda, \beta \cdot s)$ mit beliebigen $\beta > 0$ diese Bedingungen. Zur Vermeidung dieser Mehrdeutigkeiten wird zusätzlich gefordert, dass die Punkte ξ ausschließlich der Menge*

$$\Omega := \left\{ \xi \in \mathbb{R}^k \mid \xi \geq 0, \sum_{i=1}^k \xi_k = 1 \right\} \quad (4.15)$$

zugehören.

Satz 4.2.2 bildet die theoretische Grundlage für eine Methode zur Berechnung der Paretopunkte, der **Skalarisierungsmethode**. Wenn zu jedem Paretopunkt x^* , die Regularitätsbedingungen erfüllt sind und damit notwendigerweise ein KKT-Punkt $(x, \xi, \lambda, s) \in \mathcal{K}(f, X)$ existiert, gilt

$$\mathcal{P}(f, X) \subseteq \text{Proj}(\mathcal{K}(f, X)). \quad (4.16)$$

Dabei handelt es sich bei Proj um die Projektion

$$\text{Proj} : \mathcal{K}(f, X) \rightarrow \mathbb{R}^n, \quad \text{Proj}(x, \xi, \lambda, s) \mapsto x \quad (4.17)$$

Sind zusätzlich f, g konvex und h affin, gilt wegen Satz 4.2.2 sogar Gleichheit zwischen den Mengen in (4.16).

Die Skalarisierungsmethode basiert auf der Tatsache, dass der KKT-Punkt (x, ξ, λ, s) des multikriteriellen Problems als Äquivalent den KKT-Punkt (x, λ, s) des skalaren Minimierungsproblems besitzt

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \xi^\top f(x) \\ \text{N.B} \quad & h(x) = 0 \\ & g(x) \leq 0. \end{aligned} \quad (4.18)$$

Insofern lassen sich die KKT-Punkte des multikriteriellen Optimierungsproblems mit Hilfe entsprechender skalarer Optimierungsprobleme berechnen.

4.3 Pareto- und Effizienzmengen

Dieser Abschnitt beschreibt die Zusammenhänge zwischen der multikriteriellen Optimierung, deren Begriffe und den Inhalten aus Kapitel 2 und 3. Aufgrund spezieller Eigenschaften und Bedingungen multikriterieller Optimierungsprobleme, lassen sich weiterführende Resultate ableiten, die auf denen der letzten Kapitel aufbauen.

Zuerst wird die Menge der KKT-Punkte $\mathcal{K}(f, X)$ näher betrachtet. Die Benennung des Vektors ξ der KKT-Bedingungen (4.12) mit Blick auf den gleichbenannten Parametervektor in parametrisierten Optimierungsproblemen wurde nicht ohne Grund vorgenommen. Tatsächlich übernimmt der Vektor ξ auch hier die Rolle des Parameters eines parametrisierten Optimierungsproblems, wie es am skalarisierten Ersatzproblem (4.18) deutlich wird. Unter diesem Aspekt stellt sich ein Problem der multikriteriellen Optimierung, das gleichzeitig skalarisierbar ist, als parametrisches Optimierungsproblem dar. Die Perturbationen bedingt durch den Parameter ξ wirken sich dabei lediglich auf die Zielfunktionen aus; die Restriktionen bleiben erhalten.

Zur Erinnerung: Mittels des Nullstellenproblems (2.8) lieferte die Abbildung

$$F : \mathbb{R}^{\tilde{n}+k} \rightarrow \mathbb{R}^{\tilde{n}}, \quad (w, \xi) \mapsto F(w; \xi)$$

die Menge der KKT-Punkte des parametrisierten Optimierungsproblems (2.56).

Aufgrund der speziellen Wahl der Parametervektoren $\xi \in \Omega$ entspricht $\mathcal{K}(f, X)$ der Lösungsmenge des Nullstellenproblems

$$\tilde{F}(x, \lambda, s, \pi, \xi) := \begin{pmatrix} F(w; \xi) \\ \sum_{i=1}^k \xi_i - 1 \end{pmatrix} = 0, \quad \xi > 0, \pi, s \geq 0. \quad (4.19)$$

Hillermeier zeigt in [Hil00], Kapitel 5.1, dass die Lösungsmenge der impliziten Gleichung (4.19) unter gewissen Voraussetzungen einer differenzierbaren, $(k-1)$ -dimensionalen Untermannigfaltigkeit entspricht. Die Abbildung \tilde{F} ist dort allerdings ausschließlich mit Gleichungsrestriktionen definiert. Der folgende Satz erweitert das Resultat für zusätzliche Ungleichungsrestriktionen.

Satz 4.3.1 *Es sei $\tilde{\mathcal{K}} \subseteq \mathcal{K}(f, X)$ eine Menge für deren Punkte $\tilde{w} := (w; \xi) \in \tilde{\mathcal{K}}$ die Jakobimatrix $D_w F(w; \xi)$ regulär ist und die strikten Komplementaritätsbedingungen gelten, so handelt es sich mit $\xi \in \mathbb{R}^k$ bei $\tilde{\mathcal{K}}$ um eine $(k-1)$ -dimensionale differenzierbare Untermannigfaltigkeit.*

Beweis: Nach Definition A.2.2 ist erstens zu zeigen, dass für jeden Punkt $\tilde{w} = (w, \xi)$ der Menge $\tilde{\mathcal{K}}$ bez. des Rangs der Funktionalmatrix $\text{rang}(D\tilde{F}(\tilde{w})) = n - k$ gilt; zweitens muss für jedes $\tilde{w} \in \tilde{\mathcal{K}}$ eine Umgebung U existieren mit $F(\tilde{w}) = 0$ für alle $\tilde{w} \in \tilde{\mathcal{K}} \cap U$.

Die Jacobimatrix der Abbildung

$$\tilde{F} : \mathbb{R}^{\tilde{n}+k} \rightarrow \mathbb{R}^{\tilde{n}+1}$$

hat die Gestalt

$$D\tilde{F}(w, \xi) = \begin{pmatrix} D_w F(w; \xi) & D_\xi F(w; \xi) \\ 0 & e^\top \end{pmatrix}$$

Die Matrix $D_w F(w; \xi)$ ist nach Voraussetzung regulär, d.h. ihr Rang beträgt \tilde{n} . Wegen der zusätzlichen unteren Zeile der Matrix $D\tilde{F}(w, \xi)$ hat diese den Rang $\tilde{n} + 1$.

Bleibt noch zu zeigen, dass für jeden Punkt $\tilde{w} \in \tilde{\mathcal{K}}$ eine Umgebung $U \subset \mathbb{R}^{\tilde{n}+k}$ existiert mit der die Bedingung

$$\tilde{\mathcal{K}} \cap U = \left\{ u \in U \mid \tilde{F}(u) = 0 \right\} \quad (4.20)$$

erfüllt ist.

Infolge der regulären Matrix $D_w F(w; \xi)$ existieren aufgrund der Aussage des Satzes A.0.1 Umgebungen $V \subset \mathbb{R}^k$ und $W \in \mathbb{R}^{\tilde{n}}$, so dass die Existenz einer eindeutig bestimmten \mathcal{C}^1 -Abbildung

$$w : V \rightarrow W \quad (4.21)$$

mit $F(w(\xi); \xi) = 0$ gesichert ist.

Lemma 2.1.2 liefert aufgrund der strikten Komplementaritätsbedingungen die Existenz einer Umgebung V_1 , dessen Elemente wiederum auf strikt komplementäre KKT-Punkte abgebildet werden.

Da $\tilde{w} = (w, \xi)$ Bedingung (4.19) erfüllt, wobei ξ aus der Menge Ω stammt, ist auch der Schnitt $V_2 = \Omega \cap V_1 \cap V$ nicht leer.

Man setze

$$U = V_2 \times W.$$

Dann ist Bedingung (4.20) wegen der Eindeutigkeit von w erfüllt.

□

Gleichzeitig können die Teilabbildungen x, λ, s und π als parametrisierte Flächen verstanden werden. Im nächsten Satz wird dies der Einfachheit halber lediglich für x gezeigt.

Satz 4.3.2 *Zu lösen ist das multikriterielle Optimierungsproblem (4.5) mit einer strikt konvexen Abbildung $f \in \mathcal{C}^2$. Die Restriktionsmenge X sei durch eine konvexe Abbildung $g \in \mathcal{C}^1$ und einer affin linearen h gegeben. In jedem Paretopunkt gelten die strikten Komplementaritätsbedingungen, und die Funktionalmatrix $D_w F(w; \xi)$ sei für alle $\xi \in \Omega$, $w \in \mathbb{R}^{\tilde{n}}$ regulär.*

Dann existiert eine offene Umgebung U und eine eindeutig bestimmte k -dimensionale, differenzierbare, parametrisierte Fläche

$$x : U \rightarrow \mathbb{R}^n, \quad U \supset \Omega$$

und es gilt

$$\mathcal{P}(f, X) = x(\Omega) \tag{4.22}$$

Beweis:

Aufgrund der Beschaffenheit des Optimierungsproblems folgt aus Satz 4.2.2, dass zu jedem Paretopunkt x^* Vektoren λ^*, s^*, ξ^* existieren, die mit $\tilde{w}^* = (x^*, \xi^*, \lambda^*, s^*)$ einen KKT-Punkt bilden. Umgekehrt liefert jeder KKT-Punkt \tilde{w}^* den Paretopunkt x^* .

Gleichzeitig ist $w^* = (x^*, \lambda^*, s^*)$ genau dann KKT-Punkt des skalaren Optimierungsproblems (4.18) zum Parameter ξ^* , wenn \tilde{w}^* KKT-Punkt des multikriteriellen Optimierungsproblems ist.

Da f strikt konvex und X konvexe Menge sind, ist x^* genau dann optimale Lösung von (4.18) zum Parameter ξ^* , wenn Lagrangemultiplikatoren λ^*, s^* existieren, mit denen (x^*, λ^*, s^*) KKT-Punkt des skalaren Problems ist. Das Optimum x^* ist sogar strikt globale Lösung und wird dem ξ^* somit eindeutig zugeordnet.

Deshalb existiert eine eindeutig bestimmte Abbildung

$$w : \Omega \rightarrow \mathbb{R}^{\tilde{n}}$$

und damit im speziellen auch die Abbildung

$$x : \Omega \rightarrow \mathbb{R}^n,$$

die Parametern entsprechende KKT-Punkte bzw. strikt globale Optima zuteilen.

Sowohl w als auch x sind differenzierbar nach ξ , denn mit Satz A.0.1 und Lemma 2.1.2 folgt für jedes $\xi \in \Omega$ die Existenz einer Umgebung $U_\xi \subset \mathbb{R}^k$ von ξ und eindeutig

bestimmter \mathcal{C}^1 -Abbildungen $w_\xi : U_\xi \rightarrow \mathbb{R}^n$ bzw. $x_\xi : U_\xi \rightarrow \mathbb{R}^n$. Aufgrund der Eindeutigkeit muss

$$w|_{U_\xi} = w_\xi|_{U_\xi} \quad \text{und} \quad x|_{U_\xi} = x_\xi|_{U_\xi}$$

gelten. Insgesamt sind die Abbildungen w und x stetig-differenzierbar und eindeutig bestimmt auf der offenen Menge

$$U = \bigcup_{\xi \in \Omega} U_\xi \supset \bigcup_{\xi \in \Omega} \{\xi\} = \Omega.$$

Nach Satz 4.2.2 und Bemerkung 4.2.4 werden alle KKT-Punkte über die Parametermenge Ω erreicht. Damit sind auch alle Paretopunkte bekannt. Aufgrund der Konvexität des zugrundeliegenden Problems muss somit

$$\mathcal{P}(f, X) = x(\Omega)$$

gelten.

□

Bemerkung 4.3.3 *Bereits in Satz 3.1.2 wurde aufgezeigt, dass x lediglich stückweise regulär ist. Insbesondere wenn die Werte x an Ecken oder am Rand der Restriktionsmenge liegen, können singuläre Punkte entstehen.*

Sofern eine Abbildung $x : \Omega \rightarrow \mathbb{R}^n$ im Sinne von (4.22) existiert, ist die Effizienzmenge

$$\mathcal{E} := \mathcal{E}(f(X)) = \mathcal{E}(f(\mathcal{P}(f, X))) \stackrel{(4.22)}{=} \mathcal{E}(f(x(\Omega)))$$

wegen der stetig-differenzierbaren Zielfunktionen f_1, f_2, \dots, f_k ebenfalls eine parametrisierte Fläche, nämlich

$$\varphi_\mathcal{E} : \Omega \rightarrow \mathbb{R}^k, \quad \xi \mapsto f(x(\xi)) = \begin{pmatrix} f_1(x(\xi)) \\ \vdots \\ f_k(x(\xi)) \end{pmatrix}. \quad (4.23)$$

An den regulären Punkten werden die Tangentialebenen $T_{x(\xi)}\mathcal{E}$ durch die Spalten der Funktionalmatrix

$$\varphi'_\mathcal{E}(\xi) = \nabla f(x(\xi))^\top x'(\xi) \quad (4.24)$$

aufgespannt.

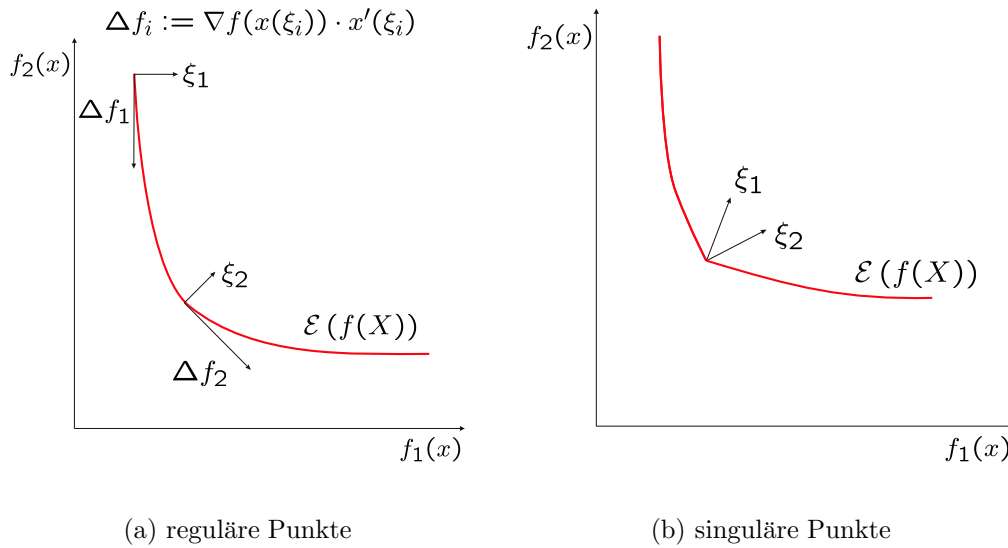


Abbildung 4.2: Parametervektoren und Effizienzmengen

Zur geometrischen Interpretation des Skalarisierungsparameters ξ bezüglich der Effizienzmenge \mathcal{E} gibt es bereits Resultate. Das und Dennis [DasDen96] beschränkten sich auf den bikriteriellen Fall. Hillermeier [Hil00] konnte die Ergebnisse für allgemeine multikriterielle Optimierungsprobleme mit Restriktionen erweitern. Unter anderem stützt sich Hillermeiers Resultat dabei auf die Forderung, dass eine Umgebung existiert, in der aktive Restriktionen ihren Status behalten. Im folgenden Lemma ist diese Voraussetzung aufgrund der Annahme strikt erfüllter Komplementaritätsbedingungen im Zusammenhang mit Lemma 2.1.2 erfüllt.

Satz 4.3.4 *Man betrachte Aufgabe (4.5) mit der Restriktionsmenge X aus (4.6). Die Abadie CQ sowie die strikten Komplementaritätsbedingungen seien in allen Paretopunkten erfüllt. Auf einer offenen Menge $U \subset \Omega$ existieren stetig-differenzierbare Abbildungen $x(\xi)$, $\lambda(\xi)$ und $s(\xi)$, die allen Skalierungsparametern $\xi \in \Omega$ die entsprechenden Paretopunkte bzw. Lagrangemultiplikatoren zuordnen. Dann gilt*

$$\xi^\top \nabla f(x(\xi))^\top x'(\xi) = 0. \quad (4.25)$$

Beweis: O.B.d.A. gebe es lediglich Ungleichungsrestriktionen $g(x) \leq 0$. Jedem Paretopunkt $x := x(\xi)$, $\xi \in U$ ist dann aufgrund der vorausgesetzten Abadie CQ und der Aussage von Satz 4.2.2 ein KKT-Punkt $(x(\xi), \xi, s)$ zugeordnet, in dem die Gleichung

$$\xi^\top \nabla f(x)^\top = -s^\top \nabla g(x)^\top \quad (4.26)$$

erfüllt ist. Somit gilt insbesondere

$$\xi^\top \nabla f(x)^\top x'(\xi) = -s^\top \nabla g(x)^\top x'(\xi). \quad (4.27)$$

Bleibt also zu zeigen, dass $s^\top \nabla g(x)^\top x'(\xi) = 0$ gilt.

Aufgrund der strikten Komplementaritätsbedingungen folgt mit Lemma 2.1.2, dass eine offene Umgebung $V \subset \Omega$ von ξ existiert, so dass der inaktive bzw. aktive Status der Restriktionen unverändert bleibt und x wieder zu einem KKT-Punkt gehört. Dementsprechend können die disjunkten Indextmengen J_s aus (2.13) und J_π aus (2.14) gebildet werden.

Gleichungssystem (4.27) ist damit äquivalent zu

$$\xi^\top \nabla f(x)^\top x'(\xi) = - \sum_{i \in J_s(\xi)} s_i \nabla g_i(x)^\top x'(\xi). \quad (4.28)$$

Da sich die Menge $J_s(\xi)$ für alle $\xi \in V$ nicht ändert, reicht es gerade die aktiven Restriktionen zu betrachten. Eine Änderung der Restriktion $g_i(x(\xi))$ durch den Parameter ξ wird durch die partielle Ableitung

$$D_\xi g(x) = \nabla g(x)^\top x'(\xi) \quad (4.29)$$

beschrieben. Da einem Parameter $\tilde{\xi}$ der Umgebung V ebenfalls der Funktionswert $g_i(x(\tilde{\xi})) = g_i(x(\xi)) = 0$ zugeordnet ist, kann die Ableitung $D_\xi g(x)$ nur Null sein. Insofern muss

$$\xi^\top \nabla f(x)^\top x'(\xi) = 0 \quad (4.30)$$

gelten.

□

Abbildung 4.2 stellt den Zusammenhang zwischen dem Skalarisierungsvektor ξ und der Ableitung $\nabla f(x)^\top x'(\xi)$ graphisch dar.

Um die singulären Stellen der Effizienzmenge, welche durch die parametrisierte Fläche φ_ε beschrieben wird, zu charakterisieren, gibt das nächste Korollar einen allgemeineren Blickwinkel auf Kurven, die wie φ_ε konstruiert sind. Auf eine Erweiterung von Kurven auf Flächen wurde hier wieder verzichtet, da singuläre Stellen dann auch in der Gestalt von Funktionalmatrizen mit nichtvollem Spaltenrang auftreten und dies die Sachlage verkomplizierte.

Korollar 4.3.5 Eine parametrisierte Kurve $\varphi : \Omega \rightarrow \mathbb{R}^k$, $\Omega \subset \mathbb{R}$ sei durch die C^1 -Abbildungen $f : \mathbb{R}^n \rightarrow \mathbb{R}^k$ und $x : \Omega \rightarrow \mathbb{R}^n$ gegeben, so dass

$$\xi \mapsto \varphi(\xi) = f(x(\xi))$$

gilt. Damit ist auch φ differenzierbar und genau dann singulär im Punkt $\xi \in \Omega$, wenn eine der drei Bedingungen

$$(a) \quad \nabla f(x(\xi)) = 0,$$

$$(b) \quad x'(\xi) = 0,$$

$$(c) \quad \nabla f_i(x(\xi)) \perp x'(\xi), \forall i = 1, 2, \dots, k, \nabla f(x(\xi)) \neq 0, x'(\xi) \neq 0.$$

gilt.

Beweis: Die Fläche $\varphi = f \circ x$ ist singulär in einem Punkt $\xi \in \Omega$, wenn die Ableitung $\varphi'(\xi) = \nabla f(x(\xi))^\top x'(\xi)$ dem Nullvektor entspricht. Das kann nur dann der Fall sein, wenn eine der Bedingungen (a), (b) oder (c) erfüllt ist. □

Gerade für numerische Anwendungen zur Approximation einer Effizienzmenge ist die Erkennung echter Knicke von großer Wichtigkeit, da in diesem Fall die Parameter einer zusammenhängenden Menge K auf den gleichen Wert abgebildet werden und damit einen unnötigen Mehraufwand mit sich bringen können. In diesem Sinne betrachte man erneut die in Korollar 4.3.5 aufgeführten Bedingungen:

Zu 4.3.5(a): Dieser Fall ist für konvexe Problemstellungen entartet. Zur Verdeutlichung gehe man davon aus, dass f konvex ist. Aufgrund der Konvexität ist jeder Parameter ξ mit $\nabla f(x(\xi)) = 0$ ein globales Minimum, dass innerhalb der Restriktionsmenge angenommen wird. Gleichzeitig liefert ξ für alle f_i , $i = 1, 2, \dots, k$ ein globales Minimum innerhalb der Restriktionsmenge. Das hieße jedoch auch, dass die optimale Lösung $x = x(\xi)$ für alle $\nu \in \Omega$ angenommen würde und die Effizienzmenge lediglich aus dem Element x bestünde.

Zu 4.3.5(b): In Bemerkung 4.3.3 wurde darauf hingewiesen, dass x i.A. nicht überall regulär ist, d.h. es existieren $\xi \in \Omega$ mit $x'(\xi) = 0$. Existiert ein echter Knick auf einer Menge $K \subset \Omega$, so gilt $x'(\xi) = 0$ für alle $\xi \in K$. Demnach muss ebenfalls

$$\varphi'_\mathcal{E}(\xi) = \nabla f(x(\xi))^\top x'(\xi) = 0$$

für alle $\xi \in K$ gelten. Womit auch $\varphi_\mathcal{E}$ in K einen echten Knick besitzt.

Zu 4.3.5(c): Die parametrisierte Fläche $\varphi_{\mathcal{E}} = f \circ x$ in (4.23) habe einen echten Knick auf der zusammenhängenden Menge K und für alle $\xi \in K$ gelte o.B.d.A. $\nabla f_i(x(\xi)) \perp x'(\xi)$, $\forall i = 1, 2, \dots, k$, mit $\nabla f(x(\xi)) \neq 0$ und $x'(\xi) \neq 0$ (ansonsten s. Bedingung 4.3.5(a),(b)). Die Ersatzprobleme (4.18) seien so gestellt, dass für jeden Parameter $\xi \in K$ genau eine optimale Lösung $x(\xi)$ existiere. Da K zusammenhängend ist und Teilmenge von \mathbb{R} , kann es sich nur um ein Intervall handeln. Für zwei Parameter $\xi, \nu \in K$, $\xi \leq \nu$ existiert folglich nach dem Mittelwertsatz ein $\phi \in [\xi, \nu]$, so dass

$$\begin{aligned} \varphi_{\mathcal{E}}(\nu) - \varphi_{\mathcal{E}}(\xi) &= \varphi'_{\mathcal{E}}(\phi) \cdot (\nu - \xi) \\ &\stackrel{\varphi'_{\mathcal{E}}(\phi)=0, \forall \phi \in K}{=} 0 \end{aligned}$$

Demnach folgt

$$\varphi_{\mathcal{E}}(\nu) = \varphi_{\mathcal{E}}(\xi) \quad \text{d.h.} \quad f(x(\nu)) = f(x(\xi)).$$

Somit ist das Bild $x(K)$ eine Teilmenge der Niveaumengen von f , d.h. es gilt

$$x(K) \subseteq N_i(\eta_i) := \{x \in \mathbb{R}^n \mid f_i(x) = \eta_i\}, \quad \eta_i = f_i(x(\xi)) \quad (4.31)$$

für alle $i = 1, 2, \dots, k$.

Da die zugrundeliegenden Ersatzprobleme (4.18) denselben Restriktionen unterliegen, sind alle Bildelemente von x zulässig, insbesondere alle $y \in x(K)$. Ferner sind alle $y \in x(K)$ auch optimale Lösungen, da sie aus den gleichen Niveaumengen $N_i(f_i(x(\xi)))$, $i = 1, 2, \dots, k$ stammen. Das wäre jedoch ein Widerspruch zur vorausgesetzten Eindeutigkeit der optimalen Lösung. Gilt Bedingung 4.3.5(c), kann somit kein echter Knick existieren.

Die letzten Beobachtungen werden im folgenden Korollar zusammengefasst.

Korollar 4.3.6 Die Effizienzmenge sei durch \mathcal{C}^1 -Abbildungen $f : \mathbb{R}^n \rightarrow \mathbb{R}^2$, und $x : \Omega \rightarrow \mathbb{R}^n$, $\Omega \subset \mathbb{R}$ wie in (4.23) durch eine parametrisierte Kurve

$$\varphi_{\mathcal{E}} : \Omega \rightarrow \mathbb{R}^2, \quad \xi \mapsto f(x(\xi))$$

beschrieben.

Ist x eine reguläre Kurve und besitzt $\varphi_{\mathcal{E}}$ keine echten Knicke, so ist $\varphi_{\mathcal{E}}$ regulär.

Beweis: Da x regulär ist, kann die Effizienzmenge in Form der parametrisierten Kurve $\varphi_{\mathcal{E}}$ nur singuläre Stellen besitzen, wenn die Fälle (a) oder (b) aus Korollar

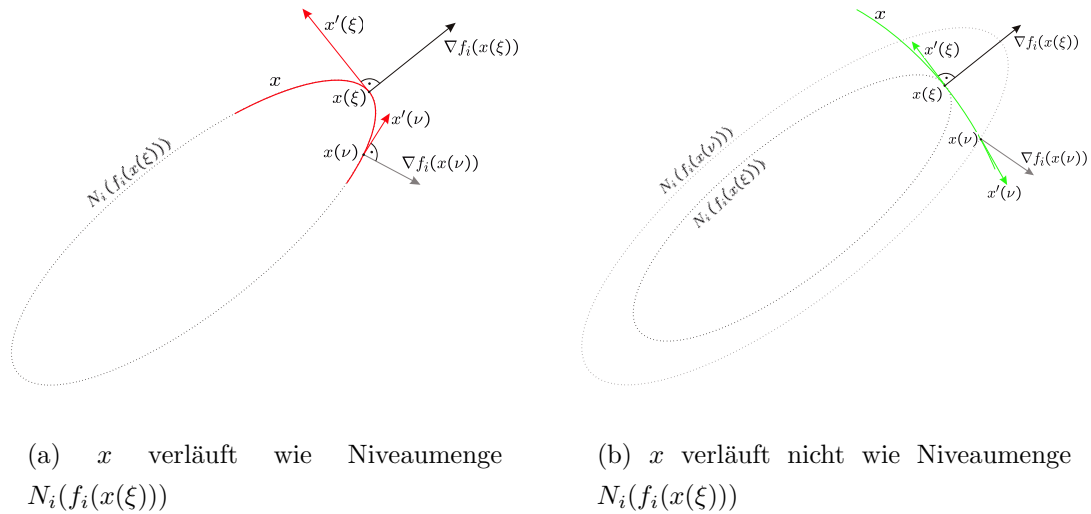


Abbildung 4.3: Fallunterscheidung: $\nabla f_i(x(\xi)) \perp x'(\xi)$

4.3.5 erfüllt sind. Diese Bedingungen führen jedoch (wie oben angeführt) zu Knicken, die nicht echt und nach Voraussetzung ausgeschlossen sind. Die Effizienzmenge muss demnach eine reguläre Kurve sein.

□

Die Parameter $\xi \in \Omega$ sind Normaleneinheitsvektoren der Tangentialebene $T_{f(x(\xi))}\mathcal{E}_U$ bezüglich $\|\cdot\|_1$, d.h. insbesondere ξ liegt in der Menge $N_{f(x(\xi))}\mathcal{E}_U$ der Normalenvektoren von \mathcal{E}_U an $x(\xi)$. Die Tangentialebene $T_{f(x(\xi))}\mathcal{E}_U$ an \mathcal{E}_U in $x(\xi)$ wird von der Ableitung $\nabla f(x(\xi)) \cdot x'(\xi)$ aufgespannt.

Satz 4.3.7 *Zum multikriteriellen Optimierungsproblem (4.5) existiere die Pareto-mengenabbildung $x : U \rightarrow \mathbb{R}^n$ als parametrisierte Fläche auf einer offenen Menge $U \subset \Omega$. Die Effizienzmenge bezüglich der Parametermenge U*

$$\mathcal{E}_U := \mathcal{E}(f(x(U)))$$

sei eine reguläre Hyperfläche.

Falls die Abadie CQ und die strikten Komplementaritätsbedingungen für jeden Paretopunkt erfüllt sind, die Restriktionsmenge X konvex ist, sowie die Zielfunktionen f_i , $i = 1, \dots, k$ konvexe Funktionen bilden, so ist \mathcal{E}_U eine konvexe Hyperfläche.

Beweis:

Die Menge Ω enthält sämtliche Elemente $\xi \in \mathbb{R}_{\geq 0}^k$, deren Komponentensumme $\sum_{i=1}^k \xi_i$

eins ergibt. Dadurch gilt für die k -te Komponente immer

$$\xi_k = 1 - \sum_{i=1}^{k-1} \xi_i.$$

Demnach existiert eine isomorphe Abbildung zwischen den Mengen Ω und

$$\tilde{\Omega} := \left\{ \xi \in \mathbb{R}_{\geq 0}^{k-1} \mid \sum_{i=1}^{k-1} \xi_i \leq 1 \right\}.$$

So kann x auch als Abbildung

$$x : \tilde{\Omega} \rightarrow \mathbb{R}^n$$

betrachtet werden, und damit ist

$$f \circ x : \tilde{\Omega} \rightarrow \mathbb{R}^k$$

nach Definition A.2.4 eine Hyperfläche.

Um die konvexe Eigenschaft der Hyperfläche nach Definition A.2.7 zu zeigen, überlege man sich zuerst, dass für differenzierbare konvexe Funktionen $H : Z \rightarrow \mathbb{R}$ mit konvexem Definitionsbereich Z stets

$$H(y) - H(x) \geq \nabla H(x)^\top (y - x) \quad (4.32)$$

für alle $x, y \in Z$ gilt.

Nach Satz 4.3.4 stehen die Parametervektoren ξ senkrecht auf der Tangentialebene, wie in Abbildung 4.2 dargestellt. Per Definition besteht die Effizienzenge \mathcal{E}_U nur aus der Teilmenge des Bildes von f , deren Urbild aus Paretopunkten $\mathcal{P}(f, U)$ besteht.

O.B.d.A. gebe es lediglich Ungleichungsnebenbedingungen $g(x) \leq 0$. Da nach Voraussetzung die Abadie CQ in jedem Paretopunkt x erfüllt sind, existieren Vektoren s und ξ , so dass (x, ξ, s) aus der Menge der KKT-Punkte $\mathcal{K}(f, X)$ stammen.

Für beliebige $x, y \in \mathcal{P}(f, U)$ betrachte man

$$\begin{aligned}
\xi^\top (f(y) - f(x)) &= \xi^\top \begin{pmatrix} f_1(y) - f_1(x) \\ \vdots \\ f_k(y) - f_k(x) \end{pmatrix} \\
&\stackrel{(4.32), \xi \geq 0}{\geq} \xi^\top \begin{pmatrix} \nabla f_1(x)^\top \\ \vdots \\ \nabla f_k(x)^\top \end{pmatrix} (y - x) \\
&\stackrel{(4.12)}{=} -s^\top \begin{pmatrix} \nabla g_1(x)^\top \\ \vdots \\ \nabla g_{m_g}(x)^\top \end{pmatrix} (y - x) \\
&\stackrel{g \text{ konvex}, s \geq 0}{\geq} -s^\top \begin{pmatrix} g_1(y) - g_1(x) \\ \vdots \\ g_{m_g}(y) - g_{m_g}(x) \end{pmatrix} \\
&\stackrel{\text{st. Kompl. bed.}}{=} - \sum_{i \in J_s(x)} s_i g_i(y) \\
&\stackrel{g_i(y) \leq 0}{\geq} 0.
\end{aligned}$$

Diese Beobachtung zeigt, dass die Punkte der Effizienzmenge immer auf einer bestimmten Seite der jeweiligen Tangentialebene liegen. Die Hyperfläche \mathcal{E}_U ist somit konvex. □

Korollar 4.3.8 *Bezüglich eines konvexen, multikriteriellen Optimierungsproblems (4.5) sei wie in Satz 4.3.7 eine reguläre und konvexe Hyperfläche \mathcal{E}_U gegeben, für die in jedem Punkt die Abadie CQ sowie die strikten Komplementaritätsbedingungen erfüllt sind. Dann liegt \mathcal{E}_U stets auf der dem Parametervektor ξ zugewandten Seite der Tangentialebenen $T_u \mathcal{E}_U$, $u \in \mathcal{E}_U$, d.h*

$$\xi^\top(y) \geq \xi^\top u \quad \text{also auch} \quad \xi^\top(y - u) \geq 0$$

für alle $y \in \mathcal{E}_U$.

Beweis: Folgt sofort aus dem Beweis von Satz 4.3.7. □

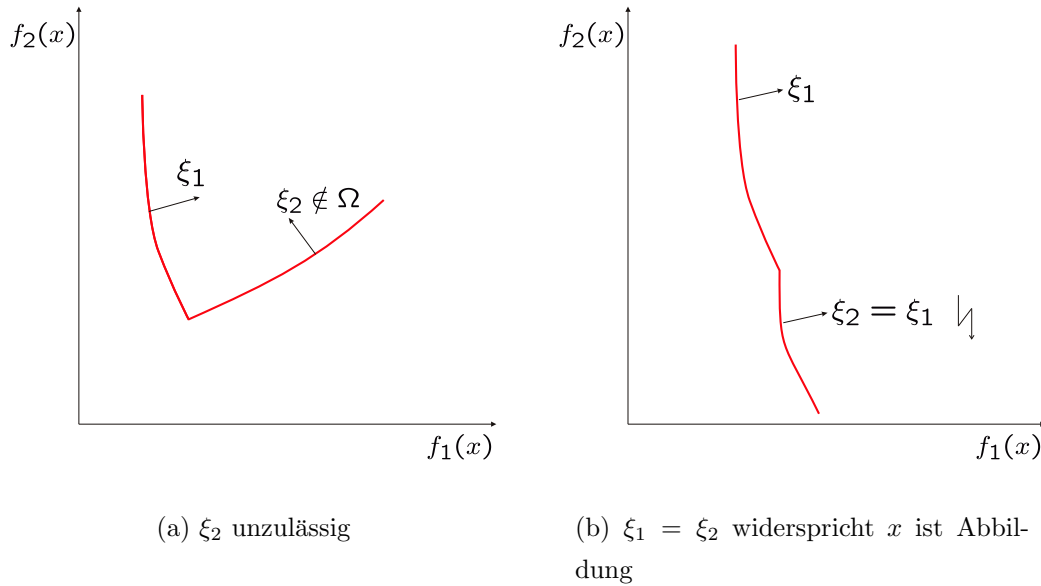
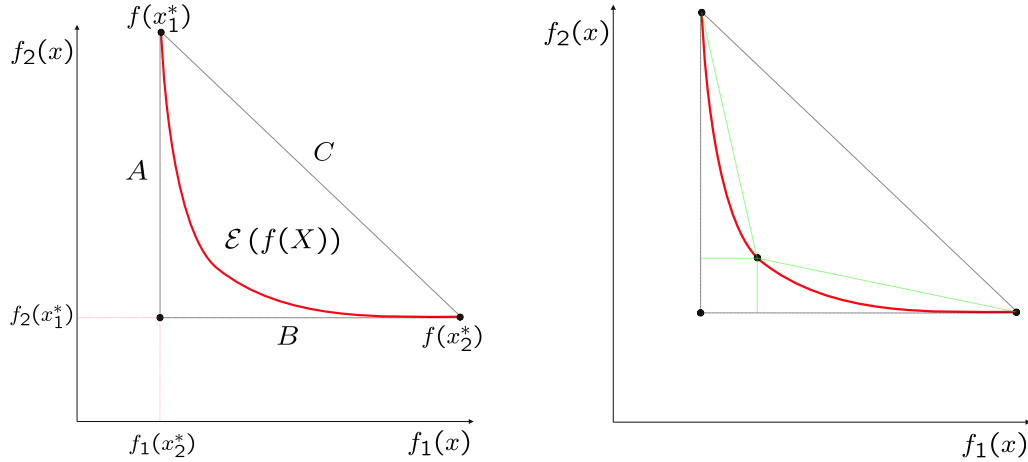


Abbildung 4.4: Beispiele nicht-konvexer Effizienzmengen

Die Konvexität der parametrisierten Fläche \mathcal{E}_U wurde in Satz 4.3.7 nur für den regulären Fall definiert und ist somit auf allgemeine Effizienzmengen, die lediglich stückweise regulär sind, nicht anwendbar. Mit Hilfe von Korollar 4.3.8 und der speziellen Wahl der Parameter aus der Menge Ω , wird jedoch klar, dass die regulären Teilstücke, aus denen die Effizienzmenge besteht, nur in spezieller Weise zusammenhängen können. Einerseits sind die Knicke niemals zu spitz, wie in Abbildung 4.4(a) dargestellt. Dafür wäre ein unzulässiger Parameter $\xi \notin \Omega$ nötig. Andererseits lassen sich keine nichtkonvexen Kurven, wie in Abbildung 4.4(b) konstruieren, da ein Parameter sonst gleichzeitig auf zwei verschiedene Werte abbildet, was dem Charakter einer Abbildung widerspricht. Insgesamt entsteht selbst bei allgemeiner Effizienzmenge wieder ein konvexer Graph, der sogar innerhalb eines durch die Einheitsvektoren des \mathbb{R}^k aufgespannten Simplex liegt (vgl. Abbildung 4.5)).

4.4 Bogenlängen von Effizienzmengen

Die besondere Form der Effizienzmenge ermöglicht eine Verfeinerung der Schrittweitenbestimmung aus Kapitel 3.2. Dazu betrachte man ein konvexes, bikriterielles Optimierungsproblem. Die Effizienzmenge ist durch die Spur der Parameterdarstellung $\varphi_{\mathcal{E}}$ gegeben (siehe 4.23), die mit nur zwei Kriterien einer parametrisierten Kurve $\varphi_{\mathcal{E}} : [0, 1] \rightarrow \mathbb{R}^2$ entspricht.



(a) Abschätzung mit zwei Punkten

(b) Abschätzung mit drei Punkten

Abbildung 4.5: Abschätzung der Effizienzmenge

Hinsichtlich der Konvexität der Effizienzmengen sowie der Beschaffenheit der Parametermenge Ω existiert ein Dreieck (s. Abbildung 4.5), dessen Länge der Hypotenuse C und der beiden Katheten A und B eine untere bzw. obere Abschätzung der Bogenlänge der Kurve $\varphi_{\mathcal{E}}$ zulässt.

Die äusseren Punkte der Effizienzmenge ergeben sich leicht aus den Optimierungsproblemen

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f_1(x) \\ \text{N.B.} \quad & x \in \arg \min f_2(x) \\ & x \in X. \end{aligned} \quad (4.33)$$

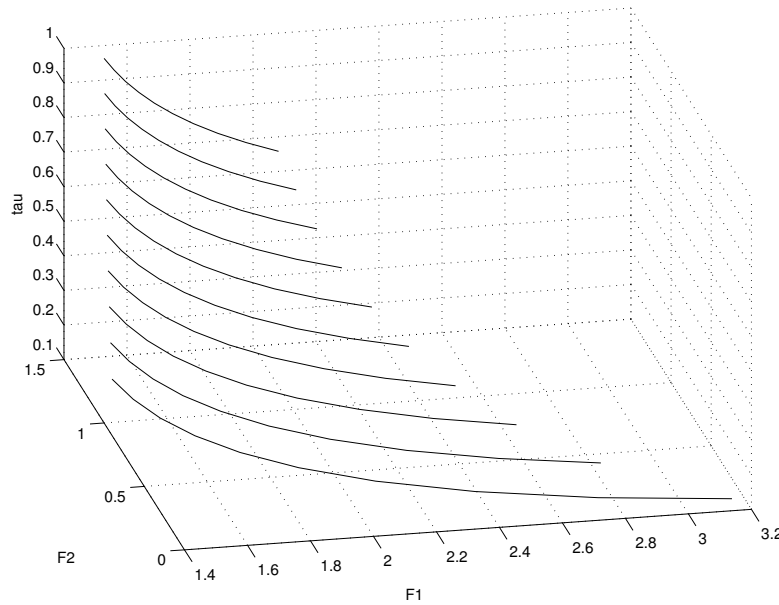
und

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f_2(x) \\ \text{N.B.} \quad & x \in \arg \min f_1(x) \\ & x \in X. \end{aligned} \quad (4.34)$$

Sei x_1^* Lösung des Optimierungsproblems (4.33) und x_2^* Lösung der Aufgabe (4.34).

Insgesamt gilt mit den Längen A , B und C aus Abbildung 4.5 die Abschätzung der Bogenlänge $L_0^1(\varphi)$

$$\begin{aligned} A + B &= \left\| f(x_1^*) - \begin{pmatrix} f_1(x_2^*) \\ f_2(x_1^*) \end{pmatrix} \right\| + \left\| \begin{pmatrix} f_1(x_2^*) \\ f_2(x_1^*) \end{pmatrix} - f(x_2^*) \right\| \\ &\geq L_0^1(\varphi) \\ &\geq \|f(x_1^*) - f(x_2^*)\|. \end{aligned}$$

Abbildung 4.6: τ -Urbildmengen

4.5 Approximationen mit τ -perturbierten Systemen

Zum Schluss dieses Kapitels geht es nochmals um Lösungen τ -perturbierter Systeme. Und zwar stellt sich die Frage, wie gut ein entsprechend τ -zulässiger Punkt $w_\tau = (x_\tau, \lambda_\tau, s_\tau, \pi_\tau)$ des skalarisierten Optimierungsproblems (4.18) zum Parameter $\xi \in \Omega$ zur Approximation der Effizienzmenge geeignet ist.

Abbildung 4.6 zeigt verschiedene τ -Urbildmengen bezüglich variierender Störungen τ aus einem numerischen Beispiel. Auffallend ist, dass die Längen der Urbildmengen mit wachsendem τ kleiner werden. Was die zusätzliche Frage aufwirft, wie sich τ -zulässige Punkte im Bild von f verteilen.

Der Einfachheit halber gehe man von einem konvexen multikriteriellen Optimierungsproblem aus, für dessen Paretopunkte die Abadie CQ erfüllt sei. Die Betrachtung von KKT- und Paretopunkten ist dann äquivalent. Zu einem Parameter $\xi \in \Omega$ existiere die optimale Lösung x^* bzw. deren KKT-Punkt $w^* = (x^*, \lambda^*, s^*, \pi^*)$ des skalarisierten Problems.

Die optimale Lösung x^* minimiert die Zielfunktion $\xi^\top f$ und demnach gilt

$$\xi^\top f(x^*) < \xi^\top f(x_\tau).$$

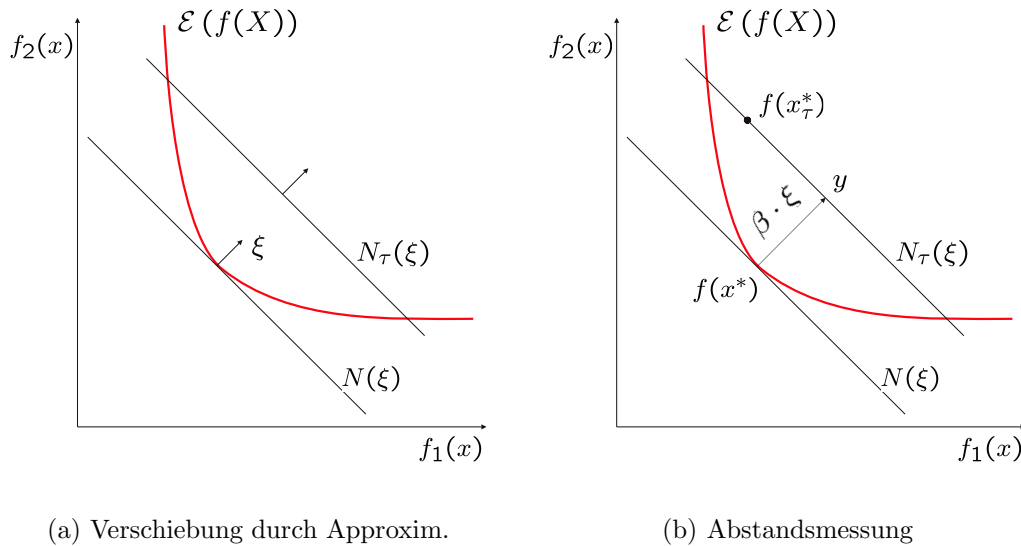


Abbildung 4.7: Approximation der Effizienzmenge

Gleichheit kann nicht herrschen, da x_τ sonst ebenfalls Lösung wäre und damit die Komplementaritätsbedingungen erfüllte, was jedoch im Widerspruch zur τ -Zulässigkeit stände.

Die parallel zueinander liegenden Niveaumengen

$$N(\xi) := \{y \in \mathbb{R}^k \mid \xi^\top y = \xi^\top f(x^*(\xi))\} \quad (4.35)$$

und

$$N_\tau(\xi) := \{y \in \mathbb{R}^k \mid \xi^\top y = \xi^\top f(x_\tau(\xi))\} \quad (4.36)$$

sind also verschieden (s. Abbildung 4.7, (a)). Da ξ senkrecht auf beiden Mengen steht, beträgt der minimale Abstand

$$d := \min_{\substack{x \in N(\xi) \\ y \in N_\tau(\xi)}} \|x - y\|_2 = \beta \cdot \|\xi\|_2,$$

wobei $\beta = \frac{\xi^\top (f(x_\tau) - f(x^*))}{\|\xi\|_2^2}$.

Die Position des Punktes $f(x_\tau)$ kann insofern eingegrenzt werden, als dass er innerhalb der Schnittmenge $N_\tau(\xi) \cap f(X)$ liegen muss, die mit einem Mindestabstand d von der exakten Lösung $f(x^*)$ entfernt ist (s. Abbildung 4.7, (b)).

Im Allgemeinen ist eine genauere Aussage zum Approximationsfehler $\|f(x_\tau) - f(x^*)\|_2$ schwierig, da die Position von $f(x_\tau)$ zusätzlich durch die Bedingung $S\Pi e - \tau$ festgelegt wird.

Für konvex-quadratische Optimierungsprobleme kann eine genauere Aussage getroffen werden, da die Komplementaritätsbedingungen in der Form $s^\top x = e^\top \tau$ die Dualitätslücke wiedergeben. In diesem Fall muss

$$\|f(x_\tau) - f(x^*)\|_2 \leq e^\top \tau$$

gelten.

Daraus ergibt sich ferner

$$\begin{aligned} d &= \beta \cdot \|\xi\|_2 \\ &= \frac{\xi^\top (f(x_\tau) - f(x^*))}{\|\xi\|_2^2} \cdot \|\xi\|_2 \\ &\leq \|f(x_\tau) - f(x^*)\|_2 \\ &\leq e^\top \tau \end{aligned}$$

Definition 4.5.1 (ε -paretominimale Punkte) Gegeben sei die Abbildung

$$f : X \rightarrow \mathbb{R}^k.$$

Ein Punkt $x^\varepsilon \in X$ heißt ε -paretominimal, falls ein $x^* \in \mathcal{P}(f, X)$ existiert, so dass

$$\|f(x^\varepsilon) - f(x^*)\|_2 \leq \varepsilon$$

gilt.

Die Menge aller ε -paretominimalen Punkte heißt ε -Paretomenge $\mathcal{P}_\varepsilon(f, X)$. Die zugehörige Bildmenge $f(\mathcal{P}_\varepsilon(f, X))$ wird als ε -Effizienzmenge $\mathcal{E}_\varepsilon(f(X))$ bezeichnet.

In diesem Sinne liefern alle τ -perturbierten Punkte $(x_\tau, \lambda_\tau, s_\tau, \pi_\tau)$ konvex-quadratischer Optimierungsprobleme die ε -paretominimale Lösung x_τ mit $\varepsilon = e^\top \tau$.

Aber auch für allgemeine konvex-nichtlineare Optimierungsprobleme ist mit dem Abstand d eine d -paretominimale Lösung gefunden, da aufgrund der Konvexität der Effizienzmenge d gleichzeitig den Maximalabstand der Niveaumenge $N_\tau(\xi)$ zur Effizienzmenge angibt. D.h. es existieren Paretopunkte $x^* \in \mathcal{P}(f, X)$ für die $\|f(x_\tau) - f(x^*)\|_2 \leq d$ ausfällt.

Teil II

Software und Tests

Kapitel 5

Software und numerische Ergebnisse

In Bemerkung 2.4.10 wurde bereits erklärt, dass die in Algorithmus 2.4.9 aufgeführte Methode OPT mit geeigneten Optimierungsverfahren besetzt werden kann. Zur Erstellung der in diesem Kapitel aufgeführten numerischen Ergebnisse wurde das (C++)-Softwarepaket QIPP implementiert, welches das Innere-Punkte Verfahren von Mehrotra (vgl. Algorithmus 5.1.3) zur Berechnung optimaler Lösungen konvex-quadratischer Aufgaben nutzt. In Abschnitt 5.1 werden Datenorganisation und implementationstechnische Umsetzungen knapp umrissen. Dann folgen die Beschreibung der Problemstellung, der zugehörigen Algebra und Erläuterungen zum angewandten Mehrotra-Algorithmus.

Das MATLAB-Programm MCQIPP ist die Umsetzung des Algorithmus 2.4.9 mit einigen Erweiterungen. Damit lassen sich die Effizienzmengen multikriterieller, konvex-quadratischer Optimierungsprobleme approximieren. Abschnitt 5.2 skizziert den Ablaufplan des Programms.

In Abschnitt 5.3 werden letztendlich Ergebnisse numerischer Tests mit MCQIPP an realen Daten aus der Steuerung von Kraftwerken eines Forschungsprojekt der VEW Energie AG (Dortmund) [Al97] vorgestellt. Die technischen Eigenschaften des verwendeten Rechners sowie Screenprints der Testauswertungen befinden sich im Anhang.

5.1 QIPP – Ein Innere-Punkte Löser

QIPP steht für **Q**uadratic **I**nterior **P**oint **P**rogram und ist ein objektorientiertes (C/C++)-Programm zur Lösung konvex-quadratischer Optimierungsprobleme. Primal-duale Innere-Punkte Methoden eignen sich dazu besonders, da ihre Iterierten gleichzeitig τ -perturbiert sind und umgekehrt τ -perturbierte Punkte wieder als Iterierte nutzbar sind. Bei anderen Optimierungsmethoden ist dies nicht ohne weiteres gegeben.

Die Umsetzung des Mehrotra-Verfahrens und die lineare Algebra basiert im wesentlichen auf Teilen der (C++)-Bibliothek OOQP von Michael Gertz und Stephen J. Wright ([GerWri01a], [GerWri01b]). Im Gegensatz zu OOQP legt sich QIPP auf eine Standardform des Optimierungsproblems fest. Somit konnten einige Verbesserungen und Vereinfachungen in der Implementierung vorgenommen werden.

Neue Matrizenklassen (`Matrix`, `SpaMat`) wurden mit Teilen der NIST FORTRAN Sparse BLAS Library umgesetzt. Der Speicherbedarf ließ sich im Vergleich zu OOQP um die Hälfte reduzieren, indem die Datenmatrizen sofort in die Form des erweitertes System (s. Abschnitt 5.1.5) abgespeichert wurden. Klassen von OOQP, die einen allgemeinen Ansatz ermöglichten, wurden größtenteils durch imperativen und damit spezialisierteren Code ersetzt. Operationen bezüglich der Linearen Algebra wurden vermehrt mit eingebundenen FORTRAN-Modulen wie BLAS, LAPACK umgesetzt. Die LR-Zerlegung für die Lösung der Gleichungssysteme organisiert das FORTRAN-Modul MA27 aus der HSL-Bibliothek [HSL].

Nach dem Prinzip der **reversen Kommunikation** wurde mit der Klasse `IPSolver` ein unabhängiger Programmkernel geschaffen, der die Schritte des Mehrotra-Verfahrens umsetzt, die unabhängig von der Matrizenverwaltung ablaufen. Darunter fallen u.a. die Bestimmung von Parametern und Schrittweiten sowie die Aktualisierung der Iterierten. Abhängige Methoden wie die Berechnung eines Gleichungssystems, die Speicherung von Matrizen und Matrix-Vektor-Operationen lassen sich auf diese Weise auslagern. Der Vorteil dieser Vorgehensweise liegt in der Flexibilität. Programmmodule zur Berechnung der Gleichungssysteme lassen sich leicht ersetzen; genauso sind Matrizenoperationen oder die Speicherverwaltung austauschbar.

Die reverse Kommunikation findet zwischen den Klassen `IPSolver` und `FQipp` statt. Letztere ist für die Koordination der Methoden zuständig, die von der Matrizenverwaltung abhängen. `FQipp` ruft dabei Methoden von `IPSolver` auf und verwertet die Ergebnisse. Tabelle 5.1 fasst die wichtigsten Klassen und ihre Funktionen zusammen.

Klassenname	Funktion
FArray	Verwaltet Vektoren ([Mo04])
FQipp	Koordiniert vom Datentyp abhängige Methoden, kommuniziert mit IPSolver ([Mo04])
IPSolver	Enthält vom Datentyp unabhängige Methoden ([Mo04])
Ma27Solver	Umsetzung der LR-Zerlegung
PArray	Ermöglicht Referenzen auf FArray ([Mo04])
QpVars	Verwaltet die Iterierten
QpResid	Verwaltet die Residuen
SpaMat	Verwaltet die dünnbesetzten Matrizen

Tabelle 5.1: Klassen von QIPP

5.1.1 Optimierung quadratischer Aufgaben

Die folgenden Ausführungen beschreiben, welche Probleme wie gelöst werden. Dabei geht es um die Umsetzung der Systeme, das Lösen der auftretenden Gleichungssysteme und führt zum verwendeten Mehrotra-Algorithmus. Die Notation wurde so gewählt, dass sie mit der Benennung der Variablen des Programms übereinstimmen.

In QIPP werden Standardprobleme der Art

$$\begin{aligned}
 \min_{x \in \mathbb{R}^n} \quad & f_P(x) := \frac{1}{2}x^\top Qx + c^\top x \\
 \text{N.B.} \quad & Ax = b \\
 & c_l \leq Cx \leq c_u \\
 & x_l \leq x \leq x_u
 \end{aligned} \tag{P}$$

gelöst. Dabei sind mit $C \in \mathbb{R}^{m_z \times n}$, $c_l, c_u \in \mathbb{R}^{m_z}$, sowie x_l und $x_u \in \mathbb{R}^n$ die Ungleichungsnebenbedingungen gegeben, mit $A \in \mathbb{R}^{m_y \times n}$ und $b \in \mathbb{R}^{m_y}$ die Gleichheitsnebenbedingungen. Es wird vorausgesetzt, dass C vollen Spalten- und A vollen Zeilenrang besitzen. Wegen der geforderten Konvexität muss die Matrix $Q \in \mathbb{R}^{n \times n}$ symmetrisch und positiv semidefinit sein.

5.1.2 Das dualisierte Optimierungsproblem

Ein allgemeines konvex-quadratisches Optimierungsproblem

$$\begin{aligned}
 \min_{x \in \mathbb{R}^n} \quad & \frac{1}{2}x^\top Qx + c^\top x \\
 \text{N.B.} \quad & Ax = b \\
 & \tilde{A}x \leq \tilde{b},
 \end{aligned} \tag{5.1}$$

wobei Q wieder als symmetrisch und positiv definit vorausgesetzt wird, korrespondiert mit dem Lagrangedualen Problem

$$\begin{aligned} \max_{x, y, \tilde{y}} \quad & -\frac{1}{2}x^\top Qx - b^\top y - \tilde{b}^\top \tilde{y} \\ \text{N.B.} \quad & Qx + A^\top y + \tilde{A}^\top \tilde{y} = -c \\ & \tilde{y} \geq 0. \end{aligned} \tag{5.2}$$

(Siehe z.B.: [BazShe93], [Mang94]) Hierbei sind y bzw. \tilde{y} die entsprechenden Lagrangemultiplikatoren zu $Ax = b$ bzw. $\tilde{A}x = \tilde{b}$. Fasst man die Ungleichungsnebenbedingungen aus (P) folgendermaßen zusammen

$$\tilde{A} = \begin{pmatrix} -C \\ C \\ -I_n \\ I_n \end{pmatrix}, \quad \tilde{b} = \begin{pmatrix} -c_l \\ c_u \\ -x_l \\ x_u \end{pmatrix},$$

stellt sich das duale Problem zu (P) mit $\tilde{y} = (\lambda, \pi, \gamma, \varphi)$, $\pi, \lambda \in \mathbb{R}^{m_z}$, $\varphi, \gamma \in \mathbb{R}^n$, wie folgt dar:

$$\begin{aligned} \max_{x, y, \tilde{y}} \quad & f_D(x, y, \lambda, \pi, \gamma, \varphi) := -\frac{1}{2}x^\top Qx + b^\top y + c_l^\top \lambda - c_u^\top \pi + x_l^\top \gamma - x_u^\top \varphi \\ \text{N.B.} \quad & Qx - A^\top y - C^\top z + \varphi - \gamma = -c \\ & z = \lambda - \pi \\ & \pi, \lambda, \varphi, \gamma \geq 0 \end{aligned} \tag{D}$$

(vgl. Tabelle 5.2)

Im Gegensatz zu (5.2) wurden einige äquivalente Vorzeichenänderungen vorgenommen. Wechseln die Vorzeichen in der 1. Nebenbedingung vor $A^\top y$ und $C^\top z$, ändert sich entsprechend das Vorzeichen vor $b^\top y$ in der dualen Zielfunktion und die Hilfsvariable z wird zu $-z$.

Mit Hilfe von Schlupfvariablen können die Ungleichungsnebenbedingungen als Gleichungsnebenbedingungen aufgefasst werden. Entsprechend ergibt sich:

$$c_l + t = Cx = c_u - u \tag{5.3}$$

$$x_l + v = x = x_u - w \tag{5.4}$$

$$t, u, v, w \geq 0. \tag{5.5}$$

Für die Dualitätslücke, der Differenz zwischen primaler und dualer Zielfunktion, folgert man schließlich:

Lemma 5.1.1 *Für die Dualitätslücke*

$$\text{gap}(x) := f_P(x) - f_D(x, y, \lambda, \pi, \gamma, \varphi)$$

bzgl. eines Punktes x mit den zugehörigen Lagrangemultiplikatoren y, λ, π, γ und φ gilt die Gleichung

$$\text{gap}(x) = \gamma^\top v + \varphi^\top w + \lambda^\top t + \pi^\top u. \quad (5.6)$$

Beweis: Per Definition und aus der 1. Nebenbedingung des Systems (D) folgt vorerst:

$$\begin{aligned} & \text{gap}(x) \\ = & f_P(x) - f_D(x, y, \lambda, \pi, \gamma, \varphi) \\ = & x^\top Qx + c^\top x - b^\top y + c_u^\top \pi - c_l^\top \lambda + x_u^\top \varphi - x_l^\top \gamma \\ = & x^\top Qx + (-Qx + A^\top y + C^\top z - \varphi + \gamma)^\top x \\ & - b^\top y + c_u^\top \pi - c_l^\top \lambda + x_u^\top \varphi - x_l^\top \gamma \\ = & x^\top Qx - x^\top Qx + y^\top Ax + z^\top Cx - \varphi^\top x + \gamma^\top x \\ & - b^\top y + c_u^\top \pi - c_l^\top \lambda + x_u^\top \varphi - x_l^\top \gamma. \end{aligned}$$

Weiterhin gilt zum einen $y^\top Ax = b^\top y$, da die Nebenbedingung $Ax = b$ des primalen Problems (P) gelten muss und zum anderen - wegen $z = \pi - \lambda$ und der Gleichungen (5.3) -

$$\begin{aligned} z^\top Cx &= \lambda^\top Cx - \pi^\top Cx \\ &= \lambda^\top (c_l - t) - \pi^\top (c_u + u) \end{aligned}$$

Mit den Gleichungen (5.4) lassen sich analog

$$\varphi^\top x = \varphi^\top (x_u - w) \quad \text{und} \quad \gamma^\top x = \gamma^\top (x_l + v)$$

herleiten. Damit heben sich einige Terme weg und es bleibt

$$\text{gap}(x) = \gamma^\top v + \varphi^\top w + \lambda^\top t + \pi^\top u.$$

□

Offensichtlich gilt

$$\tilde{y}^\top (\tilde{A} - \tilde{b}) = \tilde{y}^\top \begin{pmatrix} u \\ t \\ v \\ w \end{pmatrix} = \text{gap}(x).$$

Tabelle 5.2 fasst noch einmal zusammen, wie die einzelnen Nebenbedingungen, Schlupfvariablen und Lagrangemultiplikatoren koordiniert sind.

Nebenbedingung	Schlupfvariable	Lagrangemultiplikatoren
$Ax = b$	-	y
$Cx \geq c_l$	t	λ
$Cx \leq c_u$	u	π
$x \geq x_l$	v	γ
$x \leq x_u$	w	φ

Tabelle 5.2: Zuordnung der Variablen

5.1.3 Das gestörte KKT-System

Das KKT-System beinhaltet und vereint die Informationen des primalen Problems (P) und des dualen (D). Wegen der Konvexität des Optimierungsproblems ist die Erfüllung des KKT-Systems notwendige und hinreichende Bedingung für ein globales Minimum an entsprechender Stelle.

Optimierungsprobleme bilden eine Unterklasse der parametrisierten Optimierung. Insofern findet der Begriff des τ -pertubierten KKT-Systems auch hier Verwendung. Primal-duale Innere-Punkte Methoden ermitteln einen KKT-Punkt, indem sie das KKT-System wie τ -perturbierte Systeme in den Komplementaritätsbedingungen stören und diese Störung τ gegen Null konvergieren lassen. Die τ -perturbierten Systeme liefern Newton-Schritte mit denen neue Iterierte erzeugt werden. Die Störung verhindert die Aktivierung der Ungleichungsnebenbedingungen, weswegen auch von inneren Punkten gesprochen wird. Tatsächlich ermöglicht sie eine numerisch stabile Methode zur Berechnung eines KKT-Punktes.

Die Gleichungssysteme, welche für die Berechnung der Newton-Schritte verwendet werden, sind in Tabelle 5.3 dargestellt. Innerhalb des Doppelrahmens befindet sich die zugrundeliegende Matrix. Diese wurde zeilenweise nach den Blöcken (I)-(XII) und spaltenweise nach den Variablen $\Delta x, \Delta y, \dots, \Delta \varphi$, den Displacement-Vektoren (s. 5.14) unterteilt, wobei sich die Blöcke (IX)-(XII) auf die Komplementaritätsbedingungen beziehen. Die rechte Spalte beinhaltet die Residuen, welche die rechte Seite des Gleichungssystems mit umgekehrten Vorzeichen bilden. Zusätzlich wurde eine weitere Hilfsvariable s eingefügt, so dass die Gleichungen

$$Cx = s, \quad (5.7)$$

$$s = c_l + t, \quad (5.8)$$

$$s = c_u - u \quad (5.9)$$

	Δx	Δy	Δz	Δs	Δt	Δu	Δv	Δw	$\Delta \lambda$	$\Delta \pi$	$\Delta \gamma$	$\Delta \varphi$	RHS
I)	Q	$-A^\top$	$-C^\top$								$-I$	I	$rQ = Qx + c - A^\top y$ $-C^\top z - \gamma + \varphi$
II)	A												$rA = Ax - b$
III)	C			$-I$									$rC = Cx - s$
IV)				I	$-I$								$rt = s - c_t - t$
V)				I		I							$ru = s - c_u + u$
VI)	I						$-I$						$rv = x - x_l - v$
VII)	I							I					$rw = x - x_u + w$
VIII)			I						$-I$	I			$rz = z - \lambda + \pi$
IX)					Λ				T				$r\lambda_\sigma = \Lambda T e - \sigma \mu e$
X)						Π				U			$r\pi_\sigma = \Pi U e - \sigma \mu e$
XI)							Γ				V		$r\gamma_\sigma = \Gamma V e - \sigma \mu e$
XII)								Φ				W	$r\varphi_\sigma = \Phi W e - \sigma \mu e$

Tabelle 5.3: Gleichungssystem zur Berechnung der Displacementvektoren

gelten. Dabei sind (5.8) und (5.9) eine resultierende Variante zu (5.3).

Die Störung bzw. Perturbation der Komplementaritätsbedingungen ist so konzipiert, dass die Iterierten möglichst gleichzeitig und numerisch stabil gegen die Lösung konvergieren. Im Allgemeinen werden die zugehörigen Gleichungssysteme instabil, sobald eine Lösungskomponente zu nah am Restriktionsrand liegt, während andere noch zu weit davon entfernt sind. Die folgende Definition liefert eine Menge deren Elemente bezüglich des o.g. Effekts stabil sind.

Definition 5.1.2 (Zentraler Pfad) Zu einem beliebigen Parameter $\mu > 0$ sei

$$\vec{X}(\mu) = \vec{X}_\mu := (x_\mu, y_\mu, z_\mu, s_\mu, t_\mu, u_\mu, v_\mu, w_\mu, \lambda_\mu, \pi_\mu, \gamma_\mu, \varphi_\mu) \quad (5.10)$$

die Lösung des gestörten KKT-Systems aus Tabelle 5.3, dann bezeichnet die Menge

$$\mathcal{C} = \{\vec{X}(\mu), \mu > 0\} \quad (5.11)$$

den **zentralen Pfad** zu Aufgabe (P).

Existenz und Eindeutigkeit des zentralen Pfades können unter den gegebenen Voraussetzungen gezeigt werden. Auch die Konvergenz

$$\lim_{\mu \rightarrow 0} x_\mu = x^*$$

gegen eine optimale Lösung x^* ist gegeben, sofern sie existiert. Siehe dazu [Meg89], [McL80], [Son89], [BaLa89I], [BaLa89II], [BaLa90].

Damit ergeben sich zwei Berechnungsstrategien, mit denen ein Newtonverfahren angewandt werden kann. Zum einen bezüglich des Residuums (IX - XII)

$$\begin{pmatrix} r\lambda_0 \\ r\pi_0 \\ r\gamma_0 \\ r\varphi_0 \end{pmatrix}, \quad (5.12)$$

was dem wahren Fehler in den Komplementaritätsbedingungen entspricht und einen Schritt in Richtung der Lösung ergibt. Zum anderen das Residuum

$$\begin{pmatrix} r\lambda_1 \\ r\pi_1 \\ r\gamma_1 \\ r\varphi_1 \end{pmatrix}, \quad (5.13)$$

mit dem ein Schritt in Richtung des zentralen Pfades berechnet wird.

Der Steuerungsparameter $\sigma \in [0, 1]$ aus Tabelle 5.3 ermöglicht eine Gewichtung zwischen den beiden Berechnungskriterien.

Die Matrizen Λ , T usw. des letzten eingerahmten Blocks der Tabelle, sind Diagonalmatrizen mit den entsprechenden kleingeschriebenen Vektoren auf der Diagonalen (z.B. $\Pi = \text{diag}(\pi)$). Außerdem ist $e = (1, \dots, 1)^\top$. In diesem Block steht die um $\mu > 0$ perturbierte Dualitätslücke, die in jeder Iteration für verschiedene μ berechnet wird. Aufgrund der Positivität der beteiligten Variablen, kann die um μ gestörte Komplementaritätsbedingung entsprechend aufgeteilt werden. Die Residuen $r\lambda$, $r\pi$, $r\gamma$, $r\varphi$ sind hier zur Veranschaulichung für den Fall eines Prädiktorschritts aufgeführt.

Es werden verschiedene Displacement-Vektoren

$$\Delta \vec{X} := (\Delta x, \Delta y, \Delta z, \Delta s, \Delta t, \Delta u, \Delta v, \Delta w, \Delta \pi, \Delta \lambda, \Delta \gamma, \Delta \varphi) \quad (5.14)$$

berechnet. Sie sind Ergebnisse des Gleichungssystems aus Tabelle 5.3.

5.1.4 Das Mehrotra-Verfahren

Es gibt mehrere primal-duale Innere-Punkte Verfahren - meist für lineare Programme konzipiert, jedoch erweiterbar auf den quadratischen Fall -, die das oben

erläuterte Prinzip der perturbierten Komplementaritätsbedingungen in verschiedenster Weise nutzen. Zu Nennen sind zunächst das Long-Step- und Short-Step-Verfahren sowie Varianten der Prädiktor-Korrektor-Algorithmen (s. [KoMiYo89II], [Wri97], [KoMiYo89I], [MoAd89I], [MiToYe93]). Mit Hilfe gewisser Toleranzumgebungen des zentralen Pfades können sogar Komplexitätsaussagen getroffen werden.

Man spricht von Verfahren 1. Ordnung (nicht zu verwechseln mit der Konvergenzordnung des Verfahrens), weil in jedem Schritt die Ableitungen $D_\mu \vec{X}(\mu)$ der Funktion $\vec{X}_\mu = \vec{X}(\mu)$ benutzt werden, die über das Theorem der impliziten Funktionen A.0.1 ermittelt werden. Dabei entsteht das Gleichungssystem aus Tabelle 5.3 mit $\sigma = 0$. Folglich ergibt sich im Hinblick auf (5.14) die Gleichung

$$\Delta \vec{X}(\mu) = \mu \cdot D_\mu \vec{X}(\mu). \quad (5.15)$$

Der Mehrotra-Algorithmus ist ein Verfahren 2. Ordnung. Entsprechend fließen hier Informationen der zweiten Ableitung in Form von

$$\Delta^2 \vec{X}(\mu) := \frac{\mu^2}{2} \cdot D_\mu^2 \vec{X}(\mu) \quad (5.16)$$

ein. Nochmaliges Anwenden des Theorems über implizite Funktionen zeigt, dass diese ebenfalls aus einem der Tabelle 5.3 ähnlichen Gleichungssystem resultiert. Im Unterschied besteht die rechte Seite, abgesehen von den Komplementaritätsresiduen $r\lambda$, $r\pi$, $r\gamma$, $r\varphi$, jetzt lediglich aus Nullen. Displacement-Vektoren 2. Ordnung ergeben sich mit den Residuen

$$\begin{pmatrix} r\lambda \\ r\pi \\ r\gamma \\ r\varphi \end{pmatrix} = \begin{pmatrix} \Delta\Lambda\Delta T e \\ \Delta\Pi\Delta U e \\ \Delta\Gamma\Delta V e \\ \Delta\Phi\Delta W e \end{pmatrix}. \quad (5.17)$$

Es sei noch erwähnt, dass mit Hilfe dieser Informationen die Taylorapproximation 2. Ordnung von $\vec{X}(0)$ - der gesuchten optimalen Lösung - an der Stelle $\vec{X}(\mu)$ berechnet werden kann. Die Taylorentwicklung ist bekanntlich durch

$$\vec{X}(0) - \vec{X}(\mu) = \Delta \vec{X}(\mu) + \Delta^2 \vec{X}(\mu) + R(\mu) \quad (5.18)$$

gegeben, wobei $R(\mu)$ das Restglied beschreibt.

Es folgt nun der Algorithmus von Mehrotra. Für eine kompaktere Schreibweise gelte die Aufteilung

$$\vec{X} = \left(\vec{X}_f, \vec{X}_{>0}^S, \vec{X}_{>0}^L \right)$$

mit den freien Variablen $\vec{X}_f = (x, y, z, s)$, den strikt positiven Schlupfvariablen $\vec{X}_{>0}^S = (t, u, v, w)$ und den strikt positiven Lagrangemultiplikatoren $\vec{X}_{>0}^L = (\lambda, \pi, \gamma, \varphi)$.

Eine analoge Aufteilung gelte für den Displacement-Vektoren $\Delta\vec{X}$.

Algorithmus 5.1.3 (Prädiktor-Korrektor-Verfahren von Mehrotra)

Input: Zu (P) zulässiger Startpunkt $\vec{X} = (x, y, z, s, t, u, v, w, \lambda, \pi, \gamma, \varphi)$;

Setze: Parameter $\eta \in [2, 4]$; $M := 2n + 2m_C$;

REPEAT

Setze $\mu := \frac{\text{gap}(x)}{M}$;

Berechne den Prädiktor-Schritt $\Delta_P \vec{X}$ *als Lösung des LGS aus Tabelle 5.3 mit*
 $\sigma = 0$;

Berechne

$$\tilde{\alpha} := \max_{\alpha \in (0,1]} \left(\vec{X}_{>0}^S, \vec{X}_{>0}^L \right) + \alpha \left(\Delta_P \vec{X}_{>0}^S, \Delta_P \vec{X}_{>0}^L \right) \geq 0;$$

Setze $\tilde{\mu} = \frac{\text{gap}(x + \tilde{\alpha} \Delta x)}{M}$;

Setze $\sigma = (\tilde{\mu}/\mu)^\eta$;

Berechne den Korrektor-Schritt $\Delta_K \vec{X}$ *als Lösung des LGS aus Tabelle 5.3, mit*
dem Komplementaritätsresiduum (IX)-(XII)

$$\begin{pmatrix} r\lambda_\sigma \\ r\pi_\sigma \\ r\gamma_\sigma \\ r\varphi_\sigma \end{pmatrix} + \begin{pmatrix} \Delta_P \Lambda \Delta_P T e \\ \Delta_P \Pi \Delta_P U e \\ \Delta_P \Gamma \Delta_P V e \\ \Delta_P \Phi \Delta_P W e \end{pmatrix}$$

Berechne

$$\alpha_{\max} := \max_{\alpha \in (0,1]} \left(\vec{X}_{>0}^S, \vec{X}_{>0}^L \right) + \alpha \left(\Delta_K \vec{X}_{>0}^S, \Delta_K \vec{X}_{>0}^L \right) \geq 0;$$

Setze $\alpha = 0.9999 \cdot \alpha_{\max}$;

Setze

$$\vec{X} = \vec{X} + \alpha \cdot \Delta_K \vec{X}; \tag{5.19}$$

UNTIL (is_solution(\vec{X}) OR is_infeasible(\vec{X}))

Output: \vec{X}

Bemerkung 5.1.4

1. Das Update in (5.19) entspricht der Taylorentwicklung (5.18) ohne Restglied bis auf die Tatsache, dass im Residuum der Parameter $\sigma > 0$ Einfluss nimmt.
2. Meist wird $\eta = 3.0$ gesetzt.

Unter den primal-dualen Innere-Punkt Methoden hat sich das Verfahren von Mehrotra wegen seiner Schnelligkeit und Stabilität in der Praxis durchgesetzt. Zudem konnte für Methoden dieser Art superquadratische Konvergenz nachgewiesen werden ([Tra64], [Scha67]). Dabei existieren Varianten in Programmen wie LIPSOL [Zha99], LOQO [Van97], PCx [CzMeWaWr99], OOQP ([GerWri01a], [GerWri01b]), HOPDM ([AlGo99], [GoMa95]).

Im nächsten Abschnitt werden die noch offenen Abbruchkriterien $\text{is_solution}(\vec{X})$ und $\text{is_infeasible}(\vec{X})$ geklärt.

5.1.4.1 Abbruchkriterien

Um zwischen einzelnen Iterationen zu unterscheiden, werden die Größen des Algorithmus 5.1.3 mit dem Index k versehen. Darüber hinaus werden

$$r^k = \begin{pmatrix} rQ^k \\ rA^k \\ rC^k \\ rt^k \\ ru^k \\ rv^k \\ rw^k \\ rz^k \end{pmatrix}$$

und die Dateninstanz

$$d = (Q, A, C, c, b, c_l, c_u, x_l, x_u)$$

mit

$$\|d\|_\infty = \max \{ \|Q\|_\infty, \|A\|_\infty, \|C\|_\infty, \|c\|_\infty, \|b\|_\infty, \|c_l\|_\infty, \|c_u\|_\infty, \|x_l\|_\infty, \|x_u\|_\infty \}$$

definiert. Eine Iterierte \vec{X}^k ist Lösung im numerischen Sinne, falls mit vorgegebenen Toleranzparametern μ_{tol} und r_{tol} die Ungleichungen

$$\mu_k \leq \mu_{\text{tol}}, \quad (5.20)$$

$$\frac{\|r^k\|_\infty}{\|d\|_\infty} \leq r_{\text{tol}} \quad (5.21)$$

gelten.

Die Methode `is_solution` liefert genau dann den booleschen Wert „WAHR“, wenn die obigen Ungleichungen (5.20) und (5.21) erfüllt sind.

Weiterhin gibt es unlösbare Optimierungsprobleme, was für Innere-Punkte Methoden oft schwer erkennbar ist. Die Methode `is_infeasible` soll eben diesen Fall abfangen. Numerisch wird dafür zum einen eine Funktion ϕ_d definiert,

$$\phi_d(r^k, x^k) = \frac{\|r^k\|_\infty + \text{gap}(x^k)}{\|d\|_\infty} \quad (5.22)$$

Das Problem bekommt den Status UNLOESBAR, falls für $k \geq 10$ die Bedingungen

$$\phi_d(r^k, x^k) \geq 10^{-8} \text{ und } \phi_d(r^k, x^k) \geq 10^4 \min_{0 \leq i \leq k} \phi_d(r^i, x^i) \quad (5.23)$$

erfüllt sind.

Der Mehrotra-Algorithmus bricht zum einen mit dem Status UNBEKANNT ab, wenn für ein $k \geq 30$ die Ungleichung

$$\min_{0 \leq i \leq k} \phi_d(r^k, x^k) \geq \frac{1}{2} \min_{1 \leq i \leq k-30} \phi_d(r^i, x^i) \quad (5.24)$$

gilt. Zum anderen auch dann, wenn die Aussage

$$\frac{\|r^k\|_\infty}{\|d\|_\infty} > r_{\text{tol}} \quad \text{UND} \quad \frac{\|r^k\|_\infty}{\mu_k} \geq 10^8 \frac{\|r^0\|_\infty}{\mu_0} \quad (5.25)$$

wahr ist.

Zuletzt bricht der Algorithmus auch dann ab, wenn eine maximale Anzahl von Iterationen überschritten wird.

5.1.5 Das erweiterte System

Es bietet sich die Möglichkeit, das System aus Tabelle 5.3 kompakter darzustellen und damit effektiver lösen zu können, indem gewisse Gleichungen zusammengefasst

werden. Da sich das vorliegende System recht kompliziert gestaltet und gewisse Umformungen nicht immer sofort auf der Hand liegen, soll in diesem Abschnitt das erweiterte System für dieses spezielle Problem hergeleitet werden.

Wie für (5.14) bereits beschrieben, soll ein Displacement-Vektor $\Delta\vec{X}$ berechnet werden. Dieser ist das Ergebnis des Gleichungssystems beschrieben in Tabelle 5.3.

Aus Block (XI) ergibt sich $\Gamma\Delta v + V\Delta\gamma = -r\gamma$ und $\Delta v = rv + \Delta x$ aus Block (VI). Damit lässt sich $\Delta\gamma$ berechnen, es gilt:

$$\begin{aligned}\Delta\gamma &= -V^{-1}r\gamma - V^{-1}\Gamma\Delta v \\ &= -V^{-1}r\gamma - V^{-1}\Gamma(rv + \Delta x) \\ &= -V^{-1}r\gamma - V^{-1}\Gamma rv - V^{-1}\Gamma\Delta x.\end{aligned}\tag{5.26}$$

Analog entwickelt man

$$\Delta\varphi = -W^{-1}r\varphi + W^{-1}\Phi rrw + W^{-1}\Phi\Delta x.\tag{5.27}$$

aus Block (XII) und (VII),

$$\Delta\lambda = -T^{-1}r\lambda - T^{-1}\Lambda rt - T^{-1}\Lambda\Delta s.\tag{5.28}$$

aus Block (IX) und (IV) und

$$\Delta\pi = -U^{-1}r\pi + U^{-1}\Pi ru + U^{-1}\Pi\Delta s.\tag{5.29}$$

aus Block (X) und (V).

Die Folgerungen (5.28) und (5.29) ergeben mit (VII)

$$\begin{aligned}\Delta z &= -rz - \Delta\pi + \Delta\lambda \\ &= -rz + U^{-1}r\pi - U^{-1}\Pi ru - U^{-1}\Pi\Delta s \\ &\quad -T^{-1}r\lambda - T^{-1}\Lambda\Delta s - T^{-1}\Lambda rt\end{aligned}\tag{5.30}$$

\Leftrightarrow

$$\begin{aligned}-(T^{-1}\Lambda + U^{-1}\Pi)\Delta s &= \Delta z + rz - U^{-1}r\pi + U^{-1}\Pi ru \\ &\quad +T^{-1}r\lambda + T^{-1}\Lambda rt\end{aligned}\tag{5.31}$$

\Leftrightarrow

$$\begin{aligned}-\Delta s &= (T^{-1}\Lambda + U^{-1}\Pi)^{-1} \\ &\quad (\Delta z + rz - U^{-1}(r\pi - \Pi ru) + T^{-1}(r\lambda + \Lambda rt))\end{aligned}\tag{5.32}$$

	Δx	Δy	Δz	RHS
XIII)	$Q + D$	$-A^\top$	$-C^\top$	$rQ + V^{-1}r\gamma + V^{-1}\Gamma rv - W^{-1}r\varphi + W^{-1}\Phi rw$
XIV)	A			rA
XV)	C		Ω	$rC + \Omega(rz + T^{-1}r\lambda + T^{-1}\Lambda rt - U^{-1}r\pi + U^{-1}\Pi ru)$

Tabelle 5.4: Erweitertes System im Mehrotra-Verfahren

Seien nun

$$D = W^{-1}\Phi + V^{-1}\Gamma, \quad \Omega = (T^{-1}\Lambda + U^{-1}\Pi)^{-1} \quad (5.33)$$

definiert, dann ergibt sich das erweiterte System aus Tabelle 5.4

Bemerkung 5.1.5 *Aufgrund des reduzierten Gleichungssystem vermindert sich der Rechenaufwand in Algorithmus 5.1.3 erheblich, sofern die Prädiktor- und Korrektorschritte mit Hilfe des Gleichungssystems aus Tabelle 5.3 berechnet werden.*

5.2 MCQIPP – Ein multikriterieller Optimierer

Zur Erinnerung: In Kapitel 2 wurde zur Approximation von τ -Urbildmengen Algorithmus 2.4.9 für allgemein nichtlineare, konvexe Probleme vorgestellt. Kapitel 4 zeigte, dass Effizienzmengen unter gewissen Voraussetzungen auch Lösungsmengen parametrisierter Optimierungsprobleme sind, die wiederum den τ -Urbildmengen mit $\tau = 0$ entsprechen.

Angesichts auftretender singulärer und schlecht konditionierter Matrizen für $\tau = 0$ ist es im Allgemeinen notwendig τ -Urbildmengen mit $\tau > 0$ zu approximieren. Die gewonnenen Punkte der Urbild-Approximation können als Startpunkte für Lösungsverfahren benutzt werden. Da diese Startpunkte, je nach Wahl von τ , sehr nahe an der Lösung liegen, lässt sich die Lösung innerhalb weniger Iterationen gewinnen.

Das MATLAB-Programm MCQIPP (MultiCriteria QIPP) ist die Umsetzung des Algorithmus 2.4.9 mit einigen Modifikationen. Die jetzige Version ist für den bikriteriellen Fall konzipiert. Eine Erweiterung auf mehr Kriterien ist möglich, jedoch gestaltet sich dann die Parameterwahl komplizierter. Einige, diesbezügliche Anregungen sind am Ende von Kaptitel 6 zu finden.

Zur Berechnung müssen die bikriteriellen Optimierungsprobleme in die Standardform

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f_P(x) := \begin{pmatrix} \frac{1}{2}x^\top Q_1 x + c_1^\top x \\ \frac{1}{2}x^\top Q_2 x + c_2^\top x \end{pmatrix} \\ \text{N.B.} \quad & Ax = b \\ & c_l \leq Cx \leq c_u \\ & x_l \leq x \leq x_u. \end{aligned} \quad (MCP)$$

gebracht werden. Die Matrizen A und C bzw. die Vektoren c, b, c_l, c_u, x_l, x_u sind, wie in (P) gewählt. Die Matrizen $Q_1, Q_2 \in \mathbb{R}^{n \times n}$ sind symmetrisch und positiv semidefinit.

Algorithmus 2.4.9 wurde

- 1) mit der tangentialen Schrittweitenkontrolle nach Kapitel 3.2 und
- 2) einer Routine (`get_interval2`) zur Erkennung von echten Knicken in der Effizienzmenge nach Kapitel 3.1.2 erweitert.

Abbildung 5.1 zeigt den Programmablaufplan zu MCQIPP. Erläuterungen zu den dort angegebenen Methoden und Variablen sind in Tabelle 5.5 aufgeführt.

Erläuterung: Zu Anfang wird der Parameter `xi` auf Null gesetzt sowie eine Schätzung der Bogenlänge der Effizienzmenge mit `estimate_curvelength` wie in Abschnitt 4.4 vorgenommen. Die geschätzte Länge wird mit dem Eingabeparameter `step_accuracy` dividiert, womit sich eine Schrittlänge `step_length` bezüglich der Kurve ergibt, die - je nach Qualität der Abschätzung - zu nahezu äquidistanten Abständen führen kann. Die zugehörigen Parameterwechsel, welche zu dieser Schrittlänge gehören, werden mit Hilfe der tangentialen Schrittweitenbestimmung unter Hinzunahme von Ableitungsinformationen geschätzt.

Als nächstes startet eine While-Schleife, in welcher der Parameter `xi` stetig vergrößert wird, bis er den Wert Eins annimmt. Das Programm bricht außerdem ab, wenn QIPP nicht erfolgreich war oder die maximale Anzahl an Durchläufen `max_counter` erreicht ist.

Im ersten Durchlauf zu `xi = 0` beginnt QIPP mit einer konventionellen Startpunktbestimmung. Bricht QIPP erfolgreich mit einer Lösung `iter` ab, liefert es gleichzeitig die `w_tol`-letzte Iterierte als Kandidat für die Bestimmung eines Warmstartpunkts. Ein solcher Neustart wird außerdem nach zu großen Residuenbildungen, der Erkennung von Knicken und zu kleinen Parameterwechseln im Rahmen von Algorithmus

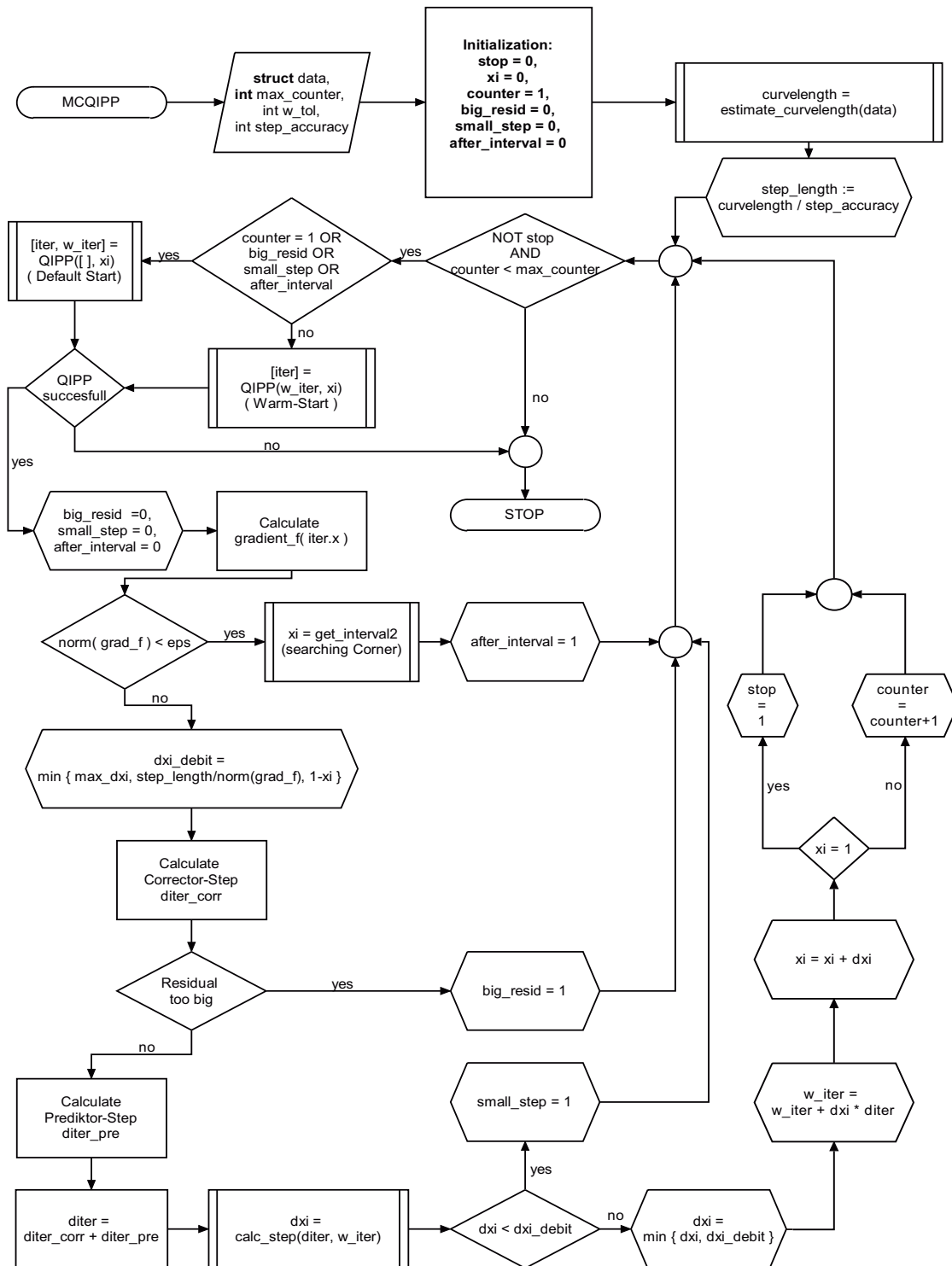


Abbildung 5.1: Programmablaufplan: MCQIPP

<code>after_interval</code>	Booleanwert: TRUE, wenn Knick entdeckt wurde
<code>big_resid</code>	Booleanwert: TRUE, wenn Residuum zu groß
<code>counter</code>	Zähler für die Hauptschleifen
<code>curvelength</code>	Geschätzte Bogenlänge der Eff.menge
<code>calc_step()</code>	Berechnet mögliche Perturbation dx_i
<code>data</code>	Struktur aus Eingabedaten Q, A, \dots
<code>diter_corr</code>	Korrektorschritt
<code>diter_pre</code>	Prädiktorschritt
<code>dx_i</code>	Mögliche Parameterperturbation von ξ
<code>dx_i_debit</code>	Sollperturbation von ξ
<code>estimate_curvelength</code>	Schätzt die Bogenlänge der Eff.menge
<code>get_interval2()</code>	Schätzt das Parameterintervall eines Knickes
<code>grad_f</code>	Ableitung d. Kurve, welche Eff.menge parametrisiert
<code>gradient_f()</code>	Berechnet Ableitung d. Kurve, welche Eff.menge parametrisiert
<code>iter</code>	Iterierte
<code>max_counter</code>	Max. Anzahl an Hauptschleifen
<code>max_dx_i</code>	Max. Perturbation von ξ
<code>small_step</code>	Booleanwert: TRUE, wenn $dx_i < dx_i_debit$
<code>step_accuracy</code>	Gewünschte Anzahl der zu berechnenden Eff.punkte
<code>step_length</code>	Geschätzte Bogenteillänge f. äquidistante Abstände auf d. Eff.menge
<code>stop</code>	Booleanwert: TRUE, wenn Eff.menge berechnet
<code>w_iter</code>	Warmstart-Iterierte
<code>w_tol</code>	Die <code>w_tol</code> -letzte Iterierte ist Warmstartpunkt <code>w_iter</code>
<code>xi</code>	Parameter ξ

Tabelle 5.5: Erläuterung zu Variablen und Methoden in MCQIPP

2.4.9(a), (b) aufgerufen. In diesen Fällen werden die booleschen Variablen `big_resid`, `after_interval` bzw. `small_step` auf `TRUE` bzw. `1` gesetzt. Sofern keiner der genannten Fälle eintritt, startet QIPP mit einem ermittelten Warmstartpunkt `w_iter`.

Nach der Bestimmung der Lösung `iter` wird die Ableitungsinformation `grad_f` berechnet. Dies ist die Ableitung der parametrisierten Kurve, welche die Effizienzmenge darstellt, an der Stelle der Lösung `iter`. Ist die Ableitung Null, so befindet sich an dieser Stelle ein Knick in der Effizienzmenge und die Methode `get_interval2` sucht einen neuen Parameter `xi` außerhalb des Knickes.

Ist die o.g. Ableitung größer als die Maschinengenauigkeit `eps`, kann ein Sollparameterwechsel `dxi_debit` geschätzt werden, der möglichst zur Schrittlänge `step_length` führt. Zur Bestimmung des Warmstartpunkts wird wie in Algorithmus 2.4.9 der Korrektorschritt berechnet und damit das Residuum überprüft, ob es hinreichend klein ist. Sofern es nicht zu groß ist, kann der entsprechende Prädiktorschritt und damit der Parameterwechsel `dxi` ermittelt werden. Hier kommt es wieder darauf an, dass `dxi` mindestens dem Sollparameterwechsel `dxi_debit` entspricht. Kann dies bejaht werden, wird der neue Warmstartpunkt `w_iter` konstruiert. Ansonsten wird der Booleanwert `small_step` `TRUE` gesetzt.

Die hier gewählte tangentielle Schrittweitenbestimmung ist nicht unproblematisch. So ist die Funktionalität der Schrittweitembestimmung auf Geraden optimal. Ein Geradenstück in der Effizienzmenge zieht die Zuordnung mehrerer Lösungen zu einem Parameter nach sich. Insofern besteht der zugehörige Abbildungswert aus einer ganzen Menge von Lösungen, was z.B. bei singulären KKT-Systemen auftritt. Algorithmus 2.4.9 arbeitet demnach genau dort am besten, wo die tangentielle Schrittweitembestimmung am schlechtesten ist und umgekehrt. Eine Umstellung auf die zirkuläre Schrittweitemsteuerung ist somit ratsam. Diese hätten sogar den Vorteil, Geradenstücke anhand einer Krümmung nahe Null zu erkennen; entsprechend könnte reagiert werden.

5.3 Numerisches Beispiel: Steuerung von Kraftwerken

MCQIPP wurde an realen Testdaten aus der Kraftwerksteuerung überprüft. Die Voreinstellungen für MCQIPP waren `step_accuracy = 100`, `w_tol = 3` und

$\max_{dx_i} = \frac{1}{30}$. Komponenten des Gradienten in einer Lösung der Effizienzkurve durften nicht kleiner als die Wurzel der MATLAB-Maschinengenauigkeit `eps` sein, damit eine reguläre Stelle erkannt wurde.

Es wurden zwei weitere Algorithmen überprüft.

Im Vergleich zu MCQIPP wurde die naive Strategie (NAIV) verwendet, die ebenfalls skalarisierte Ersatzprobleme löste, allerdings nur unter Verwendung herkömmlicher Startpunkte. Für die Bestimmung der Parameterwechsel wurde gleichfalls die tangentielle Schrittweitensteuerung gewählt. Dazu wurde der Verlauf von MCQIPP benutzt, ohne die für Warmstartpunktberechnungen nötigen Methoden aufzurufen. Unter MATLAB liefen insofern die Organisation der Parameterwechsel und der Aufruf von QIPP. Aufgrund der Performance-Unterschiede zwischen der Skriptsprache MATLAB und C++ verhilft ein Vergleich über die Rechenzeit zu keiner befriedigenden Schlussfolgerung, da der Hauptaufwand bei NAIV in QIPP, also innerhalb eines (C++)-Codes betrieben wurde. Deshalb fand außerdem ein Vergleich über die Anzahl der QIPP-Iterationen statt.

MCQIPP wurde zusätzlich mit dem MATLAB-Code NBI von Das und Dennis [DasDen96] verglichen. Hier gilt umso mehr, dass ein Zeitvergleich keine Schlussfolgerung zulässt. Dennoch wurde zwecks Einordnung zu bekannter Software ein Benchmark anhand der Testdaten erstellt. Aufgrund sehr langer Laufzeiten wurden lediglich 15 Effizienzpunkte je Datum berechnet. Eine umgekehrte Anpassung von MCQIPP mit `step_accuracy = 15` war keine Alternative, da Warmstart-Strategien unter den folglich großen Parameterwechseln versagen.

Zum Testproblem: Im Bundesland Nordrhein-Westfalen stehen 14 Kraftwerke für circa 18 Millionen Stromkonsumenten zur Verfügung. Die Stromnachfrage variiert ständig. Abhängig von Tageszeit, Wochentag, Jahreszeit usw. lassen sich jedoch sehr gute Bedarfsprognosen erstellen.

Der Betrachtungszeitraum ist ein Zeitfenster von 24 Stunden = 1440 Minuten; zur Verfügung standen 1435 Minuten. Vor jeder neuen Minute muss eine Entscheidung getroffen werden, wie der nachgefragte Strom von den zur Verfügung stehenden Kraftwerken geliefert werden kann.

Dabei müssen

1. die prognostizierte Nachfrage erfüllt und

2. die technischen Bedingungen berücksichtigt werden.

Die technischen Bedingungen sind nicht immer erfüllbar, wenn z.B. der Bedarf schneller steigt als die Kraftwerke regulierbar sind. Es ist sinnvoll mehrere Prognosezeitpunkte gleichzeitig zu betrachten um diesen starken Veränderungen vorzubeugen. Für jede weitere berücksichtigte Minute erhöht sich jedoch auch die Anzahl der Variablen um 14, der Anzahl der Kraftwerke. Es wurden jeweils 4 Minuten im Voraus betrachtet.

Gleichzeitig sollen

1. die Betriebskosten und
2. der Verschleiß der Kraftwerke

minimal gehalten werden.

Zusammen ergibt sich ein konvex-quadratisches, bikriterielles Optimierungsproblem mit linearen Nebenbedingungen. Es entstehen 56 Entscheidungsvariablen, 32 Gleichheits- und 168 Ungleichheitsnebenbedingungen. Sämtliche dazugehörigen Matrizen sind dünn besetzt.

Jede Minute muss somit eine multikriterielle Aufgabe der obigen Größenordnung gelöst werden und aufgrund der Effizienzmenge eine Entscheidung gefällt werden. Demnach sollten die einzelnen Probleme nicht länger als 60 Sekunden in der Berechnung benötigen. Ansonsten muss vorher abgebrochen werden und lediglich die bereits erhaltenen Resultate können zu einer Entscheidungsfindung herangezogen werden.

Die Ergebnisse: Die Ergebnisse der Tests sind in Tabelle 5.6 aufgeführt und ferner in Abbildung 5.2 und 5.3 dargestellt. Die Abbildungen 5.4 - 5.7 zeigen ausgewählte Beispiele berechneter Effizienzmengen, den zugehörigen Parameterwechseln, QIPP-Iterationen und der Verteilung der Häufigkeit von Warmstartpunkten. Die kompletten Ausgaben zur Auswertung der Daten befinden sich in Anhang B.

Es wurde die benötigte Rechenzeit, die Anzahl der berechneten effizienten Punkte sowie die von QIPP benötigten Iterationen gemessen. In der zweiten Spalte sind die Durchschnittswerte pro Minutenfenster eingetragen, in der dritten die gesamte

MCQIPP	\emptyset pro Zeitp.	Gesamt	Min	Max
Zeit	21.93 Sek.	7.99 Stunden	6.81 Sek.	33.36 Sek.
# eff. Punkte	98.72	129 527	29	136
# QIPP-Iter.	607.70	797 308	192	1086
Naive Strategie	\emptyset pro Zeitp.	Gesamt	Min	Max
Zeit	13.96 Sek.	5.1 Stunden	4.51 Sek.	22.54 Sek.
# eff. Punkte	92.32	121 500	25	127
# QIPP-Iter.	1231	1 620 717	297	1949
NBI	\emptyset pro Zeitp.	Gesamt	Min	Max
Zeit	45.31 Sek.	16.56 Stunden	12.92 Sek.	166.71 Sek.
# eff. Punkte	15	$15 \times 1316 = 19\,740$	15	15
QIPP-Iter.	-	-	-	-

Tabelle 5.6: Numerische Ergebnisse der Kraftwerkdaten

Summe für alle 1435 Minuten. Die vierte und fünfte Spalte gibt die Minimal- bzw. Maximalwerte an.

Insgesamt waren mit MCQIPP, NAIV und NBI 119 Probleme unlösbar. Für MCQIPP führten zusätzlich 4 Sonderfälle zu Schwierigkeiten. Demnach waren 1316 bzw. 1312 von 1435 zur Verfügung stehenden Daten lösbar.

Die Rechenzeiten sind bei NAIV wie erwartet aufgrund der Performance-Unterschiede am besten, bei NBI am schlechtesten. Aufgrund der Daten benötigt NBI durchschnittlich 3.02 Sekunden für die Berechnung eines effizienten Punkts, MCQIPP 0.22 Sekunden und NAIV 0.15 Sekunden. Damit ist NBI ca. 20-mal langsamer als NAIV und 13.8-mal langsamer als MCQIPP. Im Vergleich zu NAIV ist MCQIPP 1.47-mal langsamer.

Betrachtet man die von QIPP benötigten Iterationen in MCQIPP, verdoppeln sich diese bei NAIV. Der wirkliche Aufwand ist somit mehr als doppelt so groß. Unter zusätzlicher Berücksichtigung des Aufwands bezüglich der Ermittlung eines Warmstartpunkts kommen je Warmstartpunkt ein Prädiktor- und ein Korrektorschritt hinzu, bei zu kleinen Schrittweiten lediglich ein Korrektorschritt. Um die Rechnung einfacher zu gestalten, verzichte man auf die Tatsache, dass Korrektorschritte nach jedem Neustart entfallen. Der Aufwand beider Schritte entspricht recht genau dem einer Iteration in QIPP. Für ein genaues Bild müssten demnach für jeden Warmstartpunkt eine weitere QIPP-Iteration hinzu gezählt werden. Die Verteilung in Abbildung 5.2 zeigt, dass ca. 66 % aller effizienten Punkte mit Hilfe von Warmstartpunk-

ten ermittelt wurden. Hinzu kommen 24.71 %, die durch zu kleine Parameterwechsel entstanden und trotzdem ein Warmstartpunkt berechnet wurde. Insgesamt sind ca. 90 % zu beachten. Bei einer Gesamtheit von 129 527 effizienten Punkten wurden demnach ca. 116 600 Warmstartpunktberechnungen benötigt. Entsprechend addierte sich der Aufwand von 797 308 QIPP-Iterationen auf 913 908. Das ergäbe einen neuen Durchschnitt von 696.58 QIPP-Iterationen je Zeitpunkt, was immer knapp mehr der Hälfte des Aufwands in NAIV entspräche.

Ein Blick auf die durchschnittliche Anzahl berechneter Effizienzpunkte zeigt, dass die tangentiale Schrittweite befriedigende Ergebnisse liefert. Tatsächlich wurden in den meisten Fällen um die 100 Punkte berechnet. Die Abstände auf der Effizienzkurve sind sehr gut verteilt. In den Abbildungen 5.4 - 5.7 sollte beachtet werden, dass die Skalierungen der Achsen für die Effizienzmenge nicht einheitlich ist.

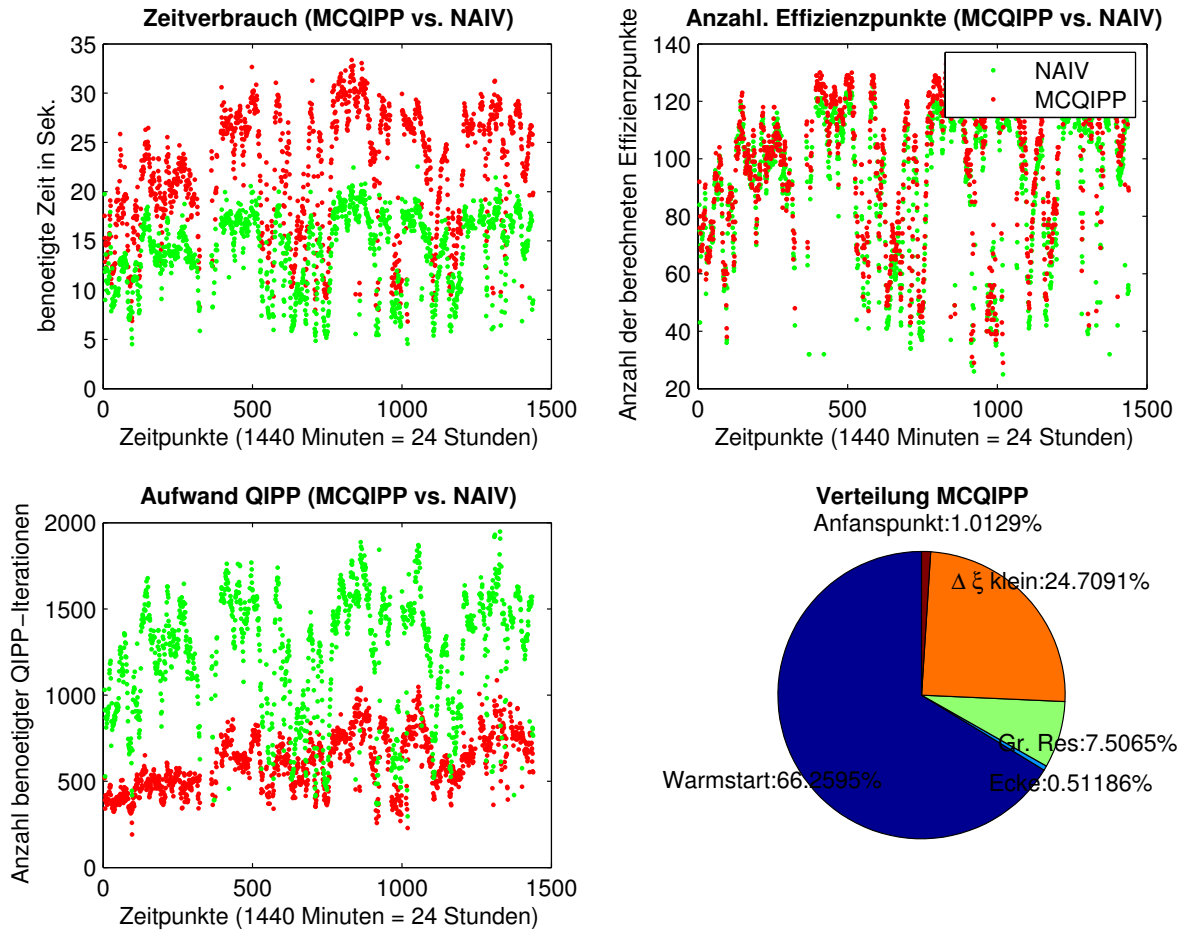


Abbildung 5.2: Benchmark: MCQIPP vs NAIV

Die roten Punkte repräsentieren MCQIPP, die grünen NAIV (vgl. Legende im Graph oben rechts). Der erste Graph (oben links) zeigt die benötigten Zeiten zur Berechnung der Effizienzmengen in jedem Zeitpunkt. Im zweiten Graph (oben rechts) werden die Anzahl der berechneten Effizienzpunkte je Zeitpunkt wiedergegeben. Der dritte Graph (unten links) stellt die benötigten QIPP-Iterationen je Zeitpunkt dar. Das Diagramm (unten rechts) zeigt die Verteilung der in MCQIPP aufgetretenen Fälle bezüglich verschiedener begründeter Neustarts und Warmstarts.

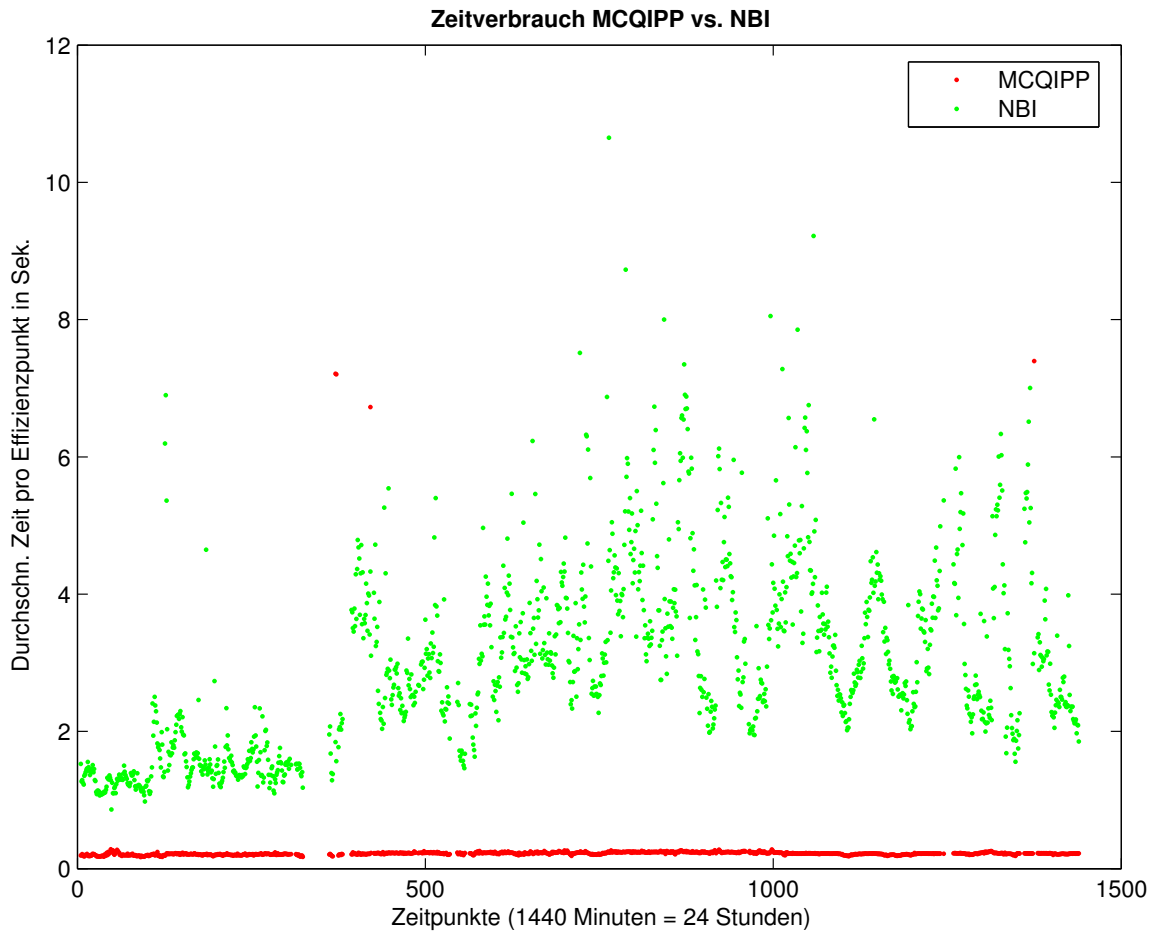


Abbildung 5.3: Benchmark: MCQIPP vs. NBI

Die roten Punkte repräsentieren MCQIPP, die grünen NBI (vgl. Legende). Es wurden die durchschnittlich benötigten Rechenzeiten für einen Effizienzpunkt zu den verschiedenen Zeitpunkten eingetragen.

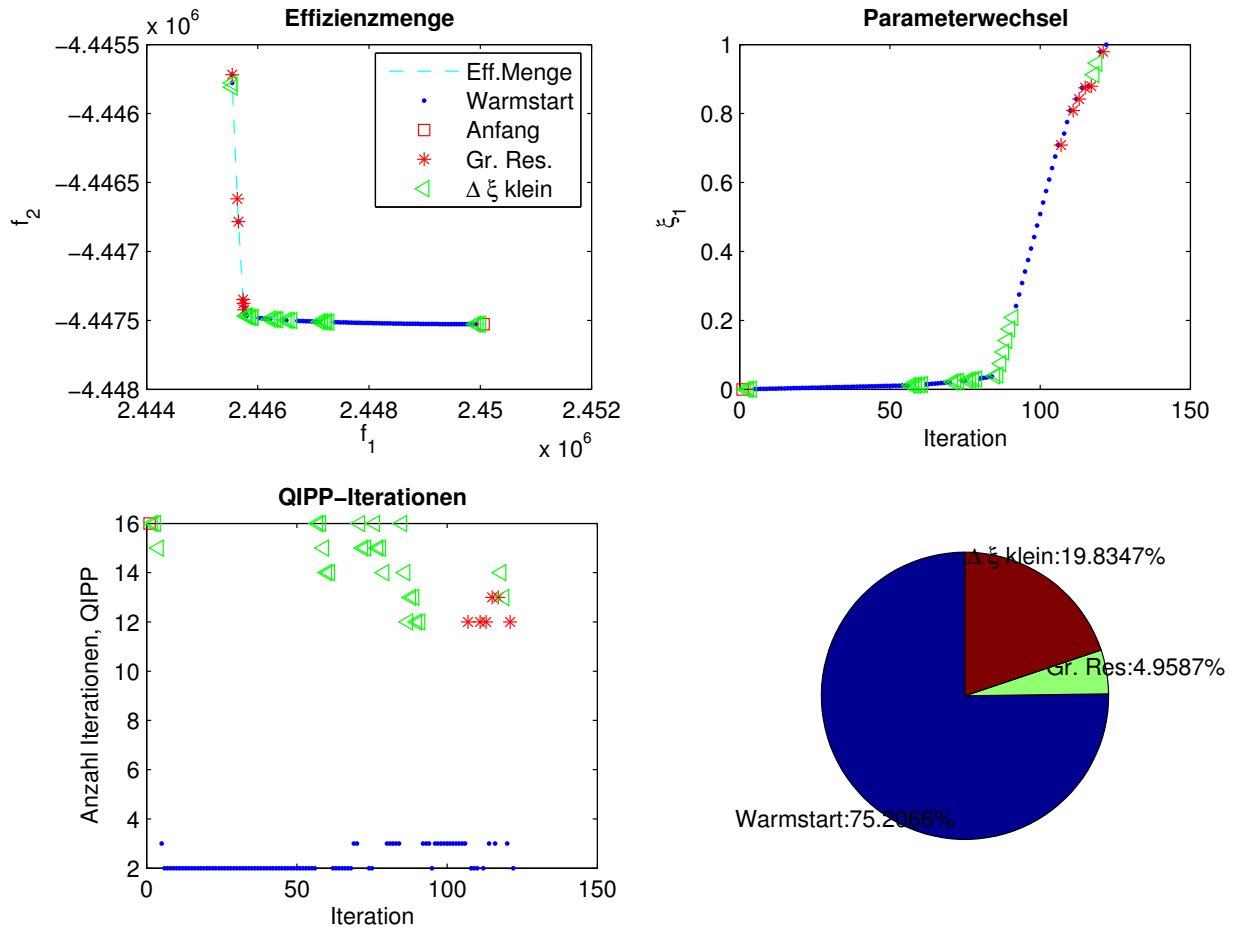


Abbildung 5.4: Zeitpunkt: Minute 580 , Zeit: 27.75 Sek., 122 Eff.pkte

Bei der Berechnung der Effizienzmengen wurde unterschieden, ob QIPP innerhalb der Iterationen mit Warmstartpunkten oder konventionellen Startpunkten aufgerufen wurde. Konventionelle Startpunkte waren immer dann nötig, wenn Neustarts durchgeführt wurden, nämlich zur Berechnung des ersten Effizienzpunkts, bei zu großen Residuen und zu kleinen Parameterperturbationen. Innerhalb der Graphen wurden diese Fälle jeweils durch ein Quadrat, einem Stern bzw. einem Dreieck kenntlich gemacht; Warmstartpunkte wurden durch Punkte dargestellt (vgl. Legende im Graph oben links).

Das Diagramm (unten rechts) gibt die Häufigkeit (in Prozent) der angeführten Fälle bei der Berechnung der Effizienzmenge zum gegebenen Zeitpunkt (hier: 580te Minute) an. Der erste Graph (oben links) stellt die approximierte Effizienzmenge bezüglich der Funktionswerte f_1 und f_2 dar. Der zweite Graph (oben rechts) gibt den Fortschritt der Parameterperturbationen in der jeweiligen MCQIPP-Iteration an. Der dritte Graph (unten links) verzeichnet die Anzahl der QIPP-Iterationen, die pro MCQIPP-Iteration benötigt wurden.

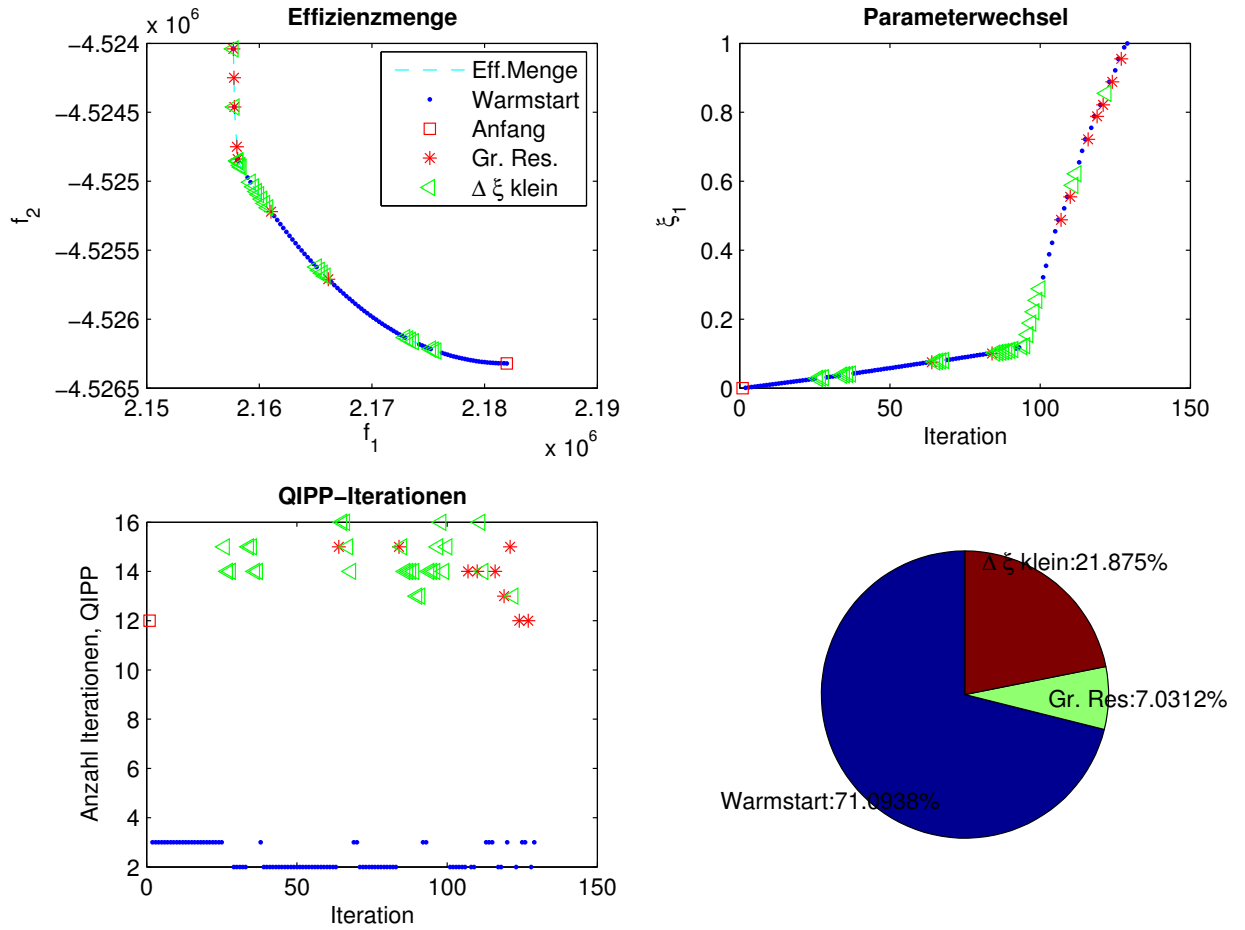


Abbildung 5.5: Zeitpunkt: Minute 396 , Zeit: 30.58 Sek., 129 Eff.pkte
(Beschreibung wie in Abbildung 5.4)

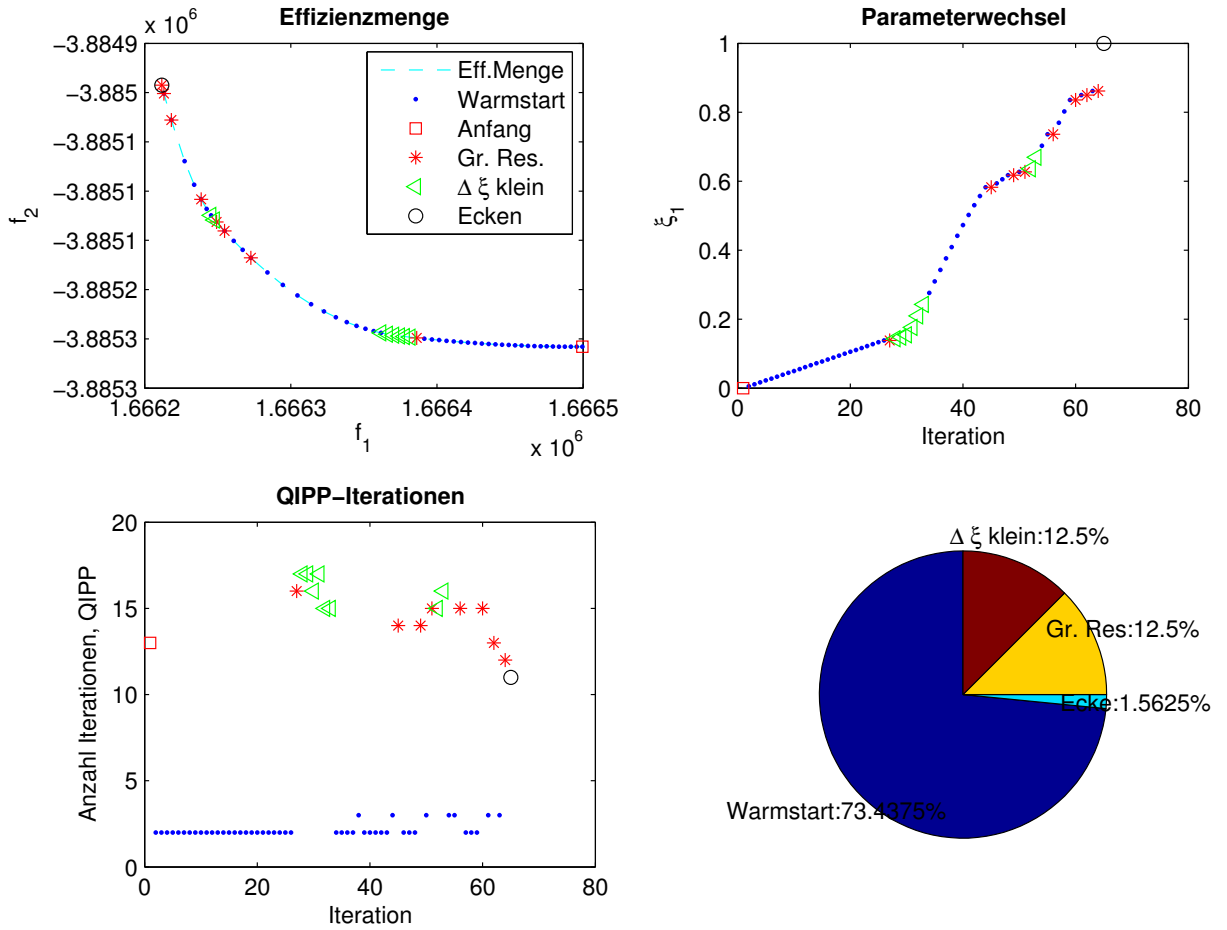


Abbildung 5.6: Zeitpunkt: Minute 29 , Zeit: 11.89 Sek., 65 Eff.Pkte
 (Beschreibung wie in Abbildung 5.4)

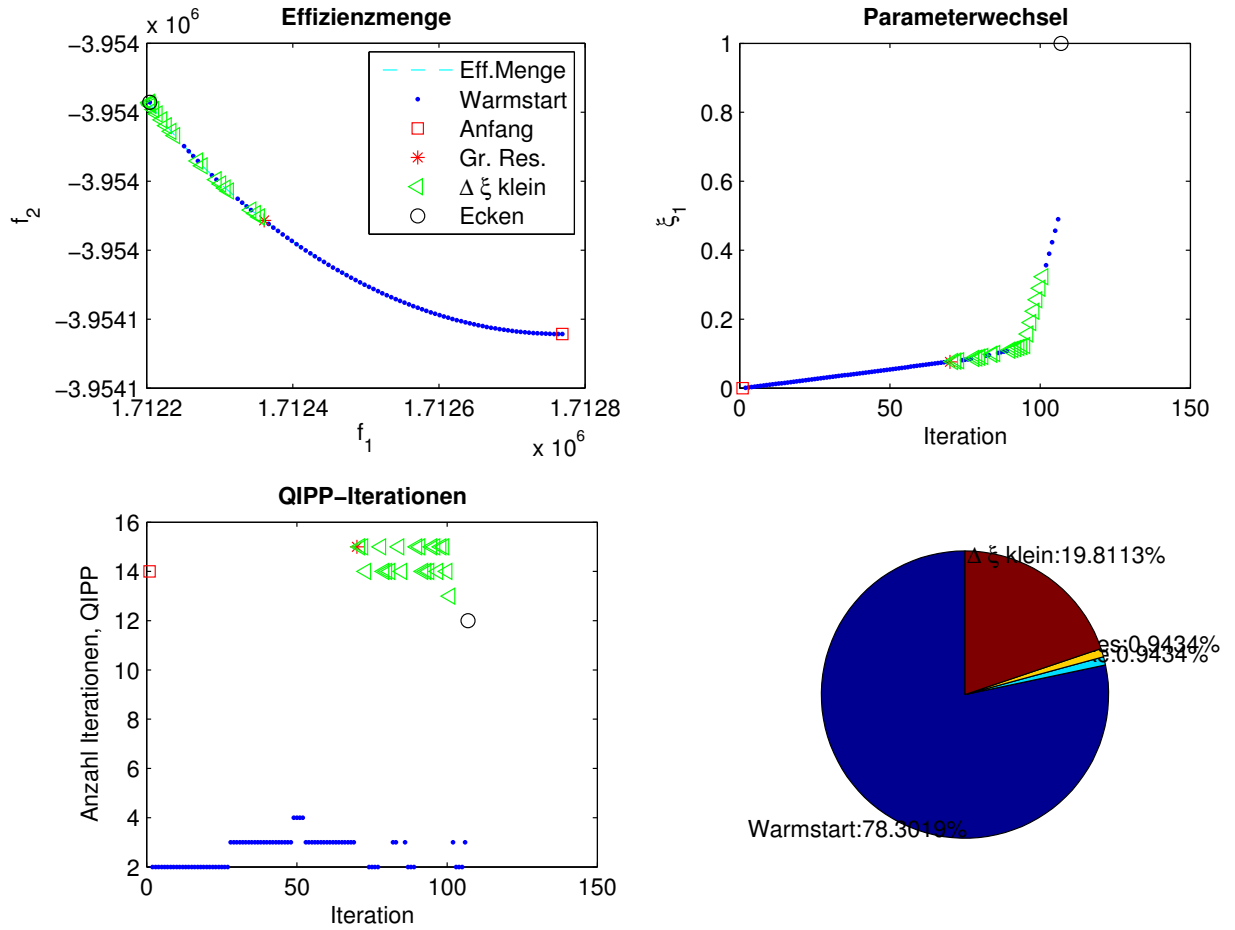


Abbildung 5.7: Zeitpunkt: Minute 144 , Zeit: 23.77 Sek., 107 Eff.pkte
(Beschreibung wie in Abbildung 5.4)

Kapitel 6

Zusammenfassung und Ausblick

In dieser Arbeit wurde ein neuartiger Pfadverfolgungsalgorithmus vorgestellt, der die Lösungsmenge konvexer, parametrisierter Optimierungsprobleme approximiert. Im Gegensatz zu vergleichbaren Verfahren werden τ -perturbierte KKT-Systeme zu Berechnung herangezogen. Basierend auf dem Prinzip Innerer-Punkte Verfahren werden zuerst lösungnahe Punkte im relativen Inneren der Restriktionsmenge ermittelt, wo die entsprechenden Systeme besser konditioniert sind.

Mit Hilfe einfachster, differentialgeometrischer Betrachtungen konnten singuläre Stellen bzw. Knicke der Lösungsmenge analysiert werden. Ferner wurden zwei Schrittweitenmethoden vorgestellt, die adaptive Parameterwechsel für möglichst äquidistante Lösungspunktabstände ermöglichen.

Angewandt auf ausgesuchte Probleme der multikriteriellen Optimierung, die wiederum parametrisierten Optimierungsprobleme entsprachen, konnte der entwickelte Pfadverfolgungsalgorithmus für die Approximation der Effizienzmenge genutzt werden. Aufgrund der speziellen Problemart ließen sich weitere Verbesserungen am Algorithmus vornehmen. Für bikriterielle, konvex-quadratische Optimierungsprobleme wurde letztendlich das MATLAB-Programm MCQIPP implementiert.

Wie in der Einführung zu Anfangs erwähnt, liefert MCQIPP kein ausgefeiltes Werkzeug. Viele Subroutinen, wie die Erkennung von Knicken und Geraden wurden nicht vollständig getestet und untersucht. Die Einstellung der Programmparameter kann durch Testläufe optimiert werden. Anstelle der tangentialen Schrittweitenbestimmung sollte eine zirkuläre umgesetzt werden.

Die Erweiterung auf einen Pfadverfolgungsalgorithmus 2. Ordnung ist vielversprechend, wenn man dahingehend die Effektivität der Mehrotra-Verfahren bei IP-Methoden mit anderen Verfahren 1. Ordnung vergleicht. Die Berechnung zweiter Ableitungen ist durch nochmaliges Anwenden des Satzes über implizite Funktionen kein großer Aufwand. Zudem werden Informationen zweiter Ordnung ebenfalls für die zirkuläre Schrittweitebestimmung benötigt.

Eine Erweiterung von MCQIPP für allgemein, nichtlineare Probleme ist möglich. Entsprechende Innere-Punkte Verfahren, wie z.B. von Byrd, Hribar und Nocedal [ByHrNo97], existieren. Einige Resultate, in denen die strikte Komplementarität vorausgesetzt wurde, fallen jedoch weg.

Die entwickelte Theorie unterstützt auch k -kriterielle Optimierung mit $k > 2$. Anstelle von reellen Parametern treten jetzt $(k - 1)$ -dimensionale, mit denen eine Schrittweitenkontrolle nicht im bisherigen Sinne möglich ist. Veränderungen am Parametervektor verursachen verschiedene rechte Seiten in den zu lösenden Gleichungssystemen, die wieder neu berechnet werden müssen - außer bei Vektorverlängerungen durch Skalare. In diesem Fall liegen Richtungsableitungen vor, die mit Schrittweitenmultiplikatoren versehen sind. Der eingesetzte Multiplikator übernimmt bei festgelegter Richtung die Funktion des Parameters, was wieder mit der entwickelten Theorie behandelt werden kann.

Eine große Schwierigkeit liegt allerdings darin, die verschiedenen Richtungsableitungen in mehreren Dimensionen zu koordinieren. Eine Möglichkeit bestände darin, die Einheitsvektoren als Richtung zu nutzen und die eingeschlagene Richtung bis zum Rand des Parameterraums zu verfolgen. Danach kehrte man zum Anfangspunkt zurück und benutzte den nächsten Einheitsvektor für die Gewinnung eines Warmstartpunktes in einer anderen Richtung und begänne von dort wieder mit dem ersten Einheitsvektor. Es entstände eine Rekursion, die eine Schar von Kurven über die Lösungsmenge legte. Das Potenzial der Warmstartpunkte nutzte man dabei jedoch nicht voll aus, da die verschiedenen Richtungen nur in den Anfangspunkten ansetzen. Adaptive Netze wären eine Alternative, die diesem Nachteil entgegen wirkte. Die Verhinderung von mehrfachen Überdeckungen verlangte jedoch eine Kommunikationsmöglichkeit zwischen den Knoten, die voraussichtlich nicht kontrollierbar ist.

Im Rahmen der multikriteriellen Optimierung könnten weitere Ansätze untersucht werden. Anstelle der Skalarisierungsmethoden bieten sich auch die Ansätze von Das und Dennis [DasDen96] oder die Methode der Beschränkungen (*engl.*: ε -Constraint

Method) [GoeNeh90], [Mie99] als Varianten parametrischer Betrachtungen an. Beide hätten ein stabileres Sensitivitätsverhalten bei Parameterwechseln zum Vorteil, was wiederum der Erzeugung adaptiver Netze zugute käme. Gäbe man sich ein Gitter im entsprechenden Parameterraum vor, ließen sich die einzelnen Knotenpunkte mit Hilfe von Warmstartstrategien ermitteln. Die nötigen Richtungsableitungen ergäben sich aus den Kanten des Netzes.

Anhang A

Grundlagen der Differentialgeometrie

Satz A.0.1 (Satz über implizite Funktionen) *Es sei $U \subset \mathbb{R}^n \times \mathbb{R}^k$ eine offene Menge und $f : U \rightarrow \mathbb{R}^m$, $(x, y) \mapsto f(x, y)$, eine C^1 -Abbildung mit der Nullstelle $(\hat{x}, \hat{y}) \in U$. Das partielle Differential $D_y f(\hat{x}, \hat{y})$ sei eine reguläre Matrix.*

Dann gibt es Umgebungen $U' \subset \mathbb{R}^n$ von \hat{x} und $U'' \subset \mathbb{R}^k$ von \hat{y} , wie auch eine C^1 -Abbildung $g : U' \rightarrow U''$, so dass die Funktionswerte von g innerhalb der Umgebung $U' \times U''$ der Nullstellenmenge von f entsprechen, d.h. explizit:

$$f(x, y) = 0, \quad (x, y) \in U' \times U'' \quad \Leftrightarrow \quad y = g(x), \quad x \in U'. \quad (\text{A.0.1})$$

(siehe [Koe93] oder [For84a])

A.1 Parametrisierte Kurven

Definition A.1.1 *Eine **parametrisierte Kurve** in \mathbb{R}^n ist eine stetige Abbildung*

$$\varphi : I \rightarrow \mathbb{R}^n, \quad t \mapsto \varphi(t) = (\varphi_1(t), \varphi_2(t), \dots, \varphi_n(t)),$$

*auf dem Intervall $I \subset \mathbb{R}$. Die Kurve φ heißt **differenzierbar**, wenn alle φ_k , $k = 1, \dots, n$ differenzierbar sind. Für $t \in I$ heißt*

$$\varphi'(t) := (\varphi'_1(t), \varphi'_2(t), \dots, \varphi'_n(t))$$

*der **Tangentenvektor** zur Kurve φ zum Parameter $t \in I$.*

*Eine differenzierbare Kurve heißt **regulär**, wenn $\varphi'(t) \neq 0$ für alle $t \in I$ gilt. Ein Parameterwert $t \in I$ heißt **singulär**, wenn $\varphi'(t) = 0$ gilt.*

Definition A.1.2 Bei der **Bogenlänge** einer differenzierbaren Kurve $\varphi : [a, b] \rightarrow \mathbb{R}^n$ handelt es sich um das Integral

$$L_a^b(\varphi) = \int_a^b \|\varphi'(t)\| dt.$$

Definition A.1.3 Für eine reguläre Kurve $\varphi \in \mathcal{C}^2(I, \mathbb{R}^n)$ und einer Stelle $t \in I$ heißt

$$\kappa(t) := \frac{\sqrt{\|\varphi''(t)\|^2 \|\varphi'(t)\|^2 - \langle \varphi''(t), \varphi'(t) \rangle}}{\|\varphi'(t)\|^3}$$

die **Krümmung** von φ an der Stelle t .

Satz A.1.4 Jede reguläre Kurve $\varphi : I \rightarrow \mathbb{R}^2$ mit konstanter Krümmung $\kappa > 0$ ist ein Kreis.

Definition A.1.5 Es sei $\phi : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$ eine zweimal differenzierbare und reguläre, parametrisierte Kurve. Im Punkt $t \in I$ existiere die Krümmung $\kappa = \kappa(t)$. Die parametrisierte Kurve

$$\gamma : [0, 2\pi r] \rightarrow \mathbb{R}^2, \quad \xi \mapsto m_0 + r \cdot \begin{pmatrix} \cos\left(\frac{\xi}{r}\right) \\ \sin\left(\frac{\xi}{r}\right) \end{pmatrix}$$

mit dem Radius $r = \frac{1}{\kappa}$ und dem Mittelpunkt $m_0 = \phi(t) + r \cdot n(t)$ heißt **Schmiegekreis** oder **Krümmungskreis** der Kurve ϕ im Punkt t . Dabei ist $n(t)$ der Normaleneinheitsvektor bezüglich $\phi'(t)$ gegen den Uhrzeigersinn gedreht.

A.2 Flächen und Mannigfaltigkeiten

Definition A.2.1 Ein **parametrisierte, k -dimensionale Fläche** ist eine \mathcal{C}^1 -Abbildung

$$\varphi : U \rightarrow \mathbb{R}^n, \quad k \leq n.$$

auf einer offenen Menge $U \subset \mathbb{R}^k$.

Eine k -dimensionale Fläche ist **regulär** falls die Vektoren der partiellen Ableitungen $\left\{ \frac{\partial \phi_i}{\partial t} \right\}_{i=1, \dots, k}$ linear unabhängig sind bzw. $\text{rang}(D\varphi(t)) = k$ für alle $t \in U$ gilt. In diesem Fall heißt φ **Immersion**.

Ist φ im Parameter $t \in U$ keine Immersion, so heißt φ **singulär** im Punkt t .

Der Begriff der parametrisierten Fläche kann zum Begriff der Mannigfaltigkeit bzw. der Untermannigfaltigkeit verallgemeinert werden.

Definition A.2.2 Eine Menge $M \subset \mathbb{R}^n$ heißt *k -dimensionale, differenzierbare Untermannigfaltigkeit*, wenn zu jedem Punkt $a \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$ und stetig-differenzierbare Abbildung

$$F : U \rightarrow \mathbb{R}^{n-k} \quad (\text{A.2.2})$$

gibt, so dass

1. $M \cap U = \{x \in U : F(x) = 0\}$,
2. Die Funktionalmatrix $\left\{ \frac{\partial F}{\partial x} \right\}_{i=1, \dots, n-k}$ hat Rang $n - k$.

Insbesondere entspricht jede reguläre k -dimensionale Fläche $\varphi : M \rightarrow V$ eine k -dimensionale Untermannigfaltigkeit, sofern φ **diffeomorph** auf V abbildet, d.h. φ ist \mathcal{C}^1 -Abbildung, bez. V bijektiv und die Umkehrabbildung φ^{-1} ist ebenfalls \mathcal{C}^1 -Abbildung.

Satz A.2.3 Eine Teilmenge $M \subset \mathbb{R}^n$ ist genau dann eine k -dimensionale, differenzierbare Untermannigfaltigkeit, falls zu jedem Punkt $a \in M$ eine offene Umgebung $V \subset M$, eine offene Teilmenge $U \subset \mathbb{R}^k$ und eine Immersion

$$\varphi : U \rightarrow \mathbb{R}^n, \quad (\text{A.2.3})$$

die T homöomorph, d.h. φ ist bijektiv und stetig sowie φ^{-1} ist stetig, auf V abbildet.

Beweis: siehe [For84], Kapitel 14, Satz 4. □

Definition A.2.4 Eine $(n - 1)$ -dimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^n$ bzw. eine parametrisierte Flächen $\varphi : U \rightarrow \mathbb{R}^n$, $U \subset \mathbb{R}^{n-1}$ wird als **Hyperfläche** bezeichnet.

Anstelle der Ableitung tritt der Begriff der Tangentialebene.

Definition A.2.5 *Es sei $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit. Ein Vektor $v \in \mathbb{R}^n$ heißt **Tangentenvektor** an M im Punkt $a \in M$, wenn eine stetig-differenzierbare Kurve*

$$\gamma :]-\varepsilon, \varepsilon[\rightarrow M, \quad (\varepsilon > 0), \quad (\text{A.2.4})$$

existiert mit $\gamma(0) = a$ und $\gamma'(0) = v$.

Die Menge der Tangentenvektoren an M in a heißt $T_a M$.

Für eine k -dimensionale Untermannigfaltigkeit, lässt sich ferner zeigen, dass $T_a M$ ein k -dimensionaler Untervektorraum des \mathbb{R}^n ist, wobei die Vektoren der partiellen Ableitungen $\frac{\partial \varphi}{\partial u_1}(c), \dots, \frac{\partial \varphi}{\partial u_k}(c)$ der Parameterdarstellung φ mit $\varphi(c) = a$ eine Basis von $T_a M$ bilden.

Entsprechend kann auch ein Normalenraum einer Untermannigfaltigkeit definiert werden.

Definition A.2.6 *Sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit. Ein Vektor $v \in \mathbb{R}^n$ heißt **Normalenvektor** von M im Punkt $a \in M$, wenn er senkrecht auf $T_a M$ steht, d.h.*

$$\langle v, w \rangle = 0 \quad \text{für alle } w \in T_a M. \quad (\text{A.2.5})$$

Die Menge aller Normalenvektoren von M in a wird mit $N_a M$ bezeichnet.

Die Menge $N_a M$ bildet einen $(n - k)$ -dimensionalen Untervektorraum des \mathbb{R}^n und die Gradienten $\nabla f_1(a), \dots, \nabla f_{n-k}(a)$ der Funktionen aus Definition A.2.2 bilden eine Basis.

Wieder lassen sich die Definitionen der Tangentialebene $T_a M$ und des Normalenraums $N_a M$ auf reguläre, parametrisierte Flächen und Kurven im \mathbb{R}^2 übertragen.

Definition A.2.7 *Eine reguläre Hyperfläche heißt **konvex**, falls M für jeden Punkt $a \in M$ auf derselben Seite der Tangentialfläche $T_a M$ liegt.*

Anhang B

Ausgaben zu numerischen Auswertungen

```
>> auswertung_both(Table4, Table_naiv)
*****
* 1) Auswertung der Daten aus mcqipp4 = MCQIPP
*****
*****
*****
*****
*****
*****
Anzahl Probleme
*****
*****
Anzahl Zeitpunkte          :1435
Anzahl unloesbarer Probleme :123
Anzahl Sonderfaelle        :4
Anzahl loesbarer Probleme   :1312
*****
*****
*****
Zeitmessungen
*****
*****
Gesamtzeit in Sekunden      :28769.4729
Gesamtzeit in Stunden       :7.9915
durchsch. Sek./Zeitpkt.    :21.928
min Sek./Zeitpkt.          :6.81
max Sek./Zeitpkt.          :33.3598
```

```

*****
*****          Effizienzpunkte          *****
*****
Anzahl Eff.pkte, Gesamt      :129527
durchsch. Anz./Zeitpkt.     :98.7248
min Anz./Zeitpkt.           :29
max Anz./Zeitpkt.           :136

```

```

*****
*****          QIPP-Iterationen          *****
*****
Anzahl Eff.pkte, Gesamt      :797308
durchsch. Anz./Zeitpkt.     :607.7043
min Anz./Zeitpkt.           :192
max Anz./Zeitpkt.           :1086

```

```

*****
*****          Neustart-Statistik          *****
*****
Neustart wg. Anfang, xi=[0 1] :1312
Neustart wg. Residuum         :9723
Neustart wg. Schrittweite     :32005
Neustart wg. Ecken            :663
Warmstart                     :9723

```

```

-----
=                               :129527
-----
-----

```

```

*****
* 2) Auswertung der Daten aus mcqipp_naiv = NAIVE STRATEGIE
*****
*****
*****
*****          Anzahl Probleme          *****
*****
Anzahl Zeitpunkte          :1435
Anzahl unloesbarer Probleme :119
Anzahl loesbarer Probleme  :1316

*****
*****          Zeitmessungen          *****
*****
Gesamtzeit in Sekunden     :18372.7013
Gesamtzeit in Stunden      :5.1035
durchsch. Sek./Zeitpkt.    :13.961
min Sek./Zeitpkt.         :4.514
max Sek./Zeitpkt.         :22.5424

*****
*****          Effizienzpunkte          *****
*****
Anzahl Eff.pkte, Gesamt    :121500
durchsch. Anz./Zeitpkt.    :92.3252
min Anz./Zeitpkt.         :25
max Anz./Zeitpkt.         :127

*****
*****          QIPP-Iterationen          *****
*****
Anzahl Eff.pkte, Gesamt    :1620717
durchsch. Anz./Zeitpkt.    :1231.5479
min Anz./Zeitpkt.         :297
max Anz./Zeitpkt.         :1949

```

```
>> auswertung_nbi(ttNBI_5_1439)
```

```
*****
```

```
*****          Anzahl Probleme          *****
```

```
*****
```

```
Anzahl Zeitpunkte          :1435
```

```
Anzahl unloesbarer Probleme :119
```

```
Anzahl loesbarer Probleme   :1316
```

```
*****
```

```
*****          Zeitmessungen          *****
```

```
*****
```

```
Gesamtzeit in Sekunden      :59628.1555
```

```
Gesamtzeit in Stunden       :16.5634
```

```
durchsch. Sek./Zeitpkt.     :45.3101
```

```
min Sek./Zeitpkt.          :12.9219
```

```
max Sek./Zeitpkt.          :166.7126
```

Anhang C

Technische Daten des Rechners

System Configuration: Sun Microsystems sun4u Sun Fire 880

System clock frequency: 150 MHz

Memory size: 12288 Megabytes

CPU: 7 × 900 MHz, US-III+

Memory: 24 × 512 MB Logical Bank Status, 256 DIMM Size

Literaturverzeichnis

- [AlGo99] A. Altman, J. Gondzio, Regularized Symmetric Indefinite Systems in Interior Point Methods for Linear and Quadratic Optimization, Optimization Methods and Software, 11-12, S. 275-302, 1999.
- [Al97] J. Albrecht, Vorausschauende Optimale Steuer- und Regelstrategien zur Verbesserung der Kraftwerkführung, Fortschr.-Ber. VDI Reihe 8 Nr. 616, VDI Verlag, Düsseldorf, 1997.
- [AlGe92] E. L. Allgower, K. Georg, Continuation and Path Following, Acta Numerica, S. 1-64, 1992.
- [BaLa89I] D. A. Bayer, J. C. Lagarias, The Nonlinear Geometry of Linear Programming. I. Affine and Projective Scaling Trajectories, Transactions of the AMS, 314, S. 499-526, 1998.
- [BaLa89II] D. A. Bayer, J. C. Lagarias, The Nonlinear Geometry of Linear Programming. II. Legendre Transform Coordinates and Central Trajectories, Transactions of the AMS, 314, S. 527-581, 1998.
- [BaLa90] D. A. Bayer, J. C. Lagarias, The Nonlinear Geometry of Linear Programming. III. Projective Legendre Transform Coordinates and Hilbert Geometry, Transactions of the AMS, 320, S. 193-225, 1990.
- [BazShe93] Mokhtar S. Bazaraa, Hanif D. Sherali, C. M. Shetty, Nonlinear Programming - Theory and Algorithms, Second Edition, John Wiley & Sons, Inc, 1993.

- [Ben62] J. F. Benders, Partitioning Procedures for Solving Mixed-Variables Programming Problems, *Numerische Mathematik*, 4, S. 238-252, 1962.
- [BoGiLeSa03] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, C. A. Sagastizábal, *Numerical Optimization - Theoretical and Practical Aspects*, Springer Verlag, Berlin Heidelberg, 2003.
- [BoMi92] B. Borchers, J. E. Mitchell, Using Interior Point Methods in a Branch and Bound Algorithm for Integer Programming, RPI Math Department Technical Report No. 195, Revised July 7, 1992.
- [BoSh00] J. F. Bonnans, A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Verlag, New York, 2000.
- [ByHrNo97] R. H. Byrd, M. E. Hribar, J. Nocedal, An Interior Point Algorithm for Large Scale Nonlinear Programming, *SIAM J. Optimization*, 9,4, S. 877-900, 1999.
- [CzMeWaWr99] J. Czyzyk, S. Mehrotra, M. Wagner, S. J. Wright, PCx: An Interior-Point Code for Linear Programming, *Optimization Methods and Software*, 11/12, S. 397-430, 1999.
- [Dan73] J. W. Daniel, Stability of the Solution of Definite Quadratic Programs, *Math. Programming* 5(1), S. 41-53, 1973.
- [DasDen96] I. Das und J. E. Dennis, Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems, Technical Report 96-11, Dept. of Computational and Applied Mathematics, Rice University, Houston, Texas, 1996.
- [dCar76] Manfred P. do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall Inc., Englewood New Jersey, 1976.
- [DeSchHe05] M. Dellnitz, O. Schütze, T. Hestermeyer, Covering Pareto Sets by Multilevel Subdivision Techniques, *Journal of Optimization, Theory and Applications*, 124(1), S. 113-136, 2005

- [Fia76] A. V. Fiacco, Sensitivity Analysis for Nonlinear Programming Using Penalty Methods, *Math. Programming*, 10(3), S. 287-311, 1976.
- [Fia83] A. V. Fiacco, *Introduction to Sensitivity Analysis in Nonlinear Programming*, Academic Press, Inc. London, 1983.
- [FiaMcCo90] A. V. Fiacco, G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM, Philadelphia, 1990
- [Fli01] J. Fliege, *Approximation Techniques for the Set of Efficient Points*, Habilitationsschrift, Fachbereich Mathematik, Universität Dortmund, 44221 Dortmund, Germany, Januar 2001.
- [Fli04] J. Fliege, *An Efficient Interior-Point Method for Convex Multicriteria Optimization Problems*, *Ergebnisbericht Angewandte Mathematik*, Nr. 248, Fachbereich Mathematik, Universität Dortmund, 2004.
- [FliHes02] J. Fliege, A. Hesel, *Constructing Approximations to the Efficient Set of Convex Quadratic Multiobjective Problems*, *Ergebnisberichte Angewandte Mathematik* 211, Fachbereich Mathematik, Universität Dortmund, 44221 Dortmund, Deutschland, Januar 2002. Submitted
- [For84a] O. Forster, *Analysis 3*, Wiesbaden, Vieweg, 1984.
- [For84] O. Forster, *Analysis 3*, Wiesbaden, Vieweg, 1984.
- [GeiKan02] C. Geiger, C. Kanzow, *Theorie und Numerik restringierter Optimierungsaufgaben*, Springer-Verlag Berlin, Heidelberg, 2002.
- [GerWri01a] E.M. Gertz, S.J. Wright, *Object-Oriented Software for Quadratic Programming*, Preprint ANL/MCS-P999-0901, Mathematics and Computer Science Division, Argonne National Laboratory, September 2001. Erhältlich unter <http://www.cs.wisc.edu/~swright/00QP/>.

- [GerWri01b] E.M. Gertz, S.J. Wright, OOQP User Guide, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Ill, Oktober 2001. Erhältlich unter <http://www.cs.wisc.edu/~swright/OOQP/>.
- [GoeNeh90] A. Göpfert und R. Nehse, Vektoroptimierung - Theorie, Verfahren und Anwendungen, Teubner Verlagsgesellschaft Leipzig, 1990.
- [GolVLo89] G. H. Golub und C. F. Van Loan, Matrix Computations, 2nd Edition, The Johns Hopkins University Press, 1989.
- [Gon98] J. Gondzio, Warm Start of the Primal-Dual Method Applied in the Cutting Plane Scheme, Mathematical Programming, 83, S. 125-143, 1998.
- [GonGro03] J. Gondzio und A. Grothey, Re-Optimization with the Primal-Dual Interior Point Method, SIAM Journal on Optimization, 13, S. 618-630, 2003.
- [GoMa95] J. Gondzio, M. Makowski, HOPDM - Modular Solver for LP Problems, User's Guide to Version 2.12, Working Paper WP-95-50, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1995.
- [GoVi99] J. Gondzio, J.-P. Vial, Warm Start and ε -Subgradients in Cutting Plane Scheme for Block-Angular Linear Programs, Computational Optimization and Applications, 14, S. 17-36, 1999.
- [GrTe97] C. Großmann, J. Terno, Numerik der Optimierung, Teubner Verlag, Stuttgart, 1997.
- [GuVaJo90] J. Guddat, F. Guerra Vazquez, H.Th. Jongen, Parametric Optimization: Singularities, Pathfollowing and Jumps, John Wiley & Sons und B. G. Teubner, 1990.
- [GuVaTaWe85] J. Guddat, F. Guerra Vasquez, K. Tammer, K. Wendler, Multiobjective and stochastic optimization based on parametric optimization, Akademie-Verlag, Berlin, 1985.

- [Hag79] W. W. Hager, Lipschitz Continuity for Constrained Process, SIAM J. Control and Optimization, 17, S. 321 - 338, 1997.
- [HeFlWe04] C. Heermann, J. Fliege, B. Weyers, Computational Experience with a New Adaptive Algorithm for Multicriteria Optimization, Working Paper, Universität Dortmund, Fachbereich Mathematik, 2004.
- [Hil00] C. Hillermeier, Nonlinear Multiobjective Optimization - A Generalized Homotopy Approach, Birkhäuser Verlag, 2000.
- [HSL] HSL: A Collection of FORTRAN Codes for Large scale scientific computation, 2000, siehe unter

<http://www.numerical.rl.ac.uk/hsl>
- [Jah04] J. Jahn, Vector Optimization - Theory, Applications and Extensions, Springer Verlag Berlin Heidelberg, 2004.
- [Jit81] K. Jittorntrum, Solution Point Differentiability without Strict Complementarity in Nonlinear Programming, Math. Programming Studies, North-Holland, Amsterdam, 1981.
- [JaSt04] F. Jarre, J. Stoer, Optimierung, Springer Verlag, Berlin Heidelberg, 2004.
- [Ka84] N. Karmarkar, A New Polynomial-Time Algorithm for Linear Programming, Combinatorica 4, S. 373-395, 1984.
- [Kel60] J. E. Kelley, The Cutting Plane Method for Solving Convex Programs, Journal of the SIAM, 8, S. 703-712, 1960.
- [Koe93] K. Königsberger, Analysis 2, Springer-Verlag Berlin-Heidelberg, 1993.
- [KoMiYo89I] M. Kojima, S. Mizuno, A. Yoshise, A Polynomial-Time Algorithm for a Class of Linear Complementary Problems, Mathematical Programming, 44, S. 1-26, 1989.
- [KoMiYo89II] M. Kojima, S. Mizuno, A. Yoshise, A Primal-Dual Interior Point Algorithm for Linear Programming, Process in Mathematical Programming: Interior-Point and Related Methods,

- N. Megiddo, Kapitel 2, S. 29-47, Springer-Verlag, New York, 1989.
- [Luc89] Dinh T. Luc, Theory of Vector Optimization, Springer Verlag, 1989.
- [Mal87] K. Malanowski, Stability of Solutions to Convex Problems of Optimization, Springer-Verlag, Berlin Heidelberg, 1987.
- [Mang94] Olvi L. Mangasarian, Nonlinear Programming, Classics in Applied Mathematics (corrected republication of work first published 1969), SIAM, Philadelphia, 1994.
- [McL80] L. McLinden, An Analogue of Moreau's Proximation Theorem, with Applications to the Nonlinear Complementarity Problem, Pacific Journal of Mathematics, 88, S. 101-161, 1980.
- [Meg89] N. Megiddo, Pathways to the Optimal Set in Linear Programming, Process in Mathematical Programming: Interior-Point and Related Methods, , Kapitel 8, S. 131-158, Springer-Verlag, New York, 1998.
- [Mie99] Kaisa Miettinen, Nonlinear Multiobjective Optimization, Kluwer Academic Publishers, 1999.
- [MiToYe93] S. Mizuno, M. Todd, Y. Ye, On Adaptive Step Primal-Dual Interior-Point Algorithms for Linear Programming, Mathematics of Operations Research, 18, S. 964-981, 1993.
- [MoAd89I] R. D. C. Monteiro, I. Adler, Interior Path-Following Primal-Dual Algorithms. Part I: Linear Programming, Mathematical Programming, 44, S. 27-41, 1989.
- [MoAd89II] R. D. C. Monteiro, I. Adler, Interior Path-Following Primal-Dual Algorithms. Part II: Convex Quadratic Programming, Mathematical Programming, 44, S. 43-66, 1989.
- [Mo04] K. Moritzen, private Mitteilung.
- [NeNe94] Y. Nesterov, A. Nemirovskii, Interior-Point Polynomial Algorithms in Convex Programming, SIAM Studies in Applied Mathematics. SIAM, Philadelphia, 1994.

- [RaHaWa91] J. Rakowska, R. T. Haftka, L. T. Watson, Tracing the Efficient Curve for Multiobjective Control-Structure Optimization, *Computing Systems in Engeneering*, 2(6), S. 271-281, 1991.
- [Rhei74] Werner C. Rheinboldt, *Methods for Solving Systems of Nonlinear Equations*, SIAM, Philadelphia 1974.
- [Rhei86] Werner C. Rheinboldt, *Numerical Analysis of Parametrized Nonlinear Equations*, John Wiley & Sons, 1986.
- [Rob74] S. M. Robinson, Perturbed Kuhn-Tucker Points and Rates of Convergence for a Class of Nonlinear-Programming Algorithms, *Mathematical Programming* 7, S. 1 - 16, 1974.
- [SawNakTan85] Y. Sawaragi, H. Nakayama, T. Tanino, *Theory of Multiobjective Optimization*, Academic Press, 1985.
- [Scha67] V. E. Shamanskii, Ob odnoj modifikatsii metody Newtona, *Ukrainskii Matematichnii Zhurnal*, 19, S. 133-138, 1967.
- [Son89] G. Sonnevend, An "Analytic Center" for Polyhedrons and New Classes of Global Algorithms for Linear (Smooth, Convex) Programming, *Sytem Modelling and Optimization: Proceedings of the 12th IFIP Conference Budapest, Ungarn, September 1985*, A.Prekopa, J. Szelezsan, B. Strazicky, *Lecture Notes in Control and Information Sciences* 84, S. 866-876, Springer-Verlag, Berlin, 1986.
- [Sto92] Tz. E. Stoyanov, Marching along surface/surface intersection curves with adaptive step length, *Computer Aided Geometric Design* 9, S. 485 - 489, 1992.
- [Tra64] J. F. Traub, *Iterative Methods for the Solution of Equations*, Prentice-Hall, Englewood Cliffs, N.J, 1964.
- [Van97] R. J. Vanderbei, *LOQO User's Manual - Version 3.10*, Statistics and Operations Research, Technical Report No. SOR-97-08, 1997.
- [Wol98] L. A. Wolsey, *Integer Programming*, John Wiley & Sons, New York, 1998.

- [Wri97] S. J. Wright, Primal-Dual Interior Point Methods, SIAM Publications, Philadelphia, 1997.
- [WuAn99] Shin-Ting Wu, Lenimar N. Andrade, Marching Along a Regular Surface/Surface Intersection with Circular Steps, Computer Aided Geometric Design, 16(1999), 249 - 268.
- [YilTod01] E. Alper Yildırım, Michael Todd, Sensitivity Analysis in Linear Programming and Semidefinite Programming using Interior-Point Method. Mathematical Programming, 90:229-261, 2001.
- [YilWri02] E. Alper Yildırım und S. J. Wright, Warm-start Strategies in Interior-Point Methods for Linear Programming, SIAM Journal on Optimization, 12(3):782-810, 2002.
- [Zha99] Yin Zhang, User's Guide to LIPSOL - Linear-Programming Interior Point Solvers v0.4, Optimization Methods and Software, Vol. 11-12, 385-396, 1999.